ORNL/SPR-2021/1950

US Department of Energy, Office of Science High Performance Computing Facility Operational Assessment 2020
Oak Ridge Leadership Computing Facility

April 2021

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY
Not for public release – Distribution is Limited

ORNL/SPR-2021/1950

Oak Ridge Leadership Computing Facility

US DEPARTMENT OF ENERGY, OFFICE OF SCIENCE HIGH PERFORMANCE COMPUTING FACILITY OPERATIONAL ASSESSMENT 2020 OAK RIDGE LEADERSHIP COMPUTING FACILITY

Subil Abraham
J. Paul Abston
Ryan M. Adamson
Valentine Anantharaj
Aaron Barlow
Ashley D. Barker
Katie L. Bethea
Kevin S. Bivens
Adam Carlyle
Daniel Dietz
Christopher Fuson
Rachel Harken
Benjamin Hernandez Arreguin
Jason J. Hill
Jason Kincl
Ryan Landfield
Don Maxwell
Veronica Melesse Vergara
Bronson Messer
J. Robert Michael
Ross Miller
Sheila Moore
Sarp Oral
Thomas Papatheodore
Ryan Prout
Sherry Ray
William Renaud
Mallikarjun Shankar
Woong Shin
Scott Simmerman
Kevin G. Thach
Aristeidis Tsaris
Georgia Tourassi
Coury Turczyn
Joseph Voss
Justin L. Whitt
Junqi Yin

April 2021

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831-6283
managed by
UT-BATTELLE, LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive Summary

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
March 2021

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; 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.

In calendar year (CY) 2020, the Summit supercomputer continued to perform quadrillions of calculations every second as part of the OLCF's traditional user and allocation programs, but the facility was also critical in the fight against COVID-19. With uninterrupted operations, the OLCF was able to pivot its work and join forces with other US federal agencies, industry, and academic leaders as part of the COVID-19 HPC Consortium. The consortium is a unique private-public effort spearheaded by the White House Office of Science and Technology Policy, the US Department of Energy (DOE), and IBM to bring together federal government, industry, and academic leaders who are volunteering compute time and resources on their world-class machines. Overall, the facility has provided more than 2 million Summit node hours to 23 projects investigating different aspects of the novel coronavirus, including identifying therapeutics and developing new theories on the body's inflammatory response to the virus. Through the Innovative and Novel Computational Impact on Theory and Experiment (INCITE), Advanced Scientific Computing Research Leadership Computing Challenge (ALCC), Director's Discretionary (DD) programs and the COVID-19 HPC Consortium, the OLCF was able to meet the challenge presented by the onset of the global COVID-19 pandemic and provided the platform for the two applications that won the Gordon Bell prize and the special Gordon Bell prize for COVID-19 research. In 2020, the OLCF also saw great growth in its newly launched Quantum Computing User Program (QCUP). The QCUP was launched in 2019 in conjunction with the Quantum Computing Institute, and it aims connect researchers with quantum computing resources through ORNL's vendor partners: currently IBM, D-Wave Systems, and Rigetti Computing. The OLCF purchased and deployed the follow-on system of Rhea named Andes and continued operations for Summit and the Alpine center-wide file system while transitioning their working environments to ensure the safe and continuous operation of the OLCF as the onset of a global pandemic changed the operating paradigm.

On March 16, 2020, the threats posed by the pandemic began impacting the operations of the OLCF, and this would continue through the entirety of CY 2020. On-site operations were significantly reduced to help the local area "flatten the curve." The OLCF was uniquely positioned by the nature and skillset of the work and successfully transitioned to a majority of the operations staff working from home. At the beginning of this reduced on-site presence it was not known that this posture would last for the remainder of the CY; the staff rose to the new challenges presented by the distributed team environment and continued to exceed the needs of the OLCF users.

CY 2020 was filled with outstanding results and accomplishments, including a very high rating from users on overall satisfaction for the seventh consecutive year; a tremendous number of core-hours delivered to researchers on Summit; and the successful delivery of the allocation split of roughly 60, 30, and 10% of core-hours offered for the INCITE, ALCC, and DD programs, respectively (Section 2). The OLCF hosted its second Quantum Computing User Forum in 2020 with a huge turn-out of more than 200 participants, illustrating the strong interest and engagement in QCUP resources. 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 2020 OLCF metric targets and the associated results. More information can be found in Section 2 for each OLCF resource.

Table ES.1. 2020 OLCF metric summary.

Metric description	**CY 2020 target**	**CY 2020 actual**
Overall OLCF score on the user survey will be 3.5/5.0 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 h) by either resolving the problem or informing the user of how the problem will be resolved.	80%	90%
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	30%	43.42%
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.	85%	99.59%
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.	80%	98.73%
OVERALL AVAILABILITY (FILESYSTEMS): *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.	85%	98.93%
OA, archive storage: Sustain availability to users measured as a percentage of maximum possible.	90%	98.33%

Metric description	CY 2020 target	CY 2020 actual
SCHEDULED AVAILABILITY (FILESYSTEMS): *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.	90%	99.43%
SA, High Performance Storage System: Sustain scheduled availability to users measured as a percentage of maximum possible scheduled.	95%	99.92%

The scientific accomplishments of OLCF users are a strong indication of long-term operational success with publications this year in such notable journals as *Science*, *Nature*, *Nature Communications, Journal of the American Chemical Society*, *Cell*, *ACS Nano*, *Nano Letters* and *Molecular Biology and Evolution*. Crucial domain-specific discoveries facilitated by resources at the OLCF are described in Section 8 of .

The OLCF supported scientific accomplishments for a broad community of researchers in 2020 from traditional modeling and simulation projects to studies exploiting artificial intelligence (AI), machine learning (ML), and big data techniques. Six teams conducted research on Summit as part of competition for the prestigious Gordon Bell Prize, an award presented 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 were nominated for a newly formed Gordon Bell Prize for HPC-based COVID-19 research. Four of the six finalists for the Gordon Bell Prize used Summit, and two of the four finalists for the special prize for HPC-based COVID-19 research used Summit.

OLCF systems continue to support data science, deep learning, and AI work. For example, an ORNL team used AI to find patterns in neutron scattering data that can lead to an understanding of the physics inside quantum or complex magnetic materials. They trained an ANN to successfully interpret data from a neutron scattering experiment performed at ORNL's Spallation Neutron Source (SNS). The team trained the network by feeding it data from neutron scattering simulations performed on systems at the OLCF, including the center's decommissioned Cray XK7 Titan. The team is now training deeper neural networks on Summit to further understand glass-like quantum materials.

COVID-19 research was a key focus for many user teams. For example, a multi-institutional team led by a group of ORNL investigators has been studying various SARS-CoV-2 protein targets, including the virus's main protease that is key in its replication process. This project began with the launch of a collaborative effort with NVIDIA and Scripps Research to create and run a new version of the AutoDock-GPU molecular modeling code, optimizing it for high-throughput molecular docking simulations on the Summit supercomputer. For each separate docking simulation, the team generated 20 possible poses, or configurations, showing how each synthetically producible compound might fit inside the viral protein structure's binding pocket. To accurately model the protein, the team used crystallographic structures from neutron scattering experiments performed at ORNL's High Flux Isotope Reactor and SNS. After generating the poses, the team calculated features of the poses to apply different ML models to reevaluate them and to better determine whether each compound was a strong binder to the main protease. To analyze the massive amount of data—1.3 TB per large-scale calculation—the team implemented a "virtual laboratory" on Summit to sort, manipulate, and join data pieces together.

Using Summit, ORNL researchers discovered new patterns in systems responsible for COVID-19 symptoms. The researchers analyzed genes from cells in the lung fluid of nine COVID-19 patients compared with 40 control patients and analyzed population-scale gene expression data—17,000 samples from uninfected individuals—to see which genes were normally co-expressed (i.e., turned on or off at the same time). The computational analyses suggest that genes related to one of the body's systems responsible for lowering blood pressure—the bradykinin system—appear to be excessively "turned on" in the lung fluid cells of those with the virus. Excessive bradykinin leads to leaky blood vessels, allowing fluid to build up in the body's soft tissues. The power of Summit enabled 2.5 billion correlation calculations in 1 week rather than spending months doing them on a desktop computer. The results helped the researchers understand the normal regulatory circuits and relationships for the genes of interest.

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 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.

Communication was important in 2020 because some traditional methods for delivering information to the public and stakeholders were unavailable due to the COVID-19 pandemic. On-site tours of the OLCF; attendance at conferences, workshops, and seminars; and the annual OLCF User Meeting are traditionally avenues to showcase the work performed and capabilities of the OLCF. Although the annual User Meeting occurred, it was virtual and was held early in the pandemic response. The OLCF's communications and outreach team undertook a significant effort to provide methods for the on-site tour "experience" but in a virtual format by using a virtual facility tour via the Matterport software system.

ES.1.2 Communication with the User Community

The OLCF's communications with users are tailored to the objectives of relating science results to the larger 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 (Sections 1.4.5–1.4.6).

The impact of OLCF communications is assessed as part of an annual user survey. Communications was one of the highest rated areas of satisfaction on the survey. The mean rating for users' overall satisfaction with OLCF communications was 4.6. Ninety-seven percent of respondents rated their overall satisfaction with communications from the OLCF as “satisfied” or “very satisfied.” 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, and
- targeted training events (i.e., GPU hackathons or tutorials).

ES.2 SUMMARY OF 2020 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 2020. The 2020 metrics, target values, and actual results as of December 31, 2020 are noted throughout this report and are summarized in Section 2. The OLCF exceeded all agreed-upon metric targets.

ES.3 RESPONSES TO RECOMMENDATIONS FROM THE 2018 OPERATIONAL ASSESSMENT REVIEW

The OLCF conducted the review of the 2020 Operational Assessment via Zoom on April 21–23, 2020. The review committee had no recommendations for the OLCF. Details are provided in Appendix A.

User Support Results

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
March 2021

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 OLCF supported 1,570 users and 479 projects, including Quantum, in calendar year (CY) 2020. The OLCF continued to leverage an established user support model 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 2020 was 4.6. Overall ratings for the OLCF were positive; 97% 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 2020, with 90% of tickets being 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.

1.1 USER SUPPORT RESULTS SUMMARY

The OLCF's user support model comprises customer support interfaces, including user satisfaction surveys, formal problem-resolution mechanisms, user assistance (UA) 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 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 continuous improvement at the OLCF, users are sent surveys that solicit their feedback regarding support services and their experience as facility users. The 2020 survey was launched on October 12, 2020 and remained open for participation through November 16, 2020. The survey was sent to 1,260 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 included the Exascale Computing Project (ECP) projects, who logged into an OLCF system between January 1, 2020, and September 30, 2020. Although the ECP allocations come from the OLCF DD program, those responses were tracked separately from the DD responses. OLCF staff members were excluded from participation. A total of 688 users completed the survey for an overall response rate of 55%. The full results of the 2020 survey can be found on the OLCF website.[1]

[1] https://www.olcf.ornl.gov/wp-content/uploads/2021/02/2020-olcf-user-survey-report.pdf

The effectiveness of the processes for supporting customers, resolving problems, and conducting outreach are partly measured by the key survey responses for user support in Table 1.1.

Table 1.1. Key survey responses for 2020.

Survey question	2019 target	2019 actual	2020 target	2020 actual
Overall OLCF satisfaction score on the user survey	3.5/5.0	4.5/5.0	3.5/5.0	4.6/5.0
Overall satisfaction with support received (UA, accounts, INCITE liaisons, advanced data/workflow liaisons)	3.5/5.0	4.6/5.0	3.5/5.0	4.6/5.0
Overall satisfaction with the website	3.5/5.0	4.4/5.0	3.5/5.0	4.4/5.0
Overall satisfaction with communications	3.5/5.0	4.6/5.0	3.5/5.0	4.6/5.0
Overall satisfaction with OLCF support services	3.5/5.0	4.5/5.0	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 2019 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 2020 survey.
Percentage of user problems addressed within 3 business days	80%	90%	80%	90%

1.2 USER SUPPORT METRICS

The OLCF exceeded all user support metrics for 2020. The OLCF metric targets and actual results by CY for user support are shown in Table 1.2.

Table 1.2. OLCF user support summary for metric targets and CY results.

Survey area	CY 2019 target	CY 2019 actual	CY 2020 target	CY 2020 actual
Overall OLCF satisfaction rating	3.5/5.0	4.5/5.0	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

1.2.1 Overall Satisfaction Rating for the Facility

Users were asked to rate their satisfaction on a five-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 US Department of Energy (DOE) and the OLCF program manager who defined 3.5/5.0 as satisfactory. Overall ratings for the OLCF were positive with 97% of users responding that they were satisfied or very satisfied with the OLCF overall, which was a 3% improvement over 2019.

Key indicators from the survey, including overall satisfaction, are shown in Table 1.3. They are summarized and presented by program respondents. The data show that satisfaction among all allocation programs is similar for the four key satisfaction indicators.

Table 1.3. Satisfaction rates by program type for key indicators.

Indicator	Mean	Program			
		INCITE	ALCC	DD	ECP
Overall satisfaction with the OLCF	4.5/5.0	4.6/5.0	4.6/5.0	4.6/5.0	4.5/5.0
Overall satisfaction with support staff	4.5/5.0	4.6/5.0	4.6/5.0	4.6/5.0	4.5/5.0
Overall satisfaction with compute resources	4.6/5.0	4.6/5.0	4.6/5.0	4.6/5.0	4.5/5.0
Overall satisfaction with data resources	4.5/5.0	4.6/5.0	4.5/5.0	4.6/5.0	4.5/5.0

1.2.2 Average Rating across All User Support Questions

The calculated mean of answers to the user support services specific questions on the 2020 survey was 4.5/5.0, indicating that the OLCF exceeded the 2020 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 overall satisfaction with support received from the wide variety of OLCF services available. Most respondents (94%) were either satisfied or very satisfied with support received from UA, accounts, INCITE Scientific Computing Liaisons, and Advanced Data/Workflow Liaisons. A thematic analysis of open-ended comments identified computing power/hardware/HPC resources/performance (39% of respondents) and user support/staff (31% of respondents) as the most valued OLCF qualities.

The following bullets include a few selected 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: (1) leadership-class computational resources that enable high risk and impactful science, (2) collaboration and support from our INCITE liaison has been extremely helpful, and (3) timeliness and quality of problem resolution is excellent."
- "OLCF provides access to computing resources of a magnitude that we would otherwise not have access to. Utilizing Summit allows us to run problems that we cannot run on any other system. The OLCF staff make using the systems straightforward and the support we receive has always been excellent."
- "I really like the OLCF support team. I have encountered a couple issues with the Summit system. The support team was very helpful and proactive in assisting me."
- "We're pleased with the user guides, which are extremely helpful for navigating the complexities of the system (especially the complexity of using jsrun and resource sets)."

1.2.3 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 2019 and 2020 surveys. Although no questions scored below satisfactory on the 2020 survey, the survey was thoroughly reviewed, and the OLCF identified areas in which the Facility could improve or new services that could be provided to enhance user experience at the OLCF. In some cases, the issue was already addressed or a solution is in the works and forthcoming shortly. The areas for improvement are identified as follows.

- Deploy a Jupyter Notebook server.
  - This service was implemented in December 2020 and is discussed in more detail in Section 1.4.2.3. There are plans to further evolve the service in 2021.

- Create the ability to search the OLCF ticketing system.
  - The OLCF migrated to a new ticketing system on January 1, 2021, and this feature is something it is planning to offer in 2021 through the new ticketing system.

- Make the account renewal process easier.
  - Improving this process is on the myOLCF roadmap. There are plans to not only make user experience when renewing accounts less burdensome but also to automate parts of the renewal process, making it less burdensome on the staff who process these renewal requests.

- Send automated notifications when projects get close to utilizing their entire allocation.
  - This request was identified as a new feature that will be offered in 2021. Plans are underway now to begin sending these reports.

- Improve software documentation.
  - Areas of software documentation that can be improved were identified, and the team is working now to implement the improvements.

1.2.4 Assessing the Effectiveness of the OLCF User Survey

1.2.4.1 Revisions to the 2020 OLCF User Survey

The OLCF worked with the Assessment and Evaluation team at Oak Ridge Associated Universities (ORAU) and substantially revised the user survey in 2019. The revisions reduced the total number of possible items in the survey by approximately 56%, from 181 to 80 items. These revisions were largely carried over to the 2020 survey. Additional revisions were made this year to account for systems that had been sunset and were no longer applicable to the user experience, as well as to add an overall satisfaction item to measure satisfaction with OLCF support staff. Before the 2019 survey revisions, the average survey completion time as measured and reported by the survey software interface was 18.5 min, and the average user responded to 90 items. In 2019, the average response time was 9.8 min, and the typical respondent answered 38 items. For the current survey, the average response time was 6.74 min, and the typical respondent answered 38 items. The team believes that shortening the survey was the primary reason for the much-improved response rate in 2019 and 2020. However, relevant feedback continues to be gained from the survey, and it is still very useful for helping determine what is needed to improve user experience at the OLCF.

1.2.4.2 Data Collection

The survey sampling frame was constituted by first collecting the names of individuals who had logged into an OLCF system between January 1, 2020 and September 30, 2020. OLCF staff, OLCF vendors, and individuals with invalid email addresses were then removed from the list. The survey was hosted online beginning on October 12, 2020 and remained open for completion through November 16, 2020. Overall, this process resulted in a sampling frame with 1,260 OLCF users. A total of 688 users responded, resulting in a response rate of 55%, which was a 9% increase in participation from 2019 and over a 20%

increase from 2018. The survey was advertised on the OLCF website and in the weekly email sent to all users. Survey responses were tracked daily to assess the effectiveness of the various communication methods. The number of responses increased after every targeted notification, but the results show that other efforts, such as including the notice in the weekly communication, also contributed to the survey response rate.

The overall experience using the OLCF was approximately evenly split between three categories with the largest proportion (38%) having more than 2 years of experience using the facility (Table 1.4)

Table 1.4. User survey participation.

OLCF users	2019 survey	2020 survey
Total number of respondents (total percentage responding to survey)	578 (46%)	688 (55%)
New users (OLCF for user <1 year)	32%	30%
OLCF user for 1–2 years	21%	32%
OLCF user for >2 years	47%	38%

1.2.4.3 Statistical Analysis of the Results

The survey collected feedback about user needs, preferences, and experience with the OLCF and its support capabilities. Attitudes and opinions on the performance, availability, and possible improvements of OLCF resources and services were also solicited. ORAU provided the OLCF with a written report that included the results and a summary of the findings. In the ORAU document, the "Findings" section presents results summarized numerically that report responded levels of satisfaction. This is followed by a summary of the verbal open-ended comments from individuals who indicated via the scaled reply that they were dissatisfied with a resource or service. (Not all dissatisfied individuals supplied open-ended comments.)

The survey assessed satisfaction with OLCF resources and services by using a five-point scale, ranging from "very dissatisfied" (1) to "very satisfied" (5). These responses were closed-ended and summarized by using frequency distributions, proportions, means, and standard deviations. The proportion of respondents indicating either a 4 ("satisfied") or 5 ("very satisfied") on an item was also typically reported as "%Sat" to provide a summary measure. This measure was also used to indicate the relative satisfaction with resources and services within categories. Respondents who were "very dissatisfied" or "dissatisfied" with OLCF resources and services were asked to provide comments explaining their dissatisfaction. To better understand the types of OLCF users and how needs and preferences varied, closed-ended responses were frequently sorted by principal investigator (PI) status and by project allocation.

All open-ended responses were examined by using categorical content analysis with complete thoughts (CTs) in responses as the unit of analysis. (The percentages of response categories might add up to more than 100% when respondents provided multiple CTs in a response.) CTs were sorted into categories for the purposes of counting, comparisons, and other forms of analysis. CTs were response text that could stand alone as a meaningful reply to survey questions. CTs were not limited to any specific grammatical unit and could vary from a single word to a phrase, sentence fragment, or complete sentence.

Some response content categories were derived a priori from survey questions or OLCF website categories (e.g., Data Management). Other categories were developed inductively through an iterative process of grouping and regrouping similar content units (e.g., Containers or Training and Tutorials). Subcategories were elaborated as new relevant concepts or useful distinctions were identified and were then organized within main categories of closely related concepts. These are used to the extent possible with variations as needed to accommodate differences in the focus of specific questions and year-to-year differences in users' specific and technical responses.

Table 1.5 displays responses for four of the overall satisfaction categories divided by allocation program. As Table 1.5 illustrates, the metrics are very comparable across all five primary allocation programs, and the variations are statistically insignificant.

Table 1.5. Statistical analysis of survey results.

Survey area	INCITE			ALCC			DD			ECP		
	Mean	**Variance**	**Standard deviation**	**Mean**	**Variance**	**Standard deviation**	**Mean**	**Variance**	**Standard deviation**	**Mean**	**Variance**	**Standard deviation**
Overall satisfaction with the OLCF	4.6	0.31	0.56	4.6	0.34	0.58	4.6	0.28	0.53	4.5	0.40	0.63
Overall satisfaction with compute resources	4.6	0.35	0.59	4.6	0.38	0.62	4.6	0.31	0.56	4.5	0.44	0.66
Overall satisfaction with data resources	4.6	0.36	0.60	4.5	0.42	0.65	4.6	0.35	0.59	4.5	0.42	0.65
Overall satisfaction with staff support	4.6	0.40	0.63	4.6	0.38	0.62	4.6	0.34	0.58	4.5	0.46	0.68

1.3 PROBLEM RESOLUTION METRICS

The following operational assessment review metrics were used for problem resolution:

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

- 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.

1.3.1 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 in which it is limited in its ability to resolve due to factors beyond the facility's control. In such a scenario, 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, prudent measures might be to recreate the issue; open a bug or ticket with the product vendor; provide the vendor with the necessary information about the issue; provide a workaround to the user, if possible; and track the issue to resolution with the product vendor, which might resolve the issue with a bug fix or workaround acknowledgment.

The OLCF used the Request Tracker (RT) software to track user problem reports (i.e., tickets) and ensure that response goals were met or exceeded in 2020. Users submit tickets in a variety of ways—including via email, an online request form, and phone—but email is the predominant method used. Statistics on tickets submitted, turnaround times, and other metrics are gathered from RT for reports. The OLCF issued 2,839 tickets in response to user queries for CY 2020. The center exceeded the problem-resolution metric and responded to 90% of the queries within 3 business days (Table 1.6).

Table 1.6. Problem resolution metric summary.

Survey area	CY 2019		CY 2020	
	Target	Actual	Target	Actual
Percentage of problems addressed in 3 business days	80%	90%	80%	90%
Average of problem resolution ratings	3.5/5.0	4.5/5.0	3.5/5.0	4.6/6.0

Users were asked to provide ratings for their overall satisfaction with OLCF's problem resolution and three specific aspects, and 94% of respondents were overall either satisfied or very satisfied with problem resolution. The quality of OLCF responses to reported issues was the highest rated specific aspect, whereas the usefulness of support and training documentation and the timeliness of OLCF responses to reported issues were rated slightly lower. Nine respondents reported reasons for dissatisfaction with problem resolution. The most common reason for dissatisfaction was the lack of visibility of the tickets to the user community and the quality and/or quantity of the center's software documentation. The OLCF is working on migrating its ticketing systems to a new platform. As part of that effort, the center will be able to offer users visibility into the ticketing system in the near future. More information on this effort is discussed as follows. Additionally, as already noted, specific areas of the software documentation that can be improved were identified, and the OLCF is working to implement those improvements into the documentation.

Tickets are categorized by the most common types. The top two reported categories in 2020 were related to access and running jobs (Figure 1.1), as was the case in 2019.

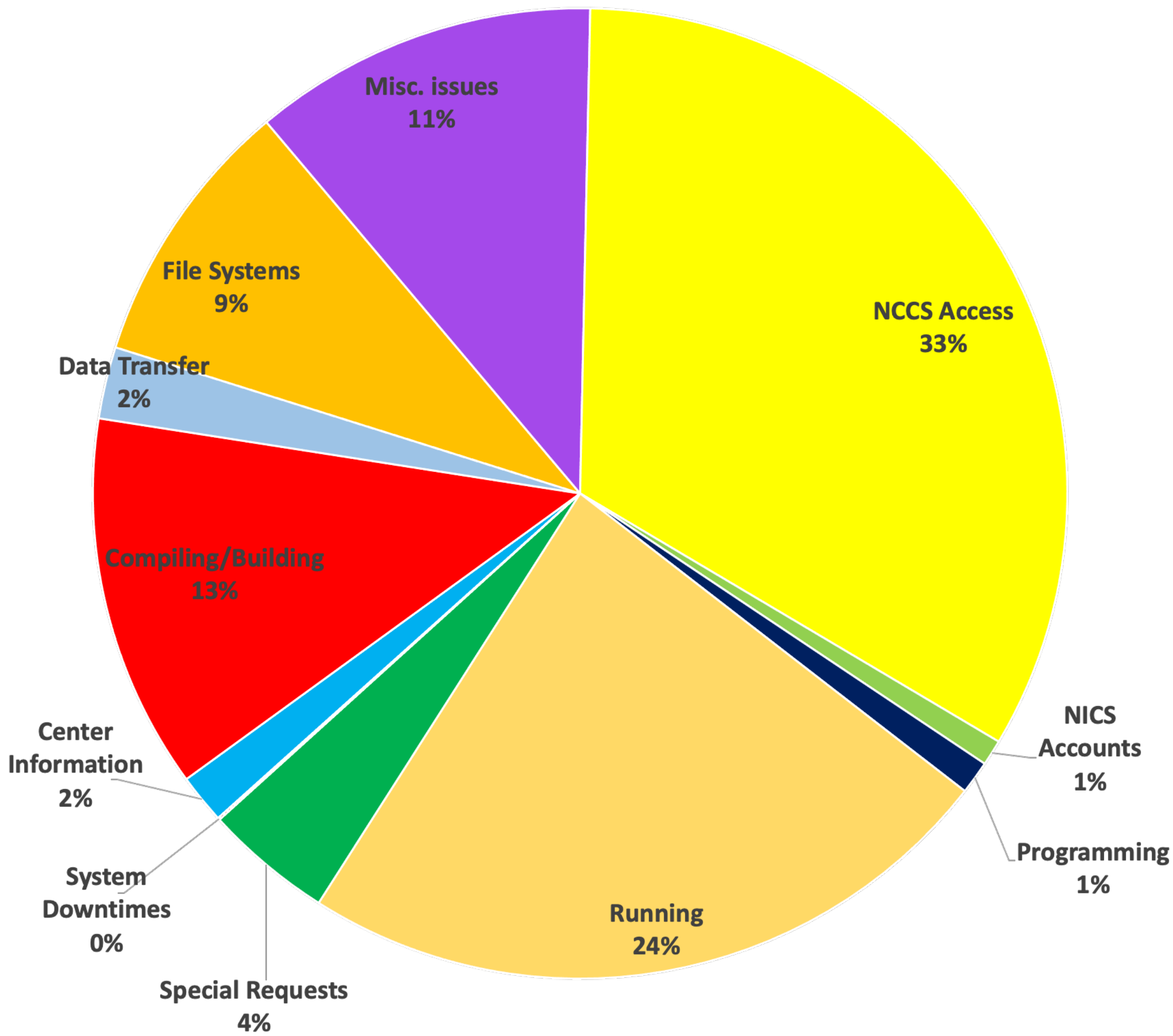


Figure 1.1. Categorization of help desk tickets.

1.4 USER SUPPORT AND OUTREACH

The Operational Assessment Report (OAR) data requested for user support and outreach include examples of in-depth collaboration between facility staff and the user community and a summary of training and outreach events conducted during this period (Appendices B–C).

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

- a user support staff that comprises account management liaisons, User Assistance and Outreach (UAO) analysts, Scientific Engagement section 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 interface with the next generation of HPC users, the external media, and the public.

1.4.1 User Support

The OLCF recognizes that users of HPC facilities have a wide range of needs that require diverse solutions, from immediate, short-term, trouble ticket-oriented support, such as assistance with debugging and optimizing code, to more in-depth support that requires total immersion in and collaboration on projects. The facility provides complementary user support vehicles that include UAO staff; liaisons in respective scientific, data, and visualization areas; and computer scientists who assist on issues surrounding the programming environments and tools. The following sections detail some of the high-level support activities during CY 2019 and the specific OLCF staff resources available to assist users.

1.4.2 User Support and Outreach

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. The following sections provide some examples of the OLCF's initiatives in 2020 that helped improve the overall user experience, albeit some were very behind the scenes.

1.4.2.1 MyOLCF

The OLCF wrote and deployed a new web-based self-service application called myOLCF,[2] which was released on October 30, 2020. The motivation for myOLCF was two-fold: to provide PIs and project members with timely, accurate data to empower decision making vis-à-vis OLCF projects and to offer more self-service tools for administering projects. To accomplish this, myOLCF makes relevant information about projects, users, account applications, and resource allocations 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.

Before myOLCF was made publicly available, cybersecurity researchers performed thousands of distinct cyberattacks on a sandboxed instance of myOLCF and identified three specific areas in which the security of the application could be improved, all of which were implemented before the production launch of the application in October 2020. Since its launch, updated versions of myOLCF with new features and user experience improvements have been deployed twice monthly on average. Documentation on myOLCF can be found on the OLCF user documentation site.[3]

In 2021, OLCF software engineers plan to deploy additional features and improvements, including:

- user-centric help ticket visibility,
- streamlined account and project application forms,
- greater compatibility on mobile devices,
- improved accessibility via accessible rich internet application best-practices, and

[2] http://my.olcf.ornl.gov
[3] https://docs.olcf.ornl.gov/services_and_applications/myolcf/overview.html

- streamlined renewal approach.

1.4.2.2 RATS CRM

The center's customer relationship management software, called the RATS Contingent Reimbursement Model (CRM), is under continuous development. In 2020, 41 versions of the software were deployed, adding many new features and improvements, such as:

- new application programming interface (API) endpoints for seamless integration with myOLCF,
- new integrations with MailChimp APIs for automated management of mailing lists,
- new "COVID" project tag and API endpoints for automated syncing of COVID-19 project information to OLCF websites, and
- redesigned operational workflows for project renewal applications.

1.4.2.3 Supporting the Launch of JupyterHub at OLCF

The OLCF launched a center-wide JupyterHub deployment in December 2020, giving users the ability to easily launch interactive, web-based environments for working with code and data. The UA team worked to test, document, train, and advertise this new service to users. Within a few weeks of announcing the service, the OLCF tallied ~60 individual users. Users can explore the OLCF's Juptyer documentation,[4] outlining the service and how to get started. A user conference was held on January 27, 2021 to discuss the service with OLCF users, and a follow-up survey was sent to those attendees to learn more about additional needs.

1.4.2.4 Moving Ticket-Based Support into Jira

The OLCF has used RT since 2003 to provide ticket-based support for various programs and for various internal request tracking purposes. The OLCF sought to replace this aging ticket system with an already-deployed Jira system to better meet the requirements of today and to allow for tighter integration with other services. For example, OLCF users have requested better visibility into their tickets via a web interface. Jira not only has such an interface but also provides an API that will be used to integrate ticket information into the new myOLCF interface.

The OLCF spent the latter half of 2020 working to gather these requirements, plan for a transition, train staff, and begin the transition from RT to Jira. Support projects (queues) were configured for the OLCF help desk (production systems) and for pre-production systems. Training sessions were held to permit support staff to learn the new interface. Documentation specific to the support staff's use of Jira was created to ensure the proper administrative handling of tickets. Included in this was information to serve as an "RT to Jira Rosetta Stone" to facilitate transition by staff who had great familiarity with RT. Additionally, effort was made to ensure the accurate collection and reporting of required DOE metrics from the new system with consideration given to how to properly support tickets that might bridge the transition from RT to Jira as the official support system. This considerable groundwork prepared the OLCF for the migration from RT to Jira, which was planned for the first weeks of 2021.

[4] https://docs.olcf.ornl.gov/services_and_applications/jupyter/index.html

1.4.2.5 Andes Transition to Operation

In late 2020, the OLCF began migrating pre- and postprocessing and data analysis user workloads from an aging cluster, Rhea, to a new cluster resource. Rhea's replacement, Andes, has 704 nodes with two AMD EPYC 7302 16 core processors and 256 GB of memory along with the nine high-memory/GPU nodes. Members of the UA group worked closely with members of the HPC Clusters group during the transition to plan a timeline that allowed a smooth user workload transfer. The timeline included staff testing, early user testing, GPU node moves, and a period during which Andes and Rhea were both available before Rhea's decommissioning. The early user testing allowed for Rhea's most active users to test for a period before transitioning Andes to production and opening to all projects with Rhea access. A 5 week overlap during which Rhea and Andes were both available was provided to all users to help provide a smooth workflow transition to the new system. Communication was very important for ensuring a smooth transition. Email notices were provided to all users in the OLCF weekly update before production availability and during the production overlap period. Targeted email notices were provided throughout the transition by directly emailing impacted user groups. For example, active Rhea users were directly notified with increasing frequency as the Rhea decommission date grew nearer. Web documentation was also used to provide details of the system, transition, and differences between both systems.

1.4.2.6 High-Memory Summit Nodes

In 2020, 54 high-memory nodes were added to Summit. The nodes have 2 TB of DDR4 memory for the POWER9 processors, 192 GB of high-bandwidth memory (HBM2) for the V100 accelerators, and 6.4 TB of nonvolatile memory. Because the available memory is higher than the standard compute node in Summit, the nodes were labeled as high-memory nodes in the batch system to allow workflows to differentiate. To help ensure that users were aware of and able to use the new nodes, the OLCF UA group notified users of the addition through the weekly user email and during the monthly user call. Details of the nodes' configuration, procedures to use, and policies were added to the Summit documentation.

1.4.2.7 Launch of the OLCF Quantum Computing User Program

In early 2020, the tremendous growth in demand for quantum computing by DOE, national labs, industry, and academia created a unique opportunity for the OLCF to establish a strategic capability to provide OLCF users with access to state-of-the-art quantum computing resource for purposes of discovery and innovation in scientific computing applications. 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. Supported research projects 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 new effort to establish the OLCF Quantum Computing User Program (QCUP) required a team effort that comprised significant contributions from staff in many different areas inside and outside the OLCF. First, the team negotiated new contracts and extensions with four quantum computing hardware vendors—IBM, Rigetti, D-Wave, and Xanadu—to provide the quantum computing resources needed to establish the user program. Each contract was a significant effort because the negotiations addressed things such as user agreements, resource availability, pricing, and policies. Additionally, the team worked with each of these vendors to establish the multistep processes required for issuing user and project accounts, monitoring allocations, enforcing user agreements, providing user support, and tracking and reporting, among other tasks. One 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 the user support model whereby QCUP users request help and report issues with the resources. Because each vendor has very specific requirements for access and support, there was no easy one-size-fits-all solution when establishing these processes. The team established a communication mechanism with users to inform them of system downtimes, changes, training events, and so on. The OLCF also launched the QCUP website[5] later in the year. The website contains background on the priorities of the QCUP program, frequently asked questions, QCUP event information, and highlighted publications. Additionally, the website seeks to provide user documentation and training materials to help approved users access and use the quantum computing resources supplied by current vendors. Additional user quick-start guide content is currently in development, and the website will be updated regularly to reflect the changing capabilities and offerings of the vendor partners.

Through the partnership between QCUP and the vendor partners, users can now participate in several quantum training events. For example, in April 2020, the OLCF hosted the Quantum Computing User Forum, which was quickly transformed into a virtual event that welcomed nearly 200 attendees from around the world. The 3-day event featured a series of technical talks that covered subjects related to quantum computing, including the development of applications, software, and simulations for quantum devices and systems.

The OLCF supported 52 QCUP projects and 117 QCUP users in 2020, and the center is already seeing publications resulting from the use of these resources.

1.4.3 Scientific Liaison Collaborations

The following sections highlight specific collaborative areas in which OLCF staff scientists partnered with INCITE and other research teams to maximize their productivity on the provided leadership-class resources.

1.4.3.1 Lightweight Task Management Prototypes

Scientific Engagement staff have developed and tested two tools, pmake and dwork, to provide task bundling functionality within ensemble HPC jobs on Summit, Rhea, and Andes. Pmake was used in production for data analysis on thousands of molecular dynamics (MD) trajectories generated by the COVID-19 project in 2020 led by Jeremy Smith (ORNL). Dwork was optimized to provide a low-latency task list during high-throughput job campaigns on Summit. Timed at 11 ms per task, dwork is a stand-alone executable that manages a task-directed acyclic graph with dynamic updates via a simple network API. Both provide a low entry barrier to projects that want to organize multiple processing steps outside the batch scheduling framework.

Publication: A. Acharya et al., "Supercomputer-Based Ensemble Docking Drug Discovery Pipeline with Application to Covid-19," *J. Chem. Inf. Model* 60, no. 12 (2020): 5,832–5,852.

Team Members Affiliated with the OLCF: D. M. Rogers, O. Henrnadez, J. V. Vermaas, J. Glaser

[5] https://www.olcf.ornl.gov/olcf-resources/compute-systems/quantum-computing-user-program/

1.4.3.2 Team Gains a 52-Fold Speedup on the Nation's Fastest Supercomputer to Accelerate COVID-19 Research

As SARS-CoV-2 tore through the world in 2020, scientists everywhere raced to understand the virus in an attempt to stop it. Almost immediately, researchers turned to the OLCF to try to help find the virus's weaknesses and information that could be useful in the development of therapeutics. One ORNL-led team studied various SARS-CoV-2 protein targets, including the virus's main protease. The team was able to harness 27,612 of Summit's NVIDIA V100 GPUs to simulate more than 1 billion compounds binding with two different structures of the main protease, with each billion-compound screening completed in less than 24 h.

With collaboration and support from OLCF scientific liaisons Jens Glaser, Josh Vermaas, and David Rogers, Ada Sedova, a biophysicist in the Molecular Biophysics Group within ORNL's Biosciences Division, and senior staff member Oscar Hernandez of ORNL's Computer Science and Mathematics Division (CSMD) worked with NVIDIA and Scripps Research to create and run a new version of the AutoDock-GPU molecular modeling code, optimizing it for high-throughput molecular docking simulations on the Summit supercomputer.

For each separate docking simulation, the team generated 20 possible poses (i.e., configurations), showing how each synthetically producible compound might fit inside the viral protein structure's binding pocket. To accurately model the protein, the team used crystallographic structures from neutron scattering experiments performed at ORNL's High Flux Isotope Reactor and Spallation Neutron Source (SNS). After generating the poses, the team had a lot of data analysis work to perform—a task that was led by Glaser.

To analyze the massive amount of data—1.3 TB per large-scale calculation—the team implemented a "virtual laboratory" on Summit to sort, manipulate, and join data pieces together. By employing the BlazingSQL platform, NVIDIA RAPIDS, the Numba compiler, and the Dask ecosystem to query data, the team reduced the results processing by orders of magnitude, achieving a 52 fold speedup for the full pipeline.

Publication: J. V. Vermaas et al., "Supercomputing Pipelines Search for Therapeutics Against COVID-19," *Computing in Science & Engineering* (2020). doi: 10.1109/MCSE.2020.3036540.

Team Members Affiliated with the OLCF: J. Vermaas, D. Rogers, J. Glaser, O. Hernandez

Awards and Recognition: Gordon Bell Prize Finalist

1.4.3.3 Simulations of Subatomic Quark Particles with Close-to-Physical Masses Using the Summit Supercomputer

Predicting the extent to which quarks interact and calculating the full distribution of quark energies and momenta have remained long-standing challenges in particle physics. Using lattice quantum chromodynamics, which represent space-time as a grid on which particles are formulated, a team that included OLCF computational scientist Bálint Joó developed a method for measuring quark interactions in their larger particles and applied it to computer simulations by using quarks with close-to-physical masses for the first time on the Summit supercomputer. The team generated snapshots of the strong-force field in a cube of space-time, weighting the snapshots to describe what the quarks were doing in the vacuum. The team then took these snapshots and simulated what would happen as quarks moved through the strong-force field. The simulations required the power of ORNL's Summit supercomputer due to the number of vacuum snapshots that the team had to generate and the number of quark propagators that needed to be calculated on them. To estimate the results at the physical quark mass, calculations needed to be performed at three different quark masses and extrapolated to the physical one. In total, the team used more than 1,000 snapshots over three different quark masses in cubes with lattices ranging from 32^3 to 64^3 points in space. The team's calculations ranged from 358 to 172 MeV, close to the experimental mass of pions, the lightest particles they were simulating. The team can apply theoretical knowledge from such simulations to experimental data to make better predictions about subatomic matter.

Publication: Bálint Joó et al., "Parton Distribution Functions from Ioffe Time Pseudodistributions from Lattice Calculations: Approaching the Physical Point," *Physical Review Letters* 125 (2020): 23,2003. doi: 10.1103/PhysRevLett.125.232003.

Team Members affiliated with the OLCF: Bálint Joó

Awards and Recognition: Best student paper at the 2020 Practice & Experience in Advanced Research Computing

1.4.3.4 Simulating Full-Scale Square Kilometre Array Data on Summit

The operational workflow for processing the full-scale low-frequency array data of the Square Kilometre Array (SKA) Phase 1 was simulated on Summit. SKA will begin construction soon. Once completed, it will be the world's largest radio telescope and one of the world's largest data generators. The results show that the workflow is ready to process a key SKA science case, an Epoch of Reionization observation. The analysis also helps reveal critical design factors for the next-generation radio telescopes. A typical 6 h observation was simulated with an end-to-end data processing and imaging pipeline. The workflow was deployed and run on 4,560 compute nodes (all available nodes at the time of job submission) by using 27,360 NVIDIA V100 GPUs to generate 2.6 PB of data. This is the first time that radio astronomical data have been processed at this scale.

Publications: R. Wang, R. Tobar, et al. "Processing Full-Scale Square Kilometre Array Data on the Summit Supercomputer," *The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20)*, November 15–20, 2020, Atlanta, Georgia. doi: 10.1109/SC41405.2020.00006

Team Members Affiliated with the OLCF: Norbert Podhorszki, Valentine Anantharaj, Scott Klasky

Awards and Recognition: Gordon Bell Prize Finalist

1.4.3.5 A Baseline for Global Weather and Climate Simulations at Unprecedented 1 km Resolution

This INCITE project generated a 4-month reference dataset in support of model development. The data are also deemed essential for simulating the impact of future global satellite-observing systems for weather and climate. The earth's atmosphere was simulated with 1.4 km grid spacing. The simulation produces a realistic global mean atmospheric circulation and an improved stratosphere through the resolved feedback of deep convection and topography with the direct indication of weather extremes. Global energy redistribution is similar in simulations with 1.4 km grid spacing with explicit representation of deep convection and simulations with 9 km grid spacing with deep convection parametrization.

Publications: Authors: N. Wedi, P. Dueben, V. Anantharaj, P. Bauer, et al., "A Baseline for Global Weather and Climate Simulations at 1 km Resolution," *Journal of Advances in Modeling Earth Systems* 12, no. 11 (2020): e2020MS002192. doi: https://doi.org/10.1029/2020MS002192.

Team Members Affiliated with the OLCF: Valentine Anantharaj

Awards and Recognitions: HPCwire Readers' Choice Award in Physical Sciences

1.4.4 OLCF User Group and Executive Board

All PIs and users on approved OLCF projects are members of the OLCF User Group (OUG) and remain so for 3 years following the conclusion of their OLCF project. The OUG meets monthly via a webinar called the OLCF User Conference Call. 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 and training events. The webinars also include a brief tutorial. A wide range of topics were covered during the year. These included a user experience talk related to SARS-CoV-2 work and talks on several software tools, including Kokkos, Cuda 11, and Open CE. Other talks highlighted tools available at or coming soon to the OLCF, including myOLCF, the Job Step Viewer, the Constellation DOI service, and the SLATE container orchestration service. Ten OLCF User Conference Calls were held in 2020 with a total of 502 attendees.

The OUG is represented by a nine to 10 member OUG Executive Board. This board typically meets monthly just after the User Conference Call for an in-depth discussion with OLCF staff to provide feedback and guidance on training topics and the facility's resources and policies. OUG Executive Board terms are 3 years and are staggered so that three new members are elected each year to replace three outgoing members. An outgoing chair will remain on the board as an ex officio member for an additional year if he/she is at the end of his/her 3-year term, hence the possibility of a 10 member board. During the year, Sarat Sreepathi replaced Mike Zingale as board chair. The board elected Eric Nielsen as vice chair, and he will become chair for the 2021–2022 term. At the 2020 user meeting, Emily Belli, Scott Callaghan, André Walker-Loud, and Mike Zingale began 3-year terms, replacing three members whose terms had ended along with an additional board member who stepped down. These terms will conclude at the 2023 user meeting. Information about the OUG and the executive board is available on the OLCF website.[6]

The OLCF hosted its annual user meeting on June 3–4, 2020. Unlike previous user meetings that featured a blend of on-site and online attendees, this meeting was entirely online/virtual due to restrictions related to the COVID-19 pandemic. The first day of the meeting focused on new user training for Summit. Presentations included an overview of best practices for OLCF users, as well as overviews of Summit and its various components, including file systems, data transfers, programming tools, the batch scheduler and job launcher, and burst buffers. Additional topics included a discussion of Python best practices and machine learning (ML)/deep learning (DL) tools. Most presentations on the first day were led by members of the UA and Outreach Group; a member from the Technology Integration Group provided one presentation and an Advanced Data and Workflows staff member led another.

The second day of the user meeting began with OLCF staff members providing OLCF operational highlights and a review of the user survey followed by overviews of several OLCF initiatives, such as the Andes data analysis cluster, the myOLCF web application, the Quantum User Program, and the exascale system. Later that day, Yongqiang Cheng, a neutron scattering scientist at ORNL, presented "Neutron

[6] https://www.olcf.ornl.gov/about-olcf/oug/

Scattering and High Performance Computing: The Synergy of Two User Facilities.” Finally, the meeting ended with an open forum for users and OLCF staff.

Attendance was up from last year’s meeting, which had 140 attendees, with 181 attendees on the first day and 199 attendees on the second day 2. Slides and videos from the different presentations were uploaded and made available via the user meeting’s event page.[7]

1.4.5 Training, Education, and Workshops

The OLCF training program serves to educate its user community with general HPC training and special topics needed to fully leverage the facility’s cutting-edge HPC resources. To do so, 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 users through the OLCF training archive.[8] In 2020, the COVID-19 pandemic caused the OLCF to adapt its training program from mostly in-person events to a fully virtual format starting in March. Because the OLCF’s user base is mostly remote, online participation was typically offered already for short (e.g., <1 day), mostly informational training events (e.g., webinars). However, multiday events and/or any training with hands-on components, especially hackathons, were typically given in person and so had to be reimagined in a fully virtual format within a relatively short time.

Each type of virtual event required its own unique set of online tools with the goal of providing participants with the same value that they previously received at in-person events. Simple webinars that are purely informational can be broadcast on most web conferencing platforms. However, hackathons, for example, require a platform that offers robust breakout room solutions that allow each team to have its own virtual room to discuss its specific application during hacking sessions but that still allow organizers to quickly bring all participants back into a common room for discussions and daily group updates. Events with hands-on components (e.g., tutorials, hackathons) also often require a separate collaborative tool (e.g., Slack) to facilitate help with coding questions/issues. It is difficult to try to make sense of a code snippet or error message in a tiny, unformatted Zoom or WebEx chat window. Some platforms, such as Microsoft Teams, are trying to bridge the gap between these different kinds of tools, but currently—and certainly at the start of 2020—multiple online tools are still needed.

Despite the challenges of adapting to a virtual training program, in 2020 the OLCF facilitated or collaborated on five multiday virtual GPU hackathons; one multiday virtual OpenMP hackathon; a two-day OpenMP target offload training; a nine-part CUDA training series; a three-part OpenACC training series; a four-day Kokkos bootcamp and training; a two-day OLCF user meeting, including a Summit new user training; a two-day introduction to DL for scientific applications webinar; a scaling up DL applications on Summit webinar; a Frontier Center of Excellence (COE) workshop; a Tuning and Analysis Utilities (TAU) performance analysis training; a NVIDIA Nsight Compute webinar; a NVIDIA Nsight systems webinar; a jsrun tutorial; 10 OLCF user conference calls/webinars; 12 Interoperable Design of Extreme-Scale Applications Software—ECP (IDEAS-ECP) webinars; and an INCITE proposal writing seminar. Appendix B provides a complete summary of these events. Some notable events are highlighted in Sections 1.4.5.1–1.4.5.5.

1.4.5.1 CUDA Training Series

The OLCF, the National Energy Research Scientific Computing Center (NERSC), and NVIDIA co-organized and delivered a nine-part CUDA training series that ran from January through September 2020.

[7] https://www.olcf.ornl.gov/calendar/2020-olcf-user-meeting/
[8] https://docs.olcf.ornl.gov/training/training_archive.html

Each module consisted of a 1-hour presentation followed by a hands-on session during which participants worked through example exercises meant to reinforce the material covered in the presentation. The topics covered were:

- introduction to CUDA C/C++,
- CUDA shared memory,
- fundamental CUDA optimization (part 1),
- fundamental CUDA optimization (part 2),
- atomics, reductions, and warp shuffle,
- managed memory,
- CUDA concurrency,
- GPU performance analysis, and
- cooperative groups.

The series[9] was intended to provide a foundational understanding of GPU programming on the CUDA platform for the OLCF and NERSC user base. In total, 333 unique users attended this series, and the slides, recordings, and hands-on materials are all available to future users through the OLCF training archive.

1.4.5.2 Introduction to Deep Learning for Scientific Applications

On October 7–8, 2020, the OLCF and IBM delivered a two-day training titled, "Introduction to Deep Learning for Scientific Applications,"[10] to show current Summit users ways to use ML and DL in traditional modeling and simulation scientific applications. The presentations started with the basics of ML on the first day and worked toward integrating DL surrogates into an application by the end of the second day. The material covered in the presentations was supplemented by hands-on labs during which participants could test their understanding of the material. Participation was capped at 100 attendees to ensure that the virtual hands-on sessions could be adequately managed, but all materials are available for future users to learn from.

1.4.5.3 Summit New User Training

On the first day of the OLCF user meeting[11] (June 3, 2020), the OLCF delivered a Summit new user training intended to deliver the information that new OLCF users need to get started running on Summit. Presentations covered topics such as OLCF best practices, system architecture, programming environments, batch scheduler, job launcher, ML/DL on Summit, and Summit tips and tricks. In total, 111 OLCF users participated in the event, and all slides and recordings are available for future users on the OLCF training archive.

1.4.5.4 2020 OLCF GPU Hackathon Series

In 2020, the OLCF and its partners successfully adapted the annual OLCF GPU hackathon series[12] to a fully virtual format. As with the in-person format, these hackathons are multiday coding events during which teams of developers prepare their own applications to run on GPUs or focus on optimizing their applications that currently run on GPUs. Each team typically comprises three or more developers of a specific application and two mentors, identified by the organizers, with GPU programming expertise.

---

[9] https://www.olcf.ornl.gov/cuda-training-series/
[10] https://www.olcf.ornl.gov/calendar/introduction-to-deep-learning-for-scientific-applications/
[11] https://www.olcf.ornl.gov/calendar/2020-olcf-user-meeting/
[12] https://docs.olcf.ornl.gov/training/olcf_gpu_hackathons.html

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 that range from workstations and local clusters to world-class HPC systems, such as the OLCF's Summit.

Historically, these hackathons were structured as 5 day (Monday–Friday) in-person events during which each team of application developers and mentors sat at its own round table in one large conference room. This structure allowed teams to hack away on their own codes but also to interact (e.g., ask questions, give advice) with members and mentors from other teams when needed. It also allowed teams to share updates, progress, and roadblocks with all hackathon participants, which allowed participants to share knowledge by helping each other with similar problems.

To recreate this environment in a virtual setting, the organizers used a combination of Zoom and Slack. Zoom was used as the main web conferencing tool due to its breakout room capabilities; one Zoom session was set up in which the main room was used for presentations and group discussions, and separate breakout rooms were created for each of the individual teams for screen sharing and verbal communication. The Slack workspace was used for communication between all participants and also provided individual team channels for chatting, sharing code snippets, and so on within teams. To combat "Zoom fatigue," the organizers chose to schedule four-day events instead of five, with day 1 on a Monday and days 2–4 on the following Monday through Wednesday. Day 1 was typically used for an overview of the compute system, team introductions, and code profiling. Then, the week in between was ideally used for further profiling, testing, and making necessary changes to the team's strategy for the following week. Finally, days 2–4 were considered the "hackathon proper" during which the teams executed their strategies through long coding sessions.

With this new format, the OLCF and its partners delivered five virtual GPU hackathons with the following organizations:

- San Diego Supercomputer Center (April 27, May 11–13),
- Princeton University (June 1, 8–10),
- NERSC (July 7, 13–15),
- Brookhaven National Laboratory (August 10, 17–19), and
- ORNL (October 19, 26–28).

These sites hosted 40 application teams, bringing dedicated GPU programming expertise to nearly 300 developers from 62 organizations. During the events, teams worked on applications that spanned a wide range of scientific domains, such as astrophysics, climate modeling, combustion, computational fluid dynamics (CFD), ML, MD, plasma physics, and quantum mechanics.

Overall, the organizers found that the virtual format provided teams with essentially the same value that they received at in-person events with the exception of personal networking. One team that saw success this year was the Center for Mobility with Vertical Lift (MOVE) CFD team from Rensselaer Polytechnic Institute. The team works with the parallel, hierarchic, adaptive, stabilized transient analysis (PHASTA) code. The team took a mini-version of PHASTA to the OLCF-hosted GPU hackathon and, with help from NVIDIA mentors at the hackathon, ported the mini-application to the OLCF's Ascent test bed system and achieved a tenfold speedup.

"This was definitely a successful hackathon for our team, as we finally took the first crack at porting a legacy CPU code to GPUs," said Abhishek Chopra, team lead and PhD student at Rensselaer Polytechnic Institute's MOVE. "These efforts have motivated us to keep working on this to make a significant difference in the speedup of the production code."

Another team, led by aerospace researcher Ioannis Nompelis at the University of Minnesota (UMN), brought a mini-application, MY3D, meant to mimic the US3D CFD solver, which is used to simulate hypersonic flows around geometrically complex shapes. The team, AeroUMN, observed a 40× speedup of MY3D when using all 42 cores on an Ascent node relative to the serial version. After porting a single-GPU version of MY3D, the team found an additional 50% speedup relative to the fully parallelized CPU version.

"The great thing about these hackathons is that they are so well-run, and there is also investment from NVIDIA, who embed their knowledgeable personnel into the hackathons so that teams can ask questions," Nompelis said. "Not only do we typically not get opportunities to access these kinds of large-scale GPU-endowed systems, but we also do not have access to mentors of this caliber."

In addition to providing application teams with development support—many of which are OLCF users—these events benefit the OLCF by introducing participants to the facility and resources, nurturing relationships with 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.

In 2020, the OLCF and its partners demonstrated that these GPU hackathons can be carried out in an on-site or virtual format with similar results for the application teams, and they even established some new best practices (e.g., the 1 + 3-day format was beneficial for most participants) for both formats moving forward. With this understanding, the OLCF is confident that this series will continue to benefit the GPU programming community during and after these uncertain times.

1.4.5.5 Frontier Center of Excellence Workshop

On September 3, 2020, the OLCF in conjunction with the Frontier Center of Excellence (COE) hosted a Frontier COE Workshop[13] for the Center for Accelerated Application Readiness (CAAR) and ECP personnel who were current users of the COE's testbed systems Poplar (CAAR) and Tulip (ECP). During the workshop, members of the Frontier COE gave updates on the latest system hardware and software, and selected CAAR and ECP teams gave lessons learned talks regarding their experiences with porting and running on the systems. The workshop was restricted to participants covered under existing CORAL-2 NDAs. In total, 192 participants attended the event.

1.4.5.6 Virtual Workshop on GPU Accelerated Data Analytics: From Drug Discovery to Radiation Physics

On October 15, 2020, OLCF staff in conjunction with CSMD and Biosciences Division staff delivered a virtual workshop[14] on GPU accelerated data analytics to show early adopters of the NVIDIA RAPIDS and BlazingSQL ecosystem how they can accelerate the analysis of large datasets on Summit. Workshop contents included an overview of NVIDIA RAPIDS, BlazingSQL, and the OLCF's JupyterHub service. A hands-on session demonstrated the OLCF's interactive data analytics capabilities to handle terabyte-sized datasets on Summit. Participants of this workshop included INCITE, ALCC, ECP, CAAR, and COVID-19 HPC Consortium teams and OLCF staff.

[13] https://www.olcf.ornl.gov/frontier-coe-workshop/
[14] https://www.olcf.ornl.gov/workshop-on-gpu-accelerated-data-analytics-virtual-lab-on-summit-from-drug-discovery-to-radiation-physics/

1.4.6 Training and Outreach Activities for Future Members of the HPC Community and the General Public

1.4.6.1 SC20 Contributions

The International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20) is one of 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 SC20 are given as follows.

As part of conference's Students@SC program, on November 5, 2020, OLCF staff led a half-day session titled, "Hands-On with the Summit Supercomputer."[15,16] This event included an overview of the leadership computing facilities, Summit's architecture, and how research teams write programs for and run workloads on Summit. Participants were also given access to the OLCF's Ascent training system—an 18 node, stand-alone compute system with the same node architecture and programming environment as Summit—to work through challenge problems of their choosing. These hands-on exercises covered a wide range of subject matter (e.g., programming models, batch scheduler, shell environments) and degrees of difficulty to accommodate the attendees' various levels of experience. The session was attended by 62 undergraduate and graduate student conference attendees.

As part of the conference's general tutorial program, on November 9, 2020, OLCF staff co-led a training titled, "Using the New SPEC HPC2021 Scientific Application Benchmark Suite,"[17] to show participants "how to leverage SPEC benchmarks for performance evaluation, tuning of system parameters, comparison of systems (e.g., for procurement), results interpretation and publication, and its power measurement capabilities." During the event, presenters provided demonstrations and hands-on guidance, focusing on the newly released SPEC HPC2021 benchmark. The OLCF's Ascent training system was used for the hands-on components.

Another tutorial co-delivered by OLCF staff as part of the conference's general tutorial program on November 10, 2020 was titled, "Better Scientific Software."[18] This training was meant to provide information on software practices, processes, and tools with the goals of improving the productivity of CSE software developers and increasing the sustainability of software artifacts. Topics included agile methodologies, collaboration via Git, scientific software design, testing, and refactoring, continuous integration (CI), and reproducibility. During the event, participants were given hands-on homework to reinforce the material covered in the presentations.

1.4.6.2 CARLA 2020—GPU Accelerated Computing with OpenACC

The Latin America High Performance Computing Conference (CARLA) is an international conference aimed at providing a forum to foster the growth and strength of the HPC community in Latin America through the exchange and dissemination of new ideas, techniques, and research in HPC and its applications areas. As part of CARLA 2020, on September 17, OLCF staff contributed a two-hour tutorial titled, "GPU Accelerated Computing with OpenACC,"[19] to introduce participants to GPU computing via the OpenACC programming model. During the event, attendees were given access to OLCF's Ascent training system to test their understanding of the material covered in the presentations.

---

[15] https://www.olcf.ornl.gov/sc20-hands-on-with-summit/
[16] https://github.com/olcf/SC20_HandsOn_with_Summit
[17] https://sc20.supercomputing.org/presentation/?id=tut134&sess=sess254
[18] https://sc20.supercomputing.org/presentation/?id=tut146&sess=sess275
[19] https://www.olcf.ornl.gov/carla2020-gpu-tutorial/

1.4.6.3 TAPIA 2020—Easy-Peasy GPU Programming Using OpenMP

The Association for Computing Machinery (ACM) Richard Tapia Celebration of Diversity in Computing Conference (TAPIA) is the premier venue to acknowledge, promote and celebrate diversity in computing. As part of TAPIA 2020, on September 16, OLCF staff contributed a 75 min tutorial titled, "Easy-Peasy GPU Programming Using OpenMP."[20] The event provided an introduction to GPU hardware, accelerator programming models, and the specifics of OpenMP. Attendees were given access to OLCF's Ascent training system for the hands-on components of the training.

1.4.6.4 Outreach

The OLCF Outreach team works to engage new and next-generation users and showcases OLCF research through strategic communication activities, such as highlights, fact sheets, posters, snapshots, the OLCF website, and center publications (Appendix C). Like their colleagues across the facility and across ORNL, the OLCF Outreach team members were able to pivot in 2020 in response to the COVID-19 pandemic. They successfully continued to deliver news and features at a regular cadence while working remotely. One significant highlight of the team's work this year involved COVID-19-related news and information, updating the OLCF homepage with real-time COVID-19 user metrics, delivering high-profile COVID-19 user-related news stories, and coordinating media coverage of OLCF users who are studying the virus and potential treatments.

In March 2020, the team released a news story titled, "ORNL Team Enlists World's Fastest Supercomputer to Combat the Coronavirus," featuring Jeremy Smith's team who used Summit to find potential drugs that could treat the virus. This work garnered extensive news coverage, including features in *Rolling Stone* magazine on *NPR.com*; it was also mentioned in the *Wall Street Journal*. In July 2020, the team wrote and disseminated another impactful research announcement that brought in extensive news coverage. This work featured Dan Jacobson's team and their "bradykinin hypothesis." The news coverage of this research grew slowly, with early coverage in *IEEE Spectrum* and other trade publications, followed by an unexpectedly popular story in a *Medium* health and wellness publication called *Elemental*. This feature was shared and favorited more than 50,000 times and became the source of a second wave of media interest that lasted through October 2020. In November, the team published features on two Summit users—out of four total finalists—who were named finalists for the 2020 Gordon Bell special prize for COVID-19. The team then published a story on Summit user Rommie Amaro's winning paper. Additionally, the team published a story on the four Summit users—out of six total—who were finalists for the traditional Gordon Bell prize and then a story on the winning team led by Lawrence Berkeley National Laboratory (LBNL); University of California (UC), Berkeley; and Princeton University.

To address pandemic-related work environment changes, the OLCF Outreach team also had to pivot with facility tours and events. On March 18, 2020, ORNL leadership sent all nonessential personnel home to work remotely through the remainder of the year. This policy put an immediate end to any on-campus tours. ORNL's Creative Services Division quickly invested in a 3D camera and software package called Matterport to develop virtual tours, and the OLCF Outreach team jumped at the opportunity. The OLCF's High Performance Computing tour was ORNL's first virtual tour released to the public. The tour received an impressive 14,222 impressions, 10,926 visitors, and 8,480 unique visitors by the end of 2020 and had the amusing honor of being listed in a popular Fodor's article entitled "The 9 Weirdest, Most Disturbing, or Just Plain Confusing Virtual Tours." An additional benefit of the virtual OLCF tour was the ability to support college students in the collaboratively taught University of Tennessee, Knoxville-ORNL course dedicated to data center management. The course typically involves time for the students to be on site at

[20] https://www.olcf.ornl.gov/tapia-2020-easy-peasy-gpu-programming-using-openmp/

ORNL, so the virtual tours of Summit and some of the surrounding data center spaces helped expose the students to the facilities almost as directly as if they were on campus in person. Some tours were able to occur on-site before the shutdown in March and for special VIP groups after the lab's doors were closed. In January, February, and March 2020, the OLCF conducted on-site tours for 51 groups, which included 438 individuals. Once the Matterport virtual tour became available, the OLCF team was able to conduct tours by video conference. The team hosted at least seven virtual tours, a much smaller number than in years past due to the reduced number of external requests from groups adapting to pandemic restrictions.

In October 2020, the team collaborated to host the second annual Exascale Day event. This year's event was virtual and extended beyond just the actual Exascale Day on October 18 with an extensive social media and features campaign, as well as a featured HPCWire interview program with Gina Tourassi, Bronson Messer, and John Turner called, "Imagine the World of Exascale." The holiday—an initiative of DOE's ECP and Cray, a Hewlett Packard Enterprise (HPE) company—honors scientists and researchers who will make groundbreaking discoveries with the help of some of the fastest supercomputers in the world, such as Frontier. The date was a creative play on the performance of exascale computers measured in exaflops. One exaflop equals one quintillion calculations per second, a number that is mathematically notated as 10^{18}.

In 2020, the team created 78 highlights—including science and technology highlights and features about Titan and Summit users, OLCF staff members, and OLCF resources—and more than 224 total outreach products.

Overall, the Outreach team produced at least 28 science highlights in 2020, including stories highlighting the Summit COVID-19 research, four stories in a new series called "Data Matters" highlighting data-related topics, and an introduction to the power and facility upgrades in preparation for the arrival of Frontier. Additionally, the team completed 23 technology stories and 11 people features.

1.5 LOOKING FORWARD

1.5.1 Application Readiness in the Exascale Era

The OLCF's CAAR is a partnership of OLCF staff, scientific application teams, vendor partners, and tool developers with the goal of readying a set of applications for the forthcoming Frontier exascale architecture. The OLCF-5 (Frontier) CAAR program is built on the legacy of the prior CAAR programs for OLCF-3 (Titan) and OLCF-4 (Summit). Eight CAAR applications were selected after a call for proposals in 2019, covering a broad range of scientific disciplines and employing a range of programming models and software designs. An additional 12 ECP applications with ECP funding are being engaged in the CAAR program, and these applications will enjoy staff and computing resources similar to the OLCF-funded CAAR projects. The applications that are part of the Frontier CAAR program are summarized in Tables 1.7 and 1.8.

Table 1.7. OLCF-funded applications at CAAR.

Application	PI	CAAR liaison	Scientific discipline
CHOLLA	Evan Schneider University of Pittsburgh	Reuben Budiardja	Astrophysics
CoMet	Dan Jacobson ORNL	Wayne Joubert	Biology
GESTS	P. K. Yeung Georgia Institute of Technology	Stephen Nichols	Fluid dynamics

LBPM	James McClure Virginia Polytechnic Institute and State University	Mark Berrill	Subsurface flow
LSMS	Markus Eisenbach ORNL	Markus Eisenbach	Materials science
NAMD	Emad Tajkhorshid University of Illinois Urbana-Champaign	Arnold Tharrington and Josh Vermaas	MD
NUCCOR	Morten Hjorth-Jensen Michigan State University/Facility for Rare Isotope Beams	Gustav Jansen	Nuclear physics
PIConGPU	Sunita Chandrasekaran University of Delaware	Ronnie Chatterjee	Plasma physics

Table 1.8. ECP-funded applications at CAAR.

ECP AD project	PI	ECP liaison	Scientific discipline
E3SM	Mark Taylor Sandia National Laboratories	Matt Norman	Climate
ExaAM	John Turner ORNL	Stephen Nichols	Additive manufacturing
ExaBiome	Kathy Yelick LBNL	Phil Roth	Biology
ExaSGD	Slaven Peles Pacific Northwest National Laboratory	Phil Roth	Electrical grid
ExaStar	Dan Kasen LBNL	Austin Harris	Astrophysics
ExaSky	Salman Habib Argonne National Laboratory	Bronson Messer	Astrophysics
ExaSMR	Steven Hamilton ORNL	Mark Berrill	Nuclear energy
GAMESS	Mark Gordon Iowa State University	Dymtro Bykov	Chemistry
LatticeQCD	Andreas Kronfeld Fermi National Accelerator Laboratory	Balint Joo	High-energy physics
NWCHEMeX	Theresa Windus Iowa State University	Dmitry Liakh	Chemistry
PELE	Jackie Chen Sandia National Laboratories	--	Combustion
WDMApp	Amitava Bhattacharjee Princeton Plasma Physics Laboratory	Ed D'Azevedo	Fusion

For the OLCF-5 Frontier project, the Application Readiness key performance parameter (KPP) metric was defined so that an application must achieve an application-specific figure of merit (FOM) that is four times greater on Frontier than on Summit. In CY 2020, the CAAR projects underwent a mid-project review to assess the progress toward meeting this KPP. The external review committee comprised Anshu Dubey (Argonne National Laboratory [ANL]), Jack DeSlippe (LBNL), and Kevin Clarno (University of Texas at Austin). The committee was asked to address the following charge questions.

- Have each of the CAAR projects developed a reasonable FOM that meaningfully and accurately reflects the required performance increase to perform their chosen challenge problem on Frontier?
- Have each of the CAAR projects realized a baseline FOM on Summit?
- Are each of the CAAR projects on track toward meeting their technical plan objectives, including a minimum FOM ratio of 4 between Summit and Frontier?

The review committee answered "yes" to all of these questions.

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
March 2021

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 2020 reporting period includes full CY production periods for the following HPC resources: the IBM AC922 (Summit), the General Parallel File System (GPFS) (Alpine), and the archival High Performance Storage System (HPSS). In 2020, the OLCF once again delivered all of the compute hours committed to the three primary allocation programs: INCITE, ALCC, and DD. The operational performance demonstrates that the OLCF delivered another prominent year of reliable and technically sufficient resources to the scientific research community.

2.1 OPERATIONAL PERFORMANCE SUMMARY

Operational performance measures the performance of the OLCF against a series of operational parameters. The two operational metrics relevant to the OLCF's operational performance are resource availability and capability use of HPC resources. Additionally, the OLCF describes resource use and GPU usage as reported numbers, not metrics.

2.2 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 comprises 4,662 high-density compute nodes, each equipped with two IBM POWER9 CPUs and six NVIDIA Volta GPUs. 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 the high-performance linpack and high-performance conjugate gradient benchmark applications from June 2018 until June 2020. The number of compute nodes changed in 2020 from 4,608 to 4,662. Three new cabinets with a higher memory footprint were added to the Summit system in July 2020 to support COVID-19 research.

2.3 GPFS (ALPINE AND WOLF) RESOURCE SUMMARY

In January 2019, the OLCF released Spider III, its next-generation global file system, to support the computational resources in the OLCF. Spider III is a single GPFS namespace named Alpine with a usable capacity of 250 PB and a file system-level performance of 2.5 TB/s. The Spider III 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 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.

2.4 DATA ANALYSIS AND VISUALIZATION CLUSTER (RHEA AND ANDES) RESOURCE SUMMARY

Andes, a new data analytics cluster, went into production on December 9, 2020. After 6 years of excellent service, the 512 node data analytics Linux cluster named Rhea was decommissioned on January 15, 2021. The primary purpose of these data analytics clusters is to provide a conduit for large-scale scientific discovery through the pre- and postprocessing 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. Each of Rhea's nodes contained two eight-core 2.0 GHz Intel Xeon processors with hyperthreading and 128 GB of main memory, which was upgraded in 2015 from 64 GB, whereas Andes is a 704 node cluster with each node containing two 16 core 3.0 GHz AMD EPYC processors and 256 GB of main memory. Rhea offered nine additional heterogeneous nodes, each of which boasted 1 TB of main memory and two NVIDIA Tesla K80 (Kepler GK210) GPUs. These nodes were moved to the Andes cluster. Data analytics clusters are currently connected to the OLCF's 250 PB high-performance GPFS, Spider III.

2.5 HPSS RESOURCE SUMMARY

The OLCF provides a long-term storage archive system based on the HPSS software product co-developed by IBM, Los Alamos National Laboratory (LANL), Sandia National Laboratories (Sandia), Lawrence Livermore National Laboratory (LLNL), LBNL, and ORNL. The ORNL HPSS instance is currently over 100 PB and provides ingestion rates of up to 30 GB/s via a 22 PB front-end disk cache backed by a 17 frame Spectra Logic TFinity tape library that houses 81 IBM TS1155 tape drives and over 10,000 tape media slots, giving ORNL a current capacity of 133 PB that is expandable well into hundreds of petabytes. The archive has ingested over 1 PB in 1 day within the last year with the average daily ingestion ranging between 100 and 150 TB. The archive environment comprises hardware from Dell, Brocade, DataDirect Networks, and Spectra Logic.

2.6 VISUALIZATION RESOURCE SUMMARY

The EVEREST facility has three computing systems and two separate state-of-the-art visualization display walls. The primary display wall spans 30.5 × 8.5 ft and consists of eighteen 1,920 × 1,080 stereoscopic Barco projection displays arranged in a 6 × 3 configuration for a 32:9 aspect ratio at 11,520 × 3,240. The secondary display wall contains sixteen 1,920 × 1,080 planar displays arranged in a 4 × 4 configuration, providing a standard 16:9 aspect ratio. 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.

2.7 OLCF COMPUTATIONAL AND DATA RESOURCE SUMMARY

The OLCF provided the Summit computational resource (Table 2.1) and the Spider III and HPSS data resources for production use in 2020. Supporting systems—such as EVEREST, Rhea, Andes, and data transfer nodes—were also offered. Metrics for these supporting systems are not provided.

Table 2.1. OLCF production computer systems for 2020.

System	Access	Type	CPU	GPU	Computational description			Interconnect
					Nodes	Node configuration	Memory configuration	
Summit	Full production	IBM AC922	3.45 GHz IBM POWER9 (22 core)	1,530 MHz NVIDIA V100 (Volta)	4,662	IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node)	(4,608) 512 GB DDR4 and 96 GB HBM2 per node; (54) 2 TB DDR4 and 192 GB HBM2 per node; >10 PB DDR4 + HBM + Nonvolatile aggregate	Mellanox EDR 100G Infiniband (Nonblocking Fat Tree)

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

2.7.1 OLCF HPC Resource Production Schedule

The OLCF computational systems entered production according to the schedule in Table 2.2. 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 2.2. OLCF HPC system production dates from 2008 to present.

System	Type	Production date[a]	Performance end date[b]	Notes
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]). Three 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	GPFS	January 1, 2019	Present	250 PB GPFS single namespace.
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, 2012 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, 2012 when partition size was reduced to 240,000 AMD Opteron cores. Beginning of Kepler installation.
JaguarPF	Cray XK6	February 2, 2012	February 13, 2012	Stability test. Restricted user access. 299,008 AMD Opteron 6274 cores. Includes 960 node Fermi-equipped partition.
JaguarPF	Cray XK6	January 5, 2012	February 1, 2012	Acceptance. No general access 299,008 AMD Opteron cores.
JaguarPF	Cray XK6	December 12, 2011	January 4, 2012	142,848 AMD Opteron cores.
JaguarPF	Cray XT5	October 17, 2011	December 11, 2011	117,120 AMD Opteron cores.
JaguarPF	Cray XT5	October 10, 2011	October 16, 2011	162,240 AMD Opteron cores.
JaguarPF	Cray XT5	September 25, 2009	October 9, 2011	224,256 AMD Opteron cores.
JaguarPF	Cray XT5	August 19, 2008	July 28, 2009	151,000 AMD Opteron cores.

[a] The production date used for computing statistics is 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.

2.7.2 Operational Performance Snapshot

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

Table 2.3. OLCF operational performance summary for Summit.

	Measurement	2019 target	2019 actual	2020 target	2020 actual
IBM AC922 (Summit)	Scheduled availability	85%	99.92%	90%	99.59%
	Overall availability	80%	98.52%	85%	98.73%
	MTTI[a] (hours)	NAM[c]	616	NAM[c]	788
	MTTF[b] (hours)	NAM	2,918	NAM	2,187
	Total usage	NAM	78.01%	NAM	90.11%
	Node-hours used[d]	NAM	31,144,061	NAM	36,572,249
	Node-hours available	NAM	39,924,230	NAM	40,585,974
	Capability usage				
	INCITE projects[e]	NAM	39.79%	NAM	49.79%
	All projects	30%	47.54%	35%	43.42%

[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.
[d] Does not include usage recorded during an outage.
[e] Does not include INCITE 2020 thirteenth month usage in January 2021.

Table 2.4. OLCF operational performance summary for HPSS.

	Measurement	2019 target	2019 actual	2020 target	2020 actual
HPSS	Scheduled availability	95%	99.79%	95%	99.92%
	Overall availability	90%	97.97%	90%	98.33%
	MTTI[a] (hours)	NAM	277	NAM	432
	MTTF[b] (hours)	NAM	874	NAM	1,463

[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 2.5. OLCF operational performance summary for Spider III, the external GPFS.

	Measurement	2019 target	2019 actual	2020 target	2020 actual
Spider III (Alpine)	Scheduled availability	90%	99.53%	95%	99.43%
	Overall availability	85%	98.18%	90%	98.93%
	MTTI[a] (hours)	NAM[c]	277	NAM	621
	MTTF[b] (hours)	NAM	581	NAM	873

[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.

For 1 year following system acceptance or a significant system upgrade, the scheduled availability (SA) target for an HPC compute resource is at least 85%, and the overall availability (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 1 year following system acceptance or a significant 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%.

2.8 RESOURCE AVAILABILITY

Details of the definitions and formulas that describe SA, OA, MTTI, and MTTF are provided in Appendix D.

2.8.1 Scheduled Availability

The SA is described by Eq. (1). The OLCF has exceeded the SA targets for the facility's computational resources for 2019 and 2020 (Table 2.9).

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

Table 2.9. OLCF operational performance summary for SA.

System	2019 target (%)	2019 actual (%)	2020 target (%)	2020 actual (%)
IBM AC922	85	99.92	90	99.59
HPSS	95	99.79	95	99.92
Alpine	90	99.53	95	99.43

2.8.1.1 Assessing Impacts to Scheduled Availability

Preventative maintenance is exercised only with the concurrence of the vendor hardware and software teams, the OLCF HPC Systems groups, and the NCCS 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.

In 2017, representatives from the OLCF, Argonne Leadership Computing Facility (ALCF), and NERSC agreed that during a scheduled maintenance, a significant event that delays the return of a system to scheduled production by more than 4 h will be counted as an adjacent unscheduled outage, unscheduled availability, and an additional interrupt. Before 2017, each facility handled this situation differently, and ASCR asked the OLCF to create a consistent definition.

2.8.2 Overall Availability

The OA of OLCF resources is derived by using Eq. (2).

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

As shown in Table 2.10, the OLCF exceeded the OA targets of the facility's resources for 2019 and 2020.

Table 2.10. OLCF operational performance summary for the OA.

System	2019 target (%)	2019 actual (%)	2020 target (%)	2020 actual (%)
IBM AC922	80	98.52	85	98.73
HPSS	90	97.97	90	98.33
Alpine	85	98.18	90	98.93

2.8.3 Mean Time to Interrupt

The MTTI for OLCF resources is derived by Eq. (3), and a summary is shown in Table 2.11.

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

Table 2.11. OLCF operational performance summary for the MTTI.

	System	2019 actual	2020 actual
MTTI[a] (hours)	IBM AC922	616	788
	HPSS	277	432
	Alpine	277	621

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

2.8.4 Mean Time to Failure

The MTTF is derived from Eq. (4), and a summary is provided in Table 2.12.

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

Table 2.12. OLCF operational performance summary for MTTF.

	System	2019 actual	2020 actual
MTTF[a] (hours)	IBM AC922	2,918	2,187
	HPSS	874	1,463
	Alpine	581	873

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

2.9 RESOURCE UTILIZATION 2020

2.9.1 Operational Assessment Guidance

The facility reports total system utilization (SU) for each HPC computational system as agreed upon with the program manager.

2.9.2 Resource Utilization Snapshot

For the IBM AC922 during the operational assessment period from January 1 to December 31, 2020, 36,572,249 Summit node-hours were used outside of outage periods from an available 40,585,974 Summit node-hours. The total SU for the IBM AC922 was 90.11%.

2.9.3 Total System Utilization

2.9.3.1 2020 Operational Assessment Guidance

SU is the percentage of time that the system's computational nodes run user jobs. No adjustment is made to exclude any user group, including staff and vendors.

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

The measurement period is for 2020, irrespective of the prescribed allocation period of any single program. For example, the INCITE allocation period follows a CY schedule. The ALCC program follows an allocation cycle that runs for 12 months, beginning on 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. The three primary 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 second production year of Summit, utilization was 90.11%, trending up from 78.01% during Summit's first year of production. Figure 2.1 summarizes the IBM AC922 utilization by month and by program for CY 2020.

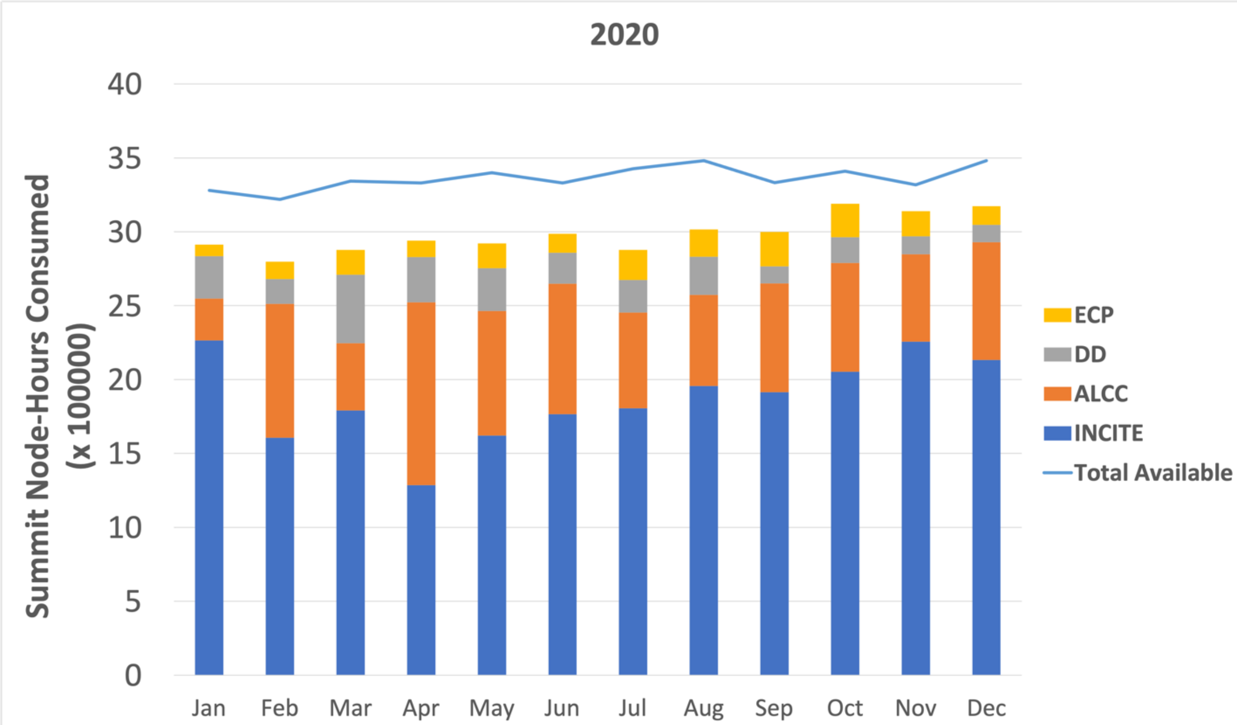


Figure 2.1. 2020 IBM AC922 resource utilization for Summit node-hours by program.

2.9.3.2 Performance of the Allocated 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 2020 INCITE allocation program was the largest program in 2020 with a commitment for 18.8 million Summit node-hours. The consumption of these allocation programs is shown in Table 2.13. As shown, all commitments were exceeded for each allocation program on Summit for 2020. This programmatic overachievement is partly due to the high uptime and diligent work of the OLCF operational staff.

Table 2.13. The 2020 allocated program performance on Summit.

Program	Allocation	Hours consumed (h)	Percent of total (%)
INCITE[a]	18,800,000	22,455,974	66.21
ALCC[b]	Allocation spans multiple CYs	8,724,048	25.72
DD	—	2,734,768	8.06
Total		33,914,790[c]	100

[a] Includes all INCITE program usage for CY 2020.
[b] Includes all ALCC program usage for CY 2020.
[c] Does not include usage outside the three primary allocation programs.

Nonrenewed INCITE projects from 2019 continued running through January 2020 under the OLCF's thirteenth 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 OAR review. It also serves to maintain high utilization while new projects establish a more predictable consumption routine. ALCC projects from the 2020 allocation period (ending June 30, 2020) were also granted extensions, as appropriate.

2.10 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 or upgrade, at least 30% of the consumed node-hours will be from jobs that request 20% or more of the available nodes. In subsequent years, at least 35% of consumed node-hours are from jobs that require 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 2020 was 35% for Summit based on years of service, as described previously. The OLCF continues to exceed expectations for capability usage of its HPC resources (Table 2.15). Keys to the successful demonstration of capability usage include the liaison role provided by Science Engagement section 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 engages with code developers to promote application portability, suitability to hybrid systems, and performance. The OLCF aggressively prioritizes capability jobs in the scheduling system.

Table 2.15. OLCF capability usage on the IBM AC922 (Summit) system.

Leadership usage	CY 2019 target	CY 2019 actual (%)	CY 2020 target	CY 2020 actual (%)
INCITE	NAM[a]	39.8	NAM	49.8
ALCC	NAM	52.7	NAM	32.3
All projects	30%	47.5	35%	43.4

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

The consumption of hours by capability jobs on Summit was 43.4%, which was above the 2020 target of 35%. There was a significant increase in the capability metric over 2019 for the INCITE program with approximately 50% of the hours reaching the capability level. ALCC saw a substantial decrease from 2019 because Summit was being used as a critical resource to the nation for COVID-19 research during the pandemic and due to lower overall capability utilization for all ALCC allocations. Most COVID-19 jobs were not at the capability level. To promote the execution of capability jobs, the OLCF provides queue prioritization for all jobs that use 20% or more of the nodes and further boost the very largest of these jobs, which use >60% of the nodes, through aging boosts. The OLCF assesses job data in 10% “bins” to understand the job size distribution. Furthermore, by assessing the aggregate bins of 20–60% and >60%, the OLCF can assess the impact of queue policy on delivered node-hours. Figure 2.2 provides the yearly average capability usage for each program, which describes the ratio of compute-hours delivered by capability jobs to the compute-hours delivered by non-capability jobs.

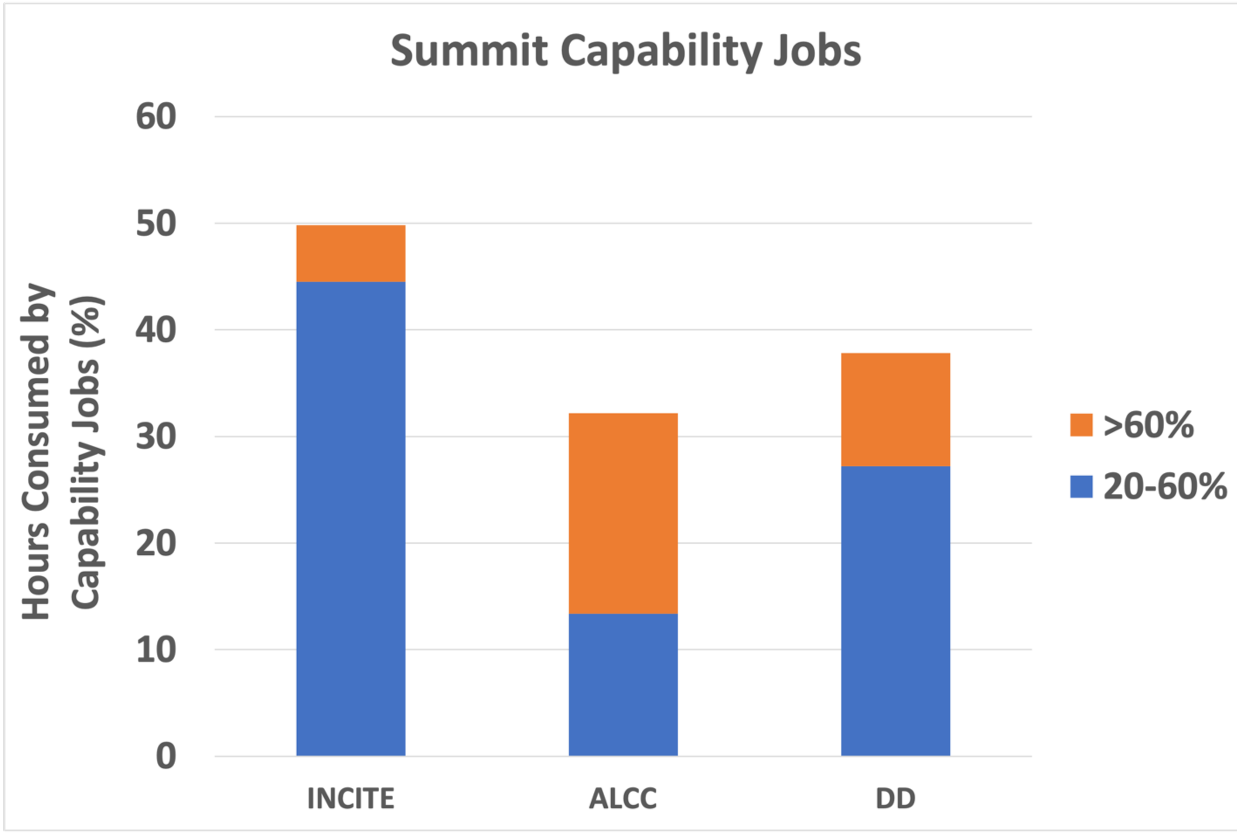


Figure 2.2. Summit capability usage by job size bins and project type.

2.11 GPU USAGE

The leadership computing systems at OLCF provide heterogeneous architectures that allow users to exploit a hybrid compute node that contains 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 2020, the OLCF continued tracking GPU usage on Summit through NVIDIA's system management interface tool. Figure 2.3 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 extraordinary 90% of hours used on Summit using the GPUs. Certainly, the overwhelming majority of the Summit node capabilities reside in the GPUs, and this number and the percentages from the ALCC and DD programs—both at approximately 85%—reflect the recognition and use of that capability from the OLCF user programs. Overall, Summit boasted a hefty 86% in 2020 for GPU-enabled applications.

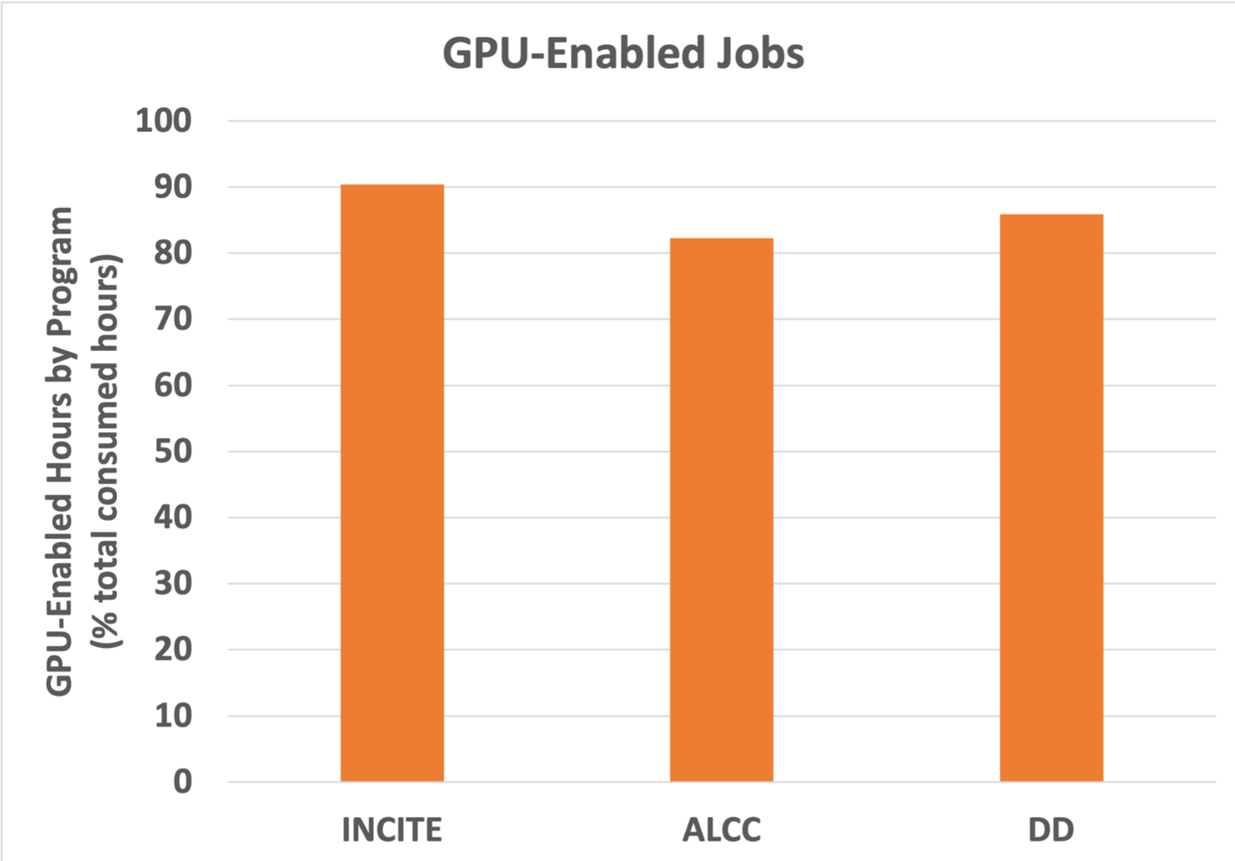


Figure 2.3. 2020 GPU-enabled usage by program.

2.12 CENTER-WIDE OPERATIONAL HIGHLIGHTS

2.12.1 Summit COVID-19 Cabinets

The COVID-19 pandemic pushed computational scientists to explore new ways of operating on and analyzing large datasets. In support of that effort, three additional Summit cabinets were purchased with supplemental federal funding for COVID-19-related research provided by DOE through the Coronavirus Aid, Relief, and Economic Security (CARES) Act. These additional cabinets uniquely provide a higher memory footprint inside Summit with 32 GB of HBM2 on the NVIDIA V100 GPUs and 2 TB of DDR4 main memory. The additional 54 nodes went into operation on July 1, 2020. During CY 2020, these new cabinets provided an additional 132,632 node-hours to Summit users. These additional resources contributed significantly to the development of an ML classifier to analyze more than 700,000 combinations of existing drugs for efficacy against COVID-19 by a team from the Mount Sinai Icahn School of Medicine. The additional larger memory nodes were also crucial to the ORNL-led development of a BlazingSQL/Dask/RAPIDS workflow for biophysical data analysis.[21]

2.12.2 Andes Data Analysis Cluster

This year, the OLCF staff successfully replaced the center's Rhea data analysis cluster with a brand-new AMD-based system called Andes.

[21] https://www.olcf.ornl.gov/2020/10/28/in-a-groundbreaking-move-summit-joins-forces-with-blazingsql-to-speed-up-data-query-processing-on-supercomputers/

For the previous 6 years, the Rhea cluster afforded OLCF users with the opportunity to perform data analysis and visualizations of simulations performed on the OLCF's supercomputer. The center's new data analysis system, which entered full production operations at the beginning of this CY, features more nodes, more memory, a faster interconnect, and new EPYC CPUs by AMD.

At its inception in 2014, Rhea—an Intel system—contained only 196 nodes and 64 GB of random-access memory (RAM) per node. Upgrades throughout the years brought it up to 512 nodes with 196 GB of RAM per node. Andes is even more powerful. At 704 nodes and 256 GB of RAM per node, Andes is the most powerful commodity cluster that the OLCF has ever deployed. It is so powerful and its chips draw so much heat that it uses active water-cooled rear door heat exchangers. The system is also connected via a Mellanox HDR InfiniBand interconnect with a water-cooled switch.

Users already familiar with the Rhea cluster should not experience a steep learning curve in transitioning to Andes. The system retains connectivity to the OLCF Spider III/Alpine center-wide GPFS and uses the same Slurm scheduler and workload manager that Rhea used in its final year.

2.12.3 Oracle SL8500 Tape Libraries Replaced

After transitioning the OLCF long-term storage from six Oracle SL8500 libraries to a new, single Spectra Logic Tfinity storage library, the OLCF needed to decommission a significant quantity of hardware. These Oracle tape libraries occupied around 1,000 ft^2 of data center space, and because of the time and cost constraints for equipment removal, several alternative options needed to be explored. These options ranged from removal for scrap, transferring to another entity, and selling the units through ORNL property sales. Two prime considerations for each option were how "environmentally friendly" the option was and how to minimize taxpayer cost. Working with ORNL's Property Management department, the Oracle libraries were listed on US General Services Administrations' GSAXcess system where four of the six units were claimed, removed, and transported at no cost to the OLCF by the new federal agencies. Finding other agencies to reuse this equipment was the ideal case from an environmental and cost perspective. The remaining two systems were again posted on the GSAXcess system, but no other federal agencies expressed interest. The OLCF then competitively contracted a services company that specialized in the removal, resale, or leasing of high-tech equipment to remove the two remaining libraries. This final effort ensured that the last two libraries were reused or recycled at a reduced cost to the center.

2.12.4 HPSS Improvements

The HPSS underwent full-stack hardware and software upgrades in 2020. To take advantage of features and performance optimizations in the new SpectraLogic Tfinity tape library, the HPSS software was upgraded from version 7.4.3 to 7.5.3. Additionally, new hardware for core servers, RAIT engines, Hierarchical Storage Interface (HSI) gateways, and small file disk movers were implemented, replacing aging hardware and modernizing the Linux operating system and configuration management tools. New all-flash metadata storage and a small file disk cache were implemented, replacing out-of-support hardware. Finally, steep growth in archival storage demand necessitated that the SpectraLogic library add four additional frames to the existing 17 frames, which introduced an additional 2,640 tapes, increasing the overall capacity by 30 PB. All upgrades were achieved with minimal downtime and will bring value to users who need long-term storage and to the OLCF by improving stability, security, and performance.

2.12.5 CADENCE Telemetry Service

Individual, bespoke data silos and pipelines have accumulated over the last decade within OLCF systems and services. Each pipeline is unique, can be brittle, and is typically designed and implemented with one use case in mind. However, the telemetry and log data are essential to the operation of the center. They are used tactically for troubleshooting and to reduce system downtime and strategically for decision making regarding the cost/feature tradeoffs of large supercomputers. These data are important not just to the OLCF but also to research scientists in the Computing and Computational Sciences Directorate who rely on this information to perform research on error detection, system optimization, and the run time variability of HPC systems.

Because this information is so critical, the Analytics and Monitoring (AM) team within the HPC Core Operations Group built a platform to finally treat telemetry data as a first-class citizen. The two main pieces of the platform, STREAM and CADENCE, are built on top of OLCF's SLATE infrastructure to take advantage of the reliability that Kubernetes provides. Altogether, the platform ensures the robustness, correctness, security, and scalability of the data. Deployment and transition to operations of these two systems were completed during the COVID-19 work-from-home phase.

The AM team and specific data owners have spent significant amounts of time together to document, troubleshoot, ingest, and provide access to their telemetry streams from many critical OLCF systems. Summit Temperature Control information, XALT data, and Atlas file system information are some of the early telemetry streams that were de-siloed and are now available via this centralized platform. Dashboards for the consumption of this information are easy to create and modify and are regularly used by operational teams for troubleshooting.

CADENCE is backed by 1.5 PB of extendable storage and will be able to keep telemetry data for the life of OLCF-5. It was designed to index 1 TB/s of information per day and can index more information than platforms with similar purposes at other national laboratories by an order of magnitude. STREAM has also exceeded performance expectations in terms of very low end-to-end latency. Data flows to dashboards and other data sinks typically within the sampling rate interval of the data source. STREAM provides a level of operational stability in the form of a stream buffer, ensuring that no system telemetry is lost, even during system downtimes for patching and upgrades.

Overall, the AM team and data owners within NCCS have developed and are continuing to support a robust data environment. The new capability provides system administration staff and other data owners and consumers with the tools and knowledge to successfully extract and use necessary data from systems to make data-driven decisions. Using the vision of treating this system information as a first-class citizen, the AM team has been able to lobby for and implement new ways of thinking. The AM team's motto to "centralize an immutable stream of facts" and "decentralize the freedom to act, adapt, and change" has finally been realized this year through the extraordinary efforts of this team.

2.12.6 eBPF System Monitoring

Greggd[22] is a new tool developed jointly between the Clusters and Storage teams to deploy specific Extended Berkeley Packet Filter (eBPF) codes in an operationally maintainable way. Written in Golang, the tool wraps the eBPF toolkit to trace and report on deep, system-level tasks and provides a standard way to output these operations as telemetry. This dynamic tracing provides operational insight into several deep system processes by adding a monitoring layer between user space and the Linux kernel. In particular, Greggd is being used to track Transmission Control Protocol (TCP) session lifetime and connection details, block device I/O latency, and all exec system calls done on a host. This provides the OLCF with an unprecedented level of system insight and user behavior monitoring within an HPC environment.

[22] https://github.com/olcf/greggd

3. ALLOCATION OF RESOURCES

CHARGE QUESTION 3: Is the allocation of resources reasonable and effective?

OLCF RESPONSE: Yes. The OLCF continues to enable high-impact science results through access to the 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.

3.1 ALLOCATION OF RESOURCES: FACILITY DD RESERVE TIME

This section provides insight into the strategic rationale behind use of the DD's reserve. It describes how the DD's reserve is allocated and lists the awarded projects, showing the PI name, sponsor organizations, hours awarded, and project title.

3.1.1 OLCF DD Program

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. Furthermore, through the ALCC program, the ASCR office allocates up to 20% of the facility's 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; and

3. to support R&D partnerships 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 and DOE offices that have large compute and data science challenges; and

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 could be entirely new areas with respect to HPC, or they could be areas in need of nurturing. Examples of projects are those associated with the ORNL Laboratory Directed Research and Development (LDRD) program; programmatic science areas, such as fusion, materials, chemistry, climate, nuclear physics, nuclear engineering, and bioenergy science and technology; and key academic partnerships, such as the UT-ORNL Joint Institute for Computational Sciences.

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 perform research that would not otherwise be possible. Section 8.3.2 provides 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 low thousands to 200,000 node-hours or more. In 2020, the OLCF DD program participants used approximately 7.5% of the total user resources on Summit, consuming almost 5.2 million Summit node-hours. Several COVID-19-related projects were initially allocated under the DD program but were subsequently moved to the ALCC program at ASCR direction. Appendix E provides a full list of DD projects for CY 2020.

4. INNOVATION

CHARGE QUESTION 4: (a) Have innovations been implemented that have improved the facility's operations? (b) Is the facility advancing research, either intramurally or through external collaborations, that will impact next-generation high performance computing platforms? (c) Is the facility effectively using its postdoctoral fellows?

OLCF RESPONSE: Yes. The OLCF actively pursues innovations that can enhance facility operations. Through collaborations with users, other facilities, vendors, and the broader digital infrastructure community, many of these innovations are disseminated and adopted across the country. Additionally, 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.

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. In 2020, the OLCF pursued innovative technological solutions and external collaborations to remain the state-of-the-art HPC facility in the United States despite the extraordinary circumstances of the COVID-19 pandemic. It is improbable to highlight all the innovative work performed by the OLCF within this report. Instead, this section highlights specific areas of innovation in operations and research in 2020.

This section also highlights the 14 postdoctoral fellows who were trained at the OLCF in CY 2020, the contributions they made to OLCF applications projects, and their career paths after their OLCF postdoctoral appointment, where appropriate.

4.1 OPERATIONAL INNOVATION

This section discusses examples of specific operational innovations to improve the facility operations: data hierarchy improvements, better system-wide metrics and benchmarking, accreditation to support protected health information, improved workflow and data services support, and application operation improvements.

4.1.1 Data Hierarchy Improvements

- HDF5 I/O middleware library: The OLCF identified a bottleneck in the HDF5 I/O middleware library on Summit, severely limiting the write I/O performance. OLCF staffer Bing Xie, in collaboration with the HDF5 team at LBNL, developed a solution. A new HDF5 module was deployed on Summit, encapsulating this solution. With this new module, the write performance of HDF5 on Summit improved up to 20 times.

- Small message communication latency improvements: In 2020, the OLCF tested and evaluated a novel GPU-to-GPU messaging mechanism called GDR_Copy that significantly reduces the latency of small message communication. This effort resulted in an OpenMPI with support for GDR_Copy, and it was released on Summit in April 2020.

- UnifyFS: UnifyFS is an ECP Software Technologies development project and is a collaboration between ORNL and LLNL. UnifyFS is a user-level burst buffer file system that aims to support diverse I/O patterns from data-intensive scientific applications. Specifically, UnifyFS provides a shared namespace to parallel applications by transparently aggregating storage devices on compute nodes, allowing parallel applications to benefit from the performance of fast node-local storage devices without code modifications. UnifyFS also supports popular I/O middlewares, including HDF5, MPIIO, and VeloC, and further simplifies the transition of applications that use such I/O middleware. In 2020, UnifyFS-v0.9.1 was released with substantial performance improvements and bug fixes. UnifyFS is currently available as a module on Summit in the OLCF.

- HPSS: HPSS is an archival storage technology that serves the backup and archival storage needs of the OLCF and dozens of other sites worldwide. In FY20, OLCF staffer Adam Disney improved the management of disk hardware in HPSS to better integrate it with the hardware's optimal configuration, allowing sites to make better use of their hardware and possibly increase performance.

4.1.2 Metrics and Benchmarking for Operational Improvements

- The OLCF enhanced the ability to acquire insights from large-scale long-term operational data accumulated from high-volume telemetry data streams from Summit. The OLCF invested in enhancing the ability to store, process, and obtain operational insights from the large-scale and long-term telemetry data from Summit and beyond.
- By recognizing the opportunities of analyzing the OLCF telemetry data streams that come with ever increasing volume and velocity, OLCF staff implemented data pipelines and methods that enable storage-efficient and analytics-efficient handling of such data. To achieve this level of efficiency, data formats, libraries, and tools—such as Apache Parquet, Apache Arrow, and Dask—were evaluated, employed in the data pipelines and workflows, and tightly integrated into center-wide resources, such as STREAM, CADENCE, SLATE, and Andes. Real-time raw data feed from the data center streaming into the center-wide Kafka bus "streams" and are streamed into a data lake for long-term storage and data analytics by data pipelines deployed on the SLATE OpenShift cluster that performs data aggregation and compression. Later, these datasets are analyzed with interactive Jupyter sessions or on-demand bulk batch jobs that leverage HPC resources, such as Andes.
- With this effort, the OLCF has significantly improved the latency (from impossible to minutes), scope (from a few hours to years), and width (from one dataset to multiple datasets) of the data analysis tasks or queries that can be run on the datasets streaming in. For example, on storage, 1 full year of Summit's 1 Hz per node and component power and temperature data that have around 100 columns and over 133 billion rows are compressed in 8 TB in-memory footprint—a 10× compression from its original 80 TB. Also on analytics, the response time of running an analysis on such a full year scope and resolution is under a few minutes, not hours, leveraging on-demand HPC resources in an interactive context. Finally, such an effort unlocked the ability to go beyond the manual visual correlation of multiple datasets limited to a very short range and scope, making it possible to perform automated correlation on multiple datasets on a range at a yearly level (i.e., full-year correlation study between job scheduler logs, per node and component thermal data, and cooling plant data). In future iterations, the OLCF plans to couple the result of such bulk analysis capability with real-time monitoring and analytics, implementing advanced monitoring that employs predictive analysis and its visualization.

4.1.3 Citadel 2.0: A Significant Improvement Using the Scalable Protected Infrastructure

- The NCCS team significantly improved the previously developed Citadel framework, which was originally designed to transfer encrypted personally identifiable information and personal health information (PHI) data from ORNL-protected enclaves to the IBM AC922 Summit. Dubbed Citadel 2.0, this new capability continues to support ORNL's Knowledge Discovery Infrastructure (KDI) enclave, which allows the highly protected enclave to use leadership-class capabilities but will now also support scalable computing on data governed by International Traffic in Arms Regulations (ITAR).

- The primary infrastructure difference between Citadel 1.0 and Citadel 2.0 is the underlying storage systems used to power the analysis while the protected data computation is occurring on Summit. Previously, the protected data were transferred via a highly secure Linux Unified Key Setup volume, which was then analyzed on Summit, allowing Summit compute resources to be used on the containerized PHI data. However, the implementation of this containerized volume approach prohibited parallel I/O—which obviously inefficiently uses the parallel file system upon which Summit resides—and also makes it impossible for users to view their results until they bring their data back into the KDI.

- To improve upon the data parallelism performance issues, usability was increased by allowing users to analyze results as soon as they are available, and to support ITAR computing, the Scalable Protected Infrastructure (SPI) enclave was built. This PHI- and ITAR-compliant enclave contains the minimal infrastructure required to support login and data transfer and an encrypted parallel GPFS. The true power behind the SPI enclave and the chief improvement of the Citadel process as a whole is the newly designed software and hardware infrastructure that allows HPC jobs to be submitted from within the SPI login node and on-the-fly provisioning of sufficient compute resources within the SPI enclave by "borrowing" the compute nodes from Summit. Great care is taken to ensure that compute nodes reprovisioned to allow for dispatch from the SPI login nodes are cleaned before and after the job's run time, including rebooting the hosts completely, which helps guarantee a level of security needed for PHI and ITAR computing. This new capability allows users to submit jobs from within the SPI login and have the true "OLCF experience" by allowing for parallel I/O and the immediate analysis of results on the encrypted SPI file system.

- The same external assessment of the review of hardware infrastructure, encryption practices, and other relevant components performed for Citadel 1.0 were performed for PHI and are currently being performed for ITAR. Citadel 2.0 has undergone internal alpha testing and is currently in the internal beta testing phase, both of which have resulted in minor improvements to security, policy, and usability. Protected data computing using Citadel 2.0 with PHI and ITAR is expected to be ready in early 2021.

4.1.4 Anchor: A New Diskless Cluster Provisioning System

Within the OLCF, large-scale commodity compute clusters are managed without local disks to ease configuration management across several hundred nodes. This diskless management relies on a collection of in-house scripts designed to build client compute images. However, the development of this tooling halted several years ago, and maintaining it as OLCF adopts new cluster software stacks has been challenging. In the years since, container technology has gained momentum and created several stable methods of solving the same problems with operating system image creation.

To this end, the OLCF developed Anchor, a diskless provisioning system that leverages container tools to build and boot cluster compute images. OLCF staffer Joseph Voss was the lead developer of the Anchor system. Image building is primary handled by integrating the Open Container Initiative tool, Buildah, with the existing configuration management system. However, node booting is done via an extensible early-boot initial RAM disk module created by the OLCF. This RAM disk first bootstraps authentication for the node to verify that it is a trusted client and is authorized to download compute node secrets. Once this is done, the image is downloaded to the node over Secure Shell, Network File System, or a direct connection to a remote Docker registry. This image is then compressed into a Squashfs file and mounted with a tmpfs read-write overlay.

This improves the OLCF deployment process and operational efficiency in several tangible ways. Primarily, using container tools ensures that system software stacks are immutable and entirely repeatable. Once an image is built, it is impossible to modify the contents of that image directly, and hardware can be easily reset to that known image state via a reboot. These images are also saved in an easily shared format with all the information needed to build them, meaning that images can be rebuilt identically and planned upgrades can be staged across the center. Additionally, the extensible RAM disk module provides deployment flexibility, which allows machines configured to boot in different methods to use the same underlying workflow and configuration management. Finally, once a cluster has booted, the compute nodes have no external dependency on a management node at run time, meaning these boot services can be spun up and down without impacting the stability of the running cluster. All these benefits combined have helped simplify large-scale cluster management at the OLCF and maintain system stability from the end user perspective. This work was presented as a State of the Practice talk at SC20.

4.1.5 BlazingSQL for HPC-COVID Consortium

The OLCF contracted BlazingDB Inc. to enable and optimize their GPU-accelerated database query engine software on Summit. BlazingSQL can perform database queries on terabytes of data in seconds by using NVIDIA GPUs, especially on the 32 GB V100 GPUs installed in the new COVID-19 cabinets. The need for accelerated database queries arose from a project that analyzed molecular docking data against SARS-CoV-2 (COVID-19) virial protein targets obtained on Summit. The contract was awarded in September 2020 and included three milestones: initial porting, optimizations for scalability using Unified Communication X, and user support, such as training. BlazingSQL participated in a data analytics workshop held at OLCF in October 2020 that reached over 50 potential users in diverse disciplines, such as bioinformatics and plasma physics, with shared common data analytics needs.

Publications: Jens Glaser, Josh V. Vermaas, David M. Rogers, Jeff Larkin, Scott LeGrand, Swen Boehm, Matthew Baker, Aaron Scheinberg, Andreas Tillack, Mathialakan Thavappiragasam, Ada Sedova, and Oscar Hernandez, "High-Throughput Virtual Laboratory for Drug Discovery Using Massive Datasets," *International Journal for High-Performance Computing Applications*, accepted 2021.

Josh V. Vermaas, Ada Sedova, Matthew Baker Swen Boehm, David Rogers, Jeff Larkin, Jens Glaser, Micholas Smith, Oscar Hernandez, and Jeremy Smith, "Supercomputing Pipelines Search for Thereapeutics Against COVID-19," *IEEE Computing in Science and Engineering* 23 (2020): 7–16. https://doi.ieeecomputersociety.org/10.1109/MCSE.2020.3036540.

Workshop: "Workshop on GPU Accelerated Data Analytics Virtual Lab on Summit: From Drug Discovery to Radiation Physics," presented at the OLCF on October 15, 2020.

4.2 RESEARCH ACTIVITIES FOR NEXT-GENERATION SYSTEMS

This section discussed examples of research innovations, including intramural and external collaborations, that will impact next-generation HPC platforms.

4.2.1 Research in Large-Scale AI, Data Systems, and Workflows

4.2.1.1 Data Optimization for Large-batch Data Parallel Training

To tackle the convergence challenge in the large-batch data-parallel training of deep neural network models, staff from the Analytics and AI Methods at Scale Group proposed a new scheme to alleviate the accuracy deterioration for DL training at scale. Because the large batch step usually results in local minima in the loss landscape, called the *generalization gap*, we introduced a workflow (Figure 4.1) to remove the noise data points in the feature space and smooth the loss landscape (i.e., fewer local minima). The iterative workflow consists of a feature extraction with a pretrained model, dimensionality reduction with t-SNE, and a noise identification with DBSCAN. We showed that the large-scale training on an optimized dataset can achieve better convergence, and the procedure can also facilitate the preparation of AI scientific datasets.

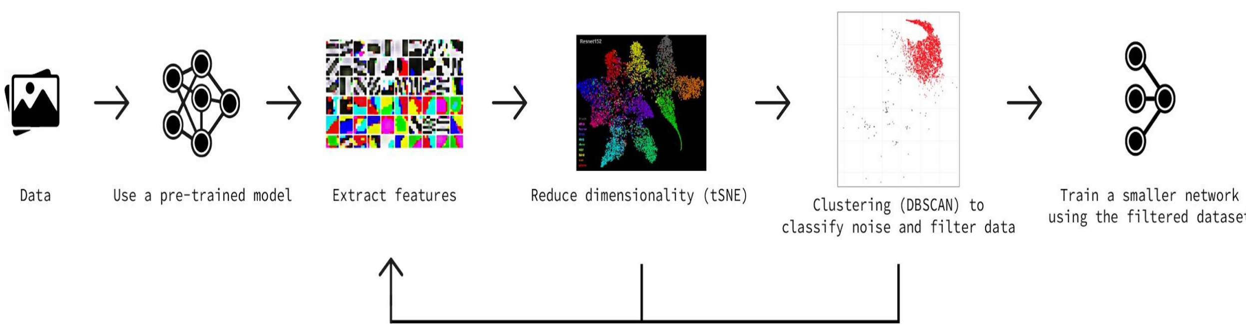


Figure 4.1. The processing pipeline of the data optimization strategy for large-batch training. Image Credit: ORNL.

Related paper: Shubhankar Gahlot, Junqi Yin, and Arjun Shankar, "Data Optimization for Large Batch Distributed Training of Deep Neural Networks," *2020 International Conference on Computational Science and Computational Intelligence* (CSCI).

4.2.1.2 DL-Based Clustering on MD Trajectories

To accelerate research for COVID-19, OLCF/AAIMS staff worked with MD experts to develop a drug discovery pipeline on Summit (Figure 4.2). The team employed the variational autoencoder (VAE) to cluster the MD trajectories to identify representative states of the spike protein. In the process, the team parallelized the VAE training and scaled up the data processing pipeline. The optimized DL-based clustering can handle much larger systems in shortened time-to-solution.

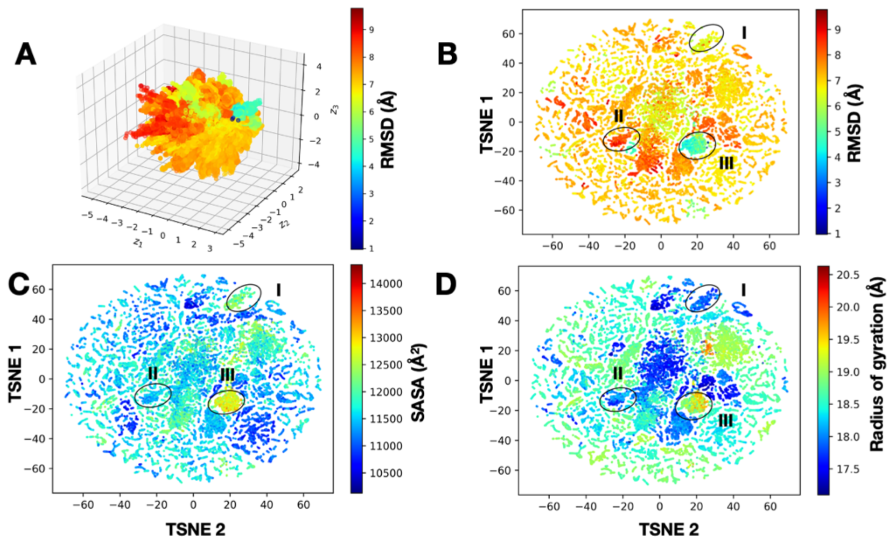


Figure 4.2. The VAE-based clustering of COVID-19 spike protein. Image Credit: ORNL.

Related paper: A. Acharya et al, "Supercomputer-Based Ensemble Docking Drug Discovery Pipeline with Application to Covid-19," *J. Chem. Inf. Model* 60, no. 12 (2020): 5,832–5,852.

4.2.1.3 Investigation of Python-Based GPU Data Analytics Frameworks on Summit

Previous OLCF User Surveys have emphasized the users' need for Python-related software and capabilities to support their scientific discovering process. Although a variety of Python software modules are available to support their data wrangling and data analysis activities, default distributions do not fully use the computational capability of the OLCF's Summit supercomputer.

OLCF, CSMD, and NVIDIA staff investigated the capabilities of NVIDIA RAPIDS and Dask to support multi-GPU and multi-node Python-based analytics frameworks on Summit. The process of testing and evaluating NVIDIA RAPIDS and Dask was critical for informing technical and user-facing aspects and to make the best use of Summit's resources. It was also important to collaborate with NVIDIA to provide feedback, report issues, and develop, extend, or optimize key features of these data analytics frameworks. Early adopters of this new OLCF software offering included teams from the COVID-19 HPC Consortium, ECP, and CAAR initiatives. NVIDIA RAPIDS and Dask frameworks will be available to Summit's users beginning in 2021.

Publications: Hao Lu, Benjamín Hernández, Suhas Somnath, and Junqi Yin, "NVIDIA Rapids on Summit Supercomputer: Early Experiences," *Nvidia GPU Technology Conference*, San Jose, California, March 2020.

Benjamín Hernández, Suhas Somnath, Junqi Yin, Hao Lu, Joe Eaton, Peter Entschev, John Kirkham, and Zahra Ronaghi, "Performance Evaluation of Python Based Data Analytics Frameworks in Summit: Early Experiences," in J. Nichols, B. Verastegui, A. Maccabe, O. Hernandez, S. Parete-Koon, and T. Ahearn (eds). *Driving Scientific and Engineering Discoveries Through the Convergence of HPC, Big Data and AI.* SMC 2020. Communications in Computer and Information Science, vol. 1,315. Springer, Cham. https://doi.org/10.1007/978-3-030-63393-6_24.

4.2.1.4 Pegasus Workflow Service Node via SLATE

The Pegasus Workflow Management System was implemented by using the NCCS SLATE services platform. This mechanism allows OLCF users to orchestrate their complex scientific workflows to marshal data from external sources and execute tasks on Summit. As more OLCF users become aware of the simplicity and benefits that this approach has on executing workflows on OLCF's resources, more scientists will begin adopting it and focus more on the science aspects of their research instead of worrying about how to access the computing resources available to them. The OLCF has created the capability to create Pegasus WSaaS, targeting the Summit supercomputer. This deployment builds upon the Kubernetes/Openshift platform (SLATE). The approach creates a fully functional Pegasus submit node in less than 25 min compared with other solutions that require experience with the workflow management system software and its dependencies, as well as require days of testing and debugging.

Publication: G. K. Papadimitriou, J. Vahi, V. Kincl, E. Deelman Anantharaj, and J. Wells, "Workflow Submit Nodes as a Service on Leadership Class Systems," *2020 Practice & Experience in Advanced Research Computing Conference Series*, July 26–30, 2020, Portland, Oregon. doi: 10.1145/3311790.3396671.

4.2.1.5 Workflows for ANL/APS Observational Data for ML/DL on Summit

Synchrotron light sources are routinely used to perform imaging experiments, and the real-time processing of synchrotron-based micro-tomography data presents a significant yet tractable computational problem. Preliminary exploration was demonstrated with an end-to-end scientific workflow on Summit based on micro-computed tomography data. The workflow comprises ANL/Advanced Photon Source (APS) data injection, TomoPy-based data reconstruction, DL-based data denoising, and simulation-based (i.e., LBPM) data segmentation. This is a promising step toward real-time analytics on APS data at the OLCF.

Publication: J. E McClure, J. Yin, R. T Armstrong, K. C Maheshwari, S. Wilkinson, L. Vlcek, Y. Wang, M. A Berrill, and M. Rivers, "Toward Real-Time Analysis of Synchrotron Micro-Tomography Data: Accelerating Experimental Workflows with AI and HPC," *Driving Scientific and Engineering Discoveries Through the Convergence of HPC, Big Data and AI. SMC 2020. Communications in Computer and Information Science*, vol. 1,315. Springer, Cham. https://doi.org/10.1007/978-3-030-63393-6_15.

4.2.2 Next-Generation Hardware and System Architecture Exploration

- ARM Testbed System: The ARM-based testbed capabilities were expanded with the inclusion of 16 HPE Apollo80 nodes in which each is equipped with Fujitsu ARM A64fx processors. OLCF staffer Ross Miller led the upgrade effort. These are the same Fujitsu ARM processors deployed on the Fugaku system, and they are the first commercially available processors to implement ARM's Scalable Vector Extension instructions. This upgrade allows OLCF users to try scientific applications on this new architecture.

- PathForward: OLCF staff were deeply engaged with the ECP PathForward program. OLCF staff member Scott Atchley served as DOE's technical representative for AMD's PathForward contract. OLCF staff members Chris Zimmer and Scott Atchley reviewed PathForward milestone reports. OLCF staff engagements with the PathForward vendors led to direct engagements with AMD and Cray initially and the HPE after the acquisition of Cray as part of the CORAL-2/OLCF-5 project. The PathForward held the final vendor reviews in September 2020. DOE particularly recognized AMD's results for being on time and for significantly impacting exascale systems.

- Emerging ML workloads: In FY20, members of Technology Integration Group working with IBM used the IBM Platform Load Sharing Facility (LSF) Simulator to model the impacts of the emerging ML workload on Summit's capability metric. The simulation helped identify techniques that might limit the impact to the capability metrics and ultimately discerned an upper bound on the workload that would degrade the ability to serve capability cycles.

- Efficient data movement within a Lustre parallel namespace: As a mitigation strategy, the Technology Integration Group started an R&D effort to implement efficient data movement techniques within the OLCF-5 Orion Lustre file system. OLCF staffer James Simmons has been working to improve data movement by accelerating data copy within the kernel on both Lustre servers and on the clients. He implemented a new communication channel within Lustre via Netlink and YAML for external tools to control data migration between storage pools on and between multiple servers. OLCF staffer Chris Brumgard has been working toward implementing a new Z File System (ZFS) virtual device (vdev) that would automate the migration of data between physical ZFS tiers on a server. This higher level vdev would enable data movement from one set of vdevs to another without modifying the rest of ZFS. This would also reduce the burden on Lustre by presenting one file system that comprises both tiers, thus drastically reducing the number of pools that must be managed and lessening the migration burden within Lustre. Ultimately, both innovations should benefit not just Orion and other smaller scale file systems at ORNL but also the Lustre and ZFS communities in which tiering has been a sought-after feature.

- Scalable Data Infrastructure for Science (SDIS): The SDIS Initiative at the OLCF aims to tackle the grand challenge of data infrastructure and management by defining and building a scalable data framework that addresses all stages of the data life cycle in support of a scientist's workflow to accelerate scientific breakthroughs. SDIS uses three focus areas to tackle this challenge.
  1. A scientist's sandbox: With an ecosystem of computational resources, data must be moved with low latency for "immediate" consumption. The sandbox framework supports scientists' creativity from inception to experiment to results across a data ecosystem while allowing facilities to deploy and maintain their own environment.
  2. Library: A key part of science is knowledge dissemination. The library supports the publication and archival aspects of the data life cycle and is part of the infrastructure that is vital to the FAIR (findable, accessible, interoperable, and reusable) data principles. SDIS provides a library framework to support the publication of scientific data, and its main feature is an ecosystem of libraries that are interoperable and searchable across facilities.
  3. Governance: Data governance is the process of managing the availability, usability, integrity, and security of data based on data standards and policies. Through the SDIS Data Assets Council, a governance framework will be defined, providing guidance and best practices in data management, interoperability, data sharing, privacy, and other data-related topics.

4.2.3 Data Science Benchmarking at Scale

OLCF staff members co-led an effort of showing in the ECP 2020 Annual Meeting results by running two mlperf-hpc DL benchmarks: cosmoflow (arXiv:1810.01993) and climate segmentation (arXiv:1808.04728). Specifically, the training throughput was shown as a function of workers for three HPC systems—Summit, Theta, and Cori—by running the same reference implementation on all of them.

In 2020, the OLCF continued experimenting with the use of the R language as an HPC analytics engine. As in recent years, great emphasis was placed on effectively using the computational power of the GPUs on Summit for analytics and big data. One motivator in this effort is the CORAL-2 BDAS suite of benchmarks whose reference implementations were originally released with only CPU backends. Since then, highly optimized GPU implementations have been provided for several of the benchmarks, and scaling evaluations were performed on Summit. Some of this work culminated in the publication of a paper titled, "A Survey of Singular Value Decomposition Methods for Distributed Tall/Skinny Data," which was published in the *IEEE/ACM 11th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA)*, a workshop at the SC20 conference. The paper explores several algorithms for computing singular value decomposition (SVD) on data with many more rows than columns. This data layout is typical for analytics and big data, and SVD is one of the most important operations for the field. The paper explores three of the most common algorithms for computing SVD, their implementation details, and various tradeoffs between them, then concludes with an extensive set of CPU and GPU benchmarking on Summit.

4.3 POSTDOCTORAL FELLOWS

4.3.1 CSEEN 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. Because of 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 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 exascale and data-intensive computing; (2) make ASCR facilities available through limited-term appointments for applied work on authentic problems with highly productive work teams and increasingly cross-disciplinary training; and (3) raise the visibility of careers in computational science and engineering to build the next generation of leaders in computational science. In CY 2019, the OLCF began to leverage additional funding from the ECP to augment the CSEEN program with additional postdoctoral fellows.

The OLCF CSEEN postdoctoral program seeks to provide opportunities to bridge the experience gap between the need to address domain science challenges and the need to 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 that is funded by OLCF-5 and the ECP. This model offers the greatest potential for scientific breakthroughs through computing and provides ample opportunities 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 trainees is the opportunity to publish and present research accomplishments.

In CY 2020, 14 postdocs were members of the OLCF workforce. Of those 14, 10 were fully supported by OLCF funds, including six postdoctoral fellows supported by ECP Application Integration. Four postdocs were supported by sources outside the OLCF, including individual ECP Applications Development (AD) projects and Scientific Discovery through Advanced Computing (SciDAC). No postdocs were partially supported by OLCF and another source. The background and current work of these postdocs in the Science Engagement section are described as follows.

- Paul Eller joined the Algorithms & 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 nonblocking 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. Since joining the OLCF, Eller has collaborated with the Combinatorial Metrics (CoMet) Comparative Genomics project team to study the performance benefit of using reduced precision numerical methods in computationally intensive GPU kernels.

- Nicholson Koukpaizan joined the Algorithms & Performance Analysis group in June 2020 after receiving his PhD in aerospace, aeronautical, and astronautical engineering from the Georgia Institute of Technology (Georgia Tech). As a member of Georgia Tech's Nonlinear Computational Aeroelasticity Laboratory, Koukpaizan researched techniques for modeling fluidic actuation for aerodynamic flow control with special focus on jet interaction fluidic oscillators sited on helicopter bodies. Since joining the OLCF, Koukpaizan has collaborated with the ECP ExaSGD applications development project to optimize parts of their software on Summit, help them prepare their software to run well on OLCF's future Frontier system, and help them deploy a CI pipeline on OLCF resources. He also continues his research into the modeling of jet interaction fluidic oscillators and is currently working to improve the scalability and performance of his model when running on Summit's GPU-accelerated nodes.

- 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, where he researched techniques for performance measurement and the 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 the message passing interface. Since joining OLCF, Eberius has collaborated with the ECP ExaBiome applications development project team to help them prepare their software to run well on OLCF's future Frontier system. In particular, he has ported parts of the MetaHipMer de novo genome assembly software and its dependencies to the emerging Frontier software stack and has compared the performance and scalability of that software when running on pre-Frontier systems with the same when running on Summit.

- Vassilios Mewes joined the Nuclear, Particle, & Astrophysics group in September 2019 after a postdoctoral stint at the Rochester Institute of Technology. Mewes 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 by using the Einstein Toolkit. He is expanding his expertise at ORNL in working on neutrino interaction physics with the ExaStar collaboration as part of the ECP. He has worked primarily on modifying and porting tabular implementations of neutrino-matter interaction kernels to GPU architectures for use in core-collapse supernova and binary neutron star merger simulations. Mewes accepted a staff position in the Nuclear, Particle, and Astrophysics group at OLCF that starts in May 2021.

- Charles (CJ) Stapleford joined the Nuclear, Particle, and Astrophysics group in October 2020 after obtaining his PhD from North Carolina State University. Stapleford's thesis work was on neutrino oscillation physics in core-collapse supernova environments, including the use of Boltzmann transport solvers coupled to approximate oscillation physics. He 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.

- Muralikrishnan (Murali) Gopalakrishnan Meena joined the Advanced Computing for Life Sciences and Engineering group in July 2020 after earning his PhD from UC Los Angeles on using network (graph) theory to characterize, model, and control vortical interactions in turbulence and wake flows. In particular, he introduced a network community-based framework to formulate reduced-order models and perform flow modification of complex laminar and turbulent vortical flows. Meena's main research at OLCF focuses on using ML to formulate sub-grid-scale turbulence closures for cloud resolving models. Additionally, he 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 via sensor measurements (e.g., flood forecasting), and identifying key interactions in various fungal communities.

- Isaac Lyngaas joined the Advanced Computing for Life Sciences & Engineering group in September 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 performance portability approaches for computing the cloud resolving model within the Energy Exascale Earth System Model (E3SM)-multiscale modeling framework (MMF). These efforts include porting the existing CRM code to a performance portable framework and investigating the use of this framework for upcoming new hardware architectures on Frontier and Aurora. Additionally, Lyngaas's research has focused developing numerical algorithms for approximating partial differential equations (PDEs) with the specific focus of performance on accelerated hardware.

- Antigoni Georgiadou joined the 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 Fermi National Accelerator Laboratory (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 collaborates with the ECP ExaStar project, which builds on the current capabilities of astrophysics codes, such as FLASH, for multiphysics 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 will also collaborate with the ExaSky project, which aims to maximize the computing power for cosmological simulations.

- Elvis Maradzike joined the Advanced Computing for Chemistry & Materials 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 and modeling the physics of strongly correlated electrons.

- Justin Lietz joined the Nuclear, Particle, and Astrophysics group in September 2019 after completing his PhD in nuclear physics from Michigan State University (MSU). Lietz's 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. He will be developing high-performance algorithms and data structures for many-body physics codes, such as NUCCOR, to enable novel nuclear structure calculations on the exascale generation of supercomputers, such as Frontier.

- Paul Mott joined the Advanced Computing for Chemistry & 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 will be collaborating with the GAMMESS ECP project to optimize their software for the upcoming Frontier exascale system.

- Swarnava Ghosh joined the Advanced Computing for Chemistry & 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. He earned his PhD from Georgia Tech 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 by using first principles real-space multiple scattering calculations, such as the dependence of magnetism on disorder in manganese-antimony-tellurium quantum material systems and the influence of spin to the mechanical behavior of magnetic structural alloys. He contributes to the development of the ORNL locally self-consistent multiple scattering (LSMS) code. In particular, he is extending the capability of this code to the efficient calculation of interatomic forces and non-spherical atomic potentials.

- 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. She contributes to the development of new capabilities of the LSMS first principles code where she is extending the calculations to capture the behavior of insulators and semiconductors by accurately calculating the electron charge by using the Lloyd's formula formalism.

- 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. Mao is working on the NUCLEI SciDAC-4 project where his research focuses on coupled-cluster theory for atomic nuclei. Specifically, he is working 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 exascale computing resources that use tightly coupled distributed algorithms.

Risk Management

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
March 2021

5. RISK MANAGEMENT

CHARGE QUESTION 5: Is the facility effectively managing operational risks?

OLCF RESPONSE: Yes, the OLCF has a very successful history of anticipating, analyzing, rating, and retiring 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.

This section lists and describes the major operational risks for the OLCF in CY 2020. Planned mitigations and implementations are included in the subsequent descriptions. As of this writing, the OLCF has no high-priority operational risks, but because 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 OLCF operations and its various projects (OLCF-5 during CY 2020). Each project execution plan refers to the main *Risk Management Plan* but might 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 during the weekly operations meeting attended 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 issues such as safety, funding, expenditures, and staffing. The specific, more focused risks are concerned with the 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 in management reserves as much as possible. This reserve is continually reevaluated throughout the year.

5.2 MAJOR RISKS TRACKED IN 2020

Table 5.1 contains the major risks tracked for OLCF operations in 2020. The full OLCF operations risk register is available upon request. The selected risks are all rated medium or high in impact.

Table 5.1. 2020 OLCF major risks.

Risk ID/description	Probability/impact	Action	Status
406: System cybersecurity failures	Low/high	Mitigating	The OLCF continues to see a rise in the quantity of cybersecurity attacks against the computer resources. This increase does not directly correlate to higher success rates because the OLCF employs various techniques to repel these attacks, such as proactive patching for zero-day exploits, the formal review of cybersecurity plans, a two-factor authentication requirement for system access, and a multifactor authentication level 4 requirement for privileged access to OLCF resources.
723: Safety—personal injury	Low/medium	Mitigating	Risk can be reduced by monitoring worker compliance with existing safety requirements, holding daily toolbox safety meetings, performing periodic surveillances using independent safety professionals, performing joint walk-downs with management and work supervisors, and encouraging the stop-work authority of all personnel. Observations from safety walk-downs are evaluated for unsatisfactory trends (e.g., recurring unsafe acts). Unsatisfactory performance will receive prompt management intervention commensurate with the severity of the safety deficiencies.
917: Robust support will not be available to ensure portability of restructured applications	Medium/medium	Mitigating	Multiple instantiations of compiler infrastructure tools will be adopted to maximize the exposure of multiple levels of concurrency in the applications. Work with vendors continues to improve compiler technology and other tools.
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 were identified for several key positions, including the NCCS division director, NCCS director of science, and technology integration group leader.
1063: Programming environment and tools might be inadequate for future architectures	Medium/medium	Mitigating	The OLCF will continue to engage with users, standards organizations (e.g., OpenMP, OpenACC, and others), and tool and hardware vendors to encourage, facilitate, and enact responses to user feedback and anticipate architectural trends in key standards and tools.

Table 5.1. 2020 OLCF major risks (continued).

Risk ID/description	Probability/impact	Action	Status
1142: OLCF cost increases due to fewer computer room customers to distribute maintenance and operation costs	Low/high	Mitigating	In 2020, the data center customer base increased by one. As 2021 begins, the 5600 E102 will remain under construction for most of the CY.
1145: Changes from external project managers cause development impacts to HPSS	Medium/medium	Mitigating	IBM has continued to push for items that are not on the development roadmap to support requests of potential customers and for features that might not meet the Technical Committee's release schedule.
1240: Failure to handle export-controlled information (ECI) properly	Low/high	Mitigating	Staff with elevated privileges on systems on which ECI can be accessed go through annual training and refreshers on how to handle ECI. Project PIs and members participate in an initial project briefing during which an ORNL export control analyst describes the categorization of the project based on the project application. OLCF cybersecurity staff contribute to this briefing and outline what storage resources can handle ECI.
1323: Lack of adequate facilities for the next OLCF system	Low/high	Accept	ORNL and the OLCF plan to house the follow-on system to OLCF-5 in Building 5600. However, the preferred approach would be to construct a new building that is designed to meet the needs of the program well into the future from the beginning. The Office of Management and Budget rejected third-party financing as a method of building such a facility, so this would need congressional line-item funding.
1245: System unavailability due to mechanical/electrical system failure	Low/high	Mitigating	The 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. All preventative maintenance activities will be performed, and inspections and monitoring will be performed where possible.

5.3 NEW OR RECHARACTERIZED RISKS SINCE LAST REVIEW

5.3.1 Recharacterized Risks

In 2020, there were no recharacterized risks in the OLCF Operations Risk Register.

5.3.2 New Risks in This Reporting Period

The following risks were created and tracked during CY 2020. They are included with their risk creation date, mitigations, and triggers.

Table 5.2. Risks created and tracked during CY 2020.

Risk ID 1328	OLCF operations impacted by COVID-19 outbreak		
Risk owner	Georgia Tourassi		
Creation Date	March 1, 2020		
Status	Accept—Current		
Probability	Low		
Impact	**Cost:** Low	**Schedule:** Low	**Scope/Tech:** Low
Mitigations	ORNL has added the capacity to remote access tools, such as VPN and Citrix. Additionally, Direct Access for Windows systems reduces consumption on VPN and Citrix.		
Triggers	Feedback from users or increased unscheduled downtime during the work-from-home period.		
Latest Update	Reliance on staff home network connectivity opens an additional vector for this risk to occur. During the first week of most staff working from home, there have been reports of receiving "all circuits busy" when using personal cell phones to dial into audio conference meetings.		

5.4 RISKS RETIRED DURING THE CURRENT YEAR

Table 5.3. Risks retired and tracked during CY 2020.

Risk ID 1246	White Oak Creek flooding causes system unavailability		
Risk owner	James P. Abston		
Status	Retired		
Retirement Comment	During construction of OLCF-4, a backflow preventer was installed on the floor drain that backed up when the creek flooded to address this issue.		
Impact	**Cost:** Low	**Schedule:** Low	**Scope/Tech:** Low
Mitigations	Properly maintaining the creek beds of White Oak Creek.		
Triggers	Examining future flooding events to determine whether the rise in the creek can predict potential future risk encounters.		

5.5 MAJOR RISKS FOR NEXT YEAR

Summit's operations are critical to the success of the OLCF in 2021, and one risk directly impacts Summit's ability to operate: Risk ID 1245, "System unavailability due to mechanical/electrical system failure."

Significant construction activities continue to occur in and around Building 5600. These activities require the closure of some areas and ingress/egress paths for staff and visitors. These closures require an increased focus on safety in and around the work areas by all staff (Risk ID 723).

ORNL's response to the COVID-19 pandemic will continue to impact the location of the majority of the OLCF's workforce in 2021, and the performance of the OLCF against the metrics provided by the sponsor will continue to be monitored (Risk ID 1328).

Finally, the programming environment for the OLCF-5 system will be different enough that experiences on Summit might not sufficiently prepare 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 were encountered and effectively mitigated in CY 2020. A short summary of the status and impact of the risk on the operations of the OLCF is included.

Table 5.4. Risks encountered and effectively mitigated in CY 2020.

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, such as VPN and Citrix. Additionally, Direct Access for Windows systems reduces consumption on VPN and Citrix.		
Triggers	Feedback from users or increased unscheduled downtime during the work-from-home period.		
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, some users experienced issues with Summit nodes when their codes were running. After a full root cause analysis with IBM, a defect was determined to exist in the power supply for the AC922 nodes that is only exacerbated by fast and high-frequency transient current. IBM had developed a strategy to replace all the power supplies in Summit in CY 2020, but this was impacted by the COVID-19 pandemic, and that activity is set to be completed in CY 2021.		
Mitigations	Working closely with the vendor to track and identify the root cause of every failure as quickly as possible.		
Triggers	Problems encountered during early science and ongoing operations.		

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 and health program to ensure that DOE contractors and their workers operate a safe workplace. Additionally, 10 CFR 851 establishes procedures for investigating if a violation of a requirement has occurred, determining the nature and extent of any such violation, and 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 called the Standards-Based Management System (SBMS). SBMS contains work control requirements that describe the processes to be used in ORNL operations and R&D activities to implement integrated safety management (ISM) functions and principles.

One key feature of the 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 safely. They are 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 nonemployees. Maintenance work in the OLCF is conducted under a work plan that is developed by Facilities and Operations Directorate 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 nonemployees or subcontractors/vendors is performed in accordance with a hazard analysis. The subcontractor/vendor hazard analysis is a requirement that is 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 safety culture at ORNL is reflected in these processes, which seek to reduce and prevent personnel injuries and potential personnel exposure to hazards associated with facility operation.

OLCF operations in the NCCS remained safe, efficient, and effective because there were no total recordable cases and no Days Away Restricted or Transferred in FY 2020. However, the center did experience one first aid case in 2020. That case and the center's actions related to the case are explained later in the section.

The following activities are ES&H highlights from CY 2020.

- COVID-19 significantly impacted how the center staffed and protected essential personnel. The OLCF provided 24/7/365 support for center operations. COVID-19 protocol as established by ORNL was implemented throughout the year to protect the operating staff. Additionally, the center changed many existing protocols, such as separating essential staff, replacing face-to-face shift turnover with telephone turnover, increasing sanitization efforts, and limiting access to the operational control room. The center also identified which operator responsibilities are critical to properly maintaining the operating status of the OLCF. Once these activities were identified, emergency fill-in guides were developed for each task, and additional OLCF staff were trained by senior operators to these guides. This effort provided additional essential personnel that could be used for center support to fill in if the operating staff were impacted by COVID quarantines or sickness. There were several staffing concerns related to quarantines that required changing shifts to provide the 24/7 support.

- First aid event: The OLCF first aid injury occurred when an employee stepped across a floor opening to look in the opening and stepped on a removed floor tile. The removed tile slid across the floor, and the action resulted in the employee sustaining a muscle strain. The tile had been placed in the same orientation from the removed location on the adjacent tile. Therefore, the employee did not notice and geometric difference of the grid lines and stepped across the floor opening, putting one foot on the removed tile. Part of ORNL's and the OLCF's safety culture is to take any safety-related event seriously, learn from that event, and take actions to prevent reoccurrence. Therefore, in response to the incident, the RSSs were revised, and information was added concerning stepping on removed tiles, barricading, placing cones on removed tiles, and rotating removed tiles 45 degrees. The event was also shared with other ORNL data centers, and information will be added to the next revision of the data center access training module concerning the event and controls. Finally, the employee discussed the event in several staff meetings as a lesson learned for others.

- The center safely deployed several new systems, including adding 29 new cabinets. COVID-19 protocols greatly impacted vendor/subcontractor support and how the center proceeded with the electrical/mechanical installation and hardware installations. However, the center was successful with working within the protocols, separating staff, using ORNL Health Services guidelines for vendors/subcontractors, and performing work with internal essential staff. All work was performed safely and on schedule to meet customer needs.

- Several operational policies were established and implemented in relation to the operational support operators. These new policies identified at-risk areas, such as center access and recording upset conditions, and established written protocol to be followed. Additional policies are in draft form, and others are yet to be drafted at the time this report was authored.

- The OLCF-5 project *Health and Safety Plan* and *Hazard Analysis* were reviewed and revised in 2020.

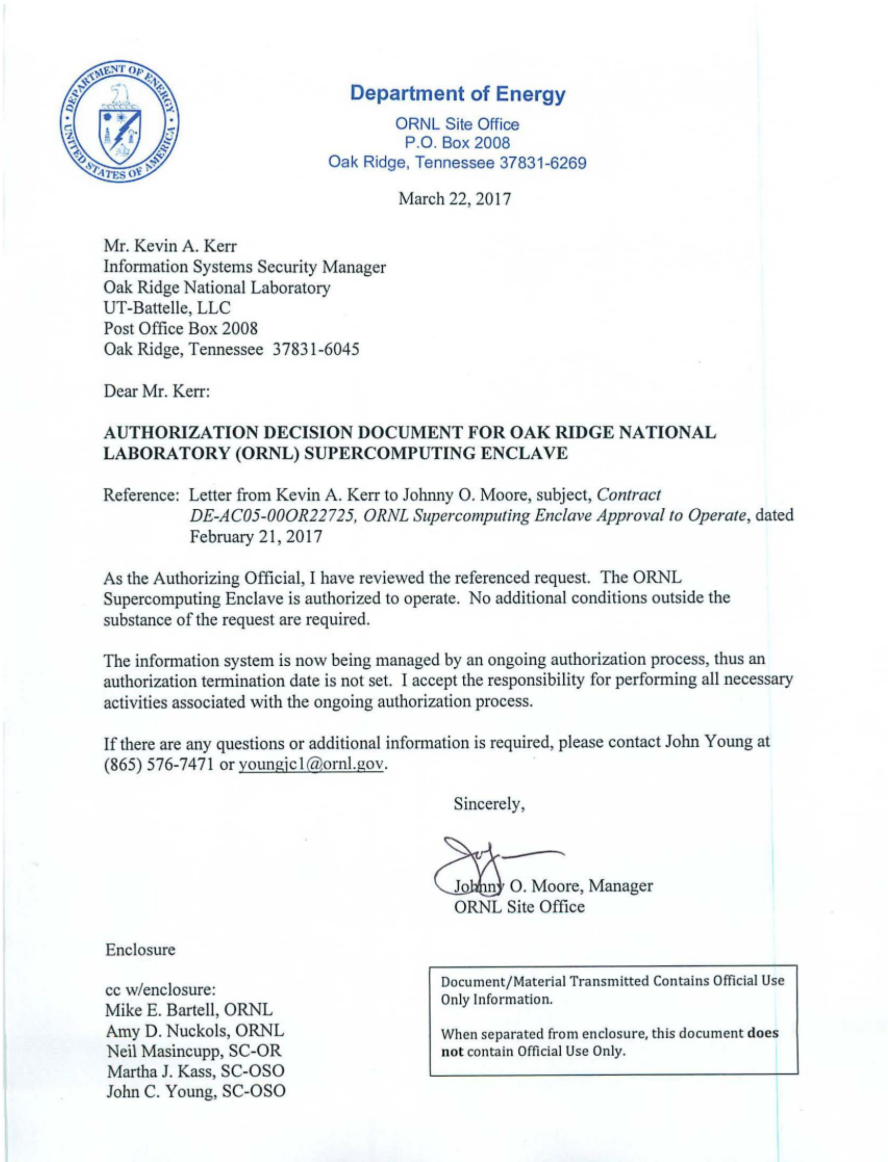
- Annual occupational exposure monitoring was conducted for noise in all three OLCF data centers. The results of these surveys are documented in the ORNL Comprehensive Tracking System. Adding new systems changed the hearing protection requirements for some of the center spaces. Additional postings were added, and the RSS was revised to include the new annual monitoring and requirements.

- The “Authorized Access to ORNL Computing Centers” access training was reviewed, and slight corrections were made to address trainee comments or fix inconsistencies.

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 information technology systems that operate for the federal government must have certification and accreditation 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 certification and 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 has 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.

In the future, cybersecurity planning will become more complex as the center continues its mission to produce great science. The facility is very proactive, viewing its cybersecurity plans as dynamic documentation to which it will preemptively respond and modify as its needs to change to provide an appropriately secure environment. The OLCF will abide 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 ITAR for projects that contain that type of sensitive information.

Over the last year, the facility has focused on implementing a new, complex processing ability—the SPI, also known as Citadel 2.0. The policy documentation and multiple third-party assessments have been completed, and operations should start in the first quarter of CY 2021. The third-party assessments granted HIPAA certifications and Cybersecurity Maturity Model Certification (CMMC) readiness to the SPI with the expectation that once the CMMCs are given, the OLCF would receive a level 3 certification.

7.2 OLCF USER VETTING

The OLCF follows a set of rigorous controls for vetting user access to ensure compliance with export-control regulations and foreign visitor access policies.

7.2.1 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 or category 2 on each project. These categories then drive the business processes that are followed for each applicant.

Table 7.1. OLCF Project Categories.

Category designation	Category description	PI actions before project activation
Category 1	The category 1 rating is applied if the project is open to 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.

Sensitive information, including proprietary and ECI, 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 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.

7.2.2 OLCF Users

All users who request access to OLCF resources are required to complete 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 is 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: o China, o Russia, o Iran, o Sudan, o Syria, o Crimea, o Cuba, and o 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 nonemployees.

[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 if 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 that users on a category 2 project are aware of the risks and rules for accessing the sensitive project.

[5] Level 2 identity proofing: The OLCF users' RSA SecurID tokens for authenticating to OLCF moderate resources. Level 2 identity proofing of all applicants is required as part of the NCCS 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) Has the facility demonstrated effective engagements with strategic stakeholders (i.e., beyond the user population)? (c) Is the facility operating in a manner that enables the delivery of the facility and DOE mission, including maintaining a vibrant US effort in science and engineering?

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 2020 OPERATIONAL ASSESSMENT GUIDANCE

The facility collects and reports annually the number of refereed publications that at least partly result from using the facility's resources. For the leadership computing facilities, tracking is done for 5 years following the project's use of the facility. This number might include publications in press, publications accepted but not submitted, or publications in preparation. This is a reported number, not a metric. Additionally, the facility may report other publications, where appropriate.

8.1.1 OLCF Publications Report

Based on a data collection completed on March 29, 2021, 427 publications that resulted from the use of OLCF resources were published in 2020. In this document, *year* refers to the CY unless it carries the prefix *FY*, indicating the fiscal year. In the 2019 OLCF OAR, 405 publications were reported. A list of 2014–2017 publications is available on the OLCF website.[23] Sponsor guidance allows accepted and in-press publications to be reported, but the OLCF only reports publications that appear 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 2012–2020 period, showing consistent growth in the total publications count and the number of publications in journals with high-impact factors. The impact on the scientific productivity of the COVID-19 pandemic has been noted

[23] https://www.olcf.ornl.gov/leadership-science/publications/

across disciplines, and much of the decrease in publications here is ascribed to this effect. Nevertheless, the number of high-impact publications shows little decrease relative to the last few years. Some of this effect is due to high-impact publications related to studies for COVID-19 pandemic response. It is also expected that the total for CY 2020 will continue to grow for several months as publications are discovered.

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

Year	Unique, confirmed OLCF publications	High-impact publications with JIF* >10
2020	427	21
2019	450	30
2018	494	20
2017	472	27
2016	459	33
2015	359	21
2014	317	16
2013	368	9
2012	347	20

*JIF = Journal impact factor

8.2 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 UA 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. Although they cannot capture the full scope and scale of achievements enabled by the OLCF in 2020, these accomplishments advance the state of the art in science and engineering R&D across diverse disciplines and are advancing DOE's science programs toward their targeted outcomes and mission goals. As an additional indication of the breadth of these achievements, OLCF users published many breakthrough publications in high-impact journals in 2020, as shown in Table 8.2.

Table 8.2. Publications in high-impact journals in 2020.

Journal	Number of publications
Science	1
Science Advances	1
Nature	1
Nature Communications	7
Nature Structural & Molecular Biology	1
Advanced Materials	1
ACS Nano	3
Chemical Reviews	1
Journal of the American Chemical Society	1
Cell	1
Nano Letters	1
Molecular Biology and Evolution	1
Genome Biology	1

Altogether in 2020, OLCF users published 54 papers in journals with a journal impact factor (JIF) of greater than 7 (a significant increase vs. the 2019 total of 43) and 21 papers in journals with a JIF greater than 10.

8.2.1 Closely Spaced Hydrogen Atoms Could Facilitate Superconductivity in Ambient Condition

PI: Timmy Ramirez-Cuesta, ORNL
Allocation Program: DD

Using neutron scattering at ORNL, researchers studied samples of zirconium vanadium hydride at atmospheric pressure and at temperatures from -450°F (5 K) to as high as -10°F (250 K)—much higher than the temperatures where superconductivity is expected to occur in these conditions. Their findings detail the first observations of such small hydrogen-hydrogen atomic distances in the metal hydride, as small as 1.6 Å compared with the 2.1 Å distances predicted for these metals. This interatomic arrangement is remarkably promising since the hydrogen contained in metals affects their electronic properties. Other materials with similar hydrogen arrangements have been found to start superconducting, but only at very high pressures.

Some of the most promising “high-temperature” superconductors, such as lanthanum decahydride, can start superconducting at about 8.0°F but unfortunately also require enormous pressures as high as 22 million pounds per square inch, or nearly 1,400 times the pressure exerted by water at the deepest part of Earth’s deepest ocean. For decades, the “holy grail” for scientists has been to find or make a material that superconducts at room temperature and atmospheric pressure, which would allow engineers to design it into conventional electrical systems and devices. Researchers are hopeful that an inexpensive, stable metal like zirconium vanadium hydride can be tailored to provide such a superconducting material.

An international team of researchers discovered that the hydrogen atoms in a metal hydride material are much more tightly spaced than was predicted for decades—a feature that could possibly facilitate superconductivity at or near room temperature and pressure. Such a superconducting material, carrying electricity without any energy loss due to resistance, would revolutionize energy efficiency in a broad range of consumer and industrial applications. The scientists conducted neutron scattering experiments at ORNL on samples of zirconium vanadium hydride at atmospheric pressure and at temperatures much higher than the temperatures where superconductivity is expected to occur in these conditions. The breakthrough in understanding occurred after the team began working with the OLCF to develop a strategy for evaluating the data. It took an ensemble of 3,200 individual simulations, a massive task that occupied around 17% of Titan's immense processing capacity for nearly a week—something a conventional computer would have required 10–20 years to accomplish. These computer simulations and additional experiments that ruled out alternative explanations proved conclusively that the unexpected spectral intensity occurs only when distances between hydrogen atoms are closer than 2.0 Å, which had never been observed in a metal hydride at ambient pressure and temperature. The team's findings represent the first known exception to the Switendick criterion in a bimetallic alloy, a rule that holds for stable hydrides at ambient temperature and pressure the hydrogen-hydrogen distance is never less than 2.1 Å.

This work was supported by ORNL's LDRD program. The research used computational resources from the OLCF, which is supported by DOE's Office of Science.

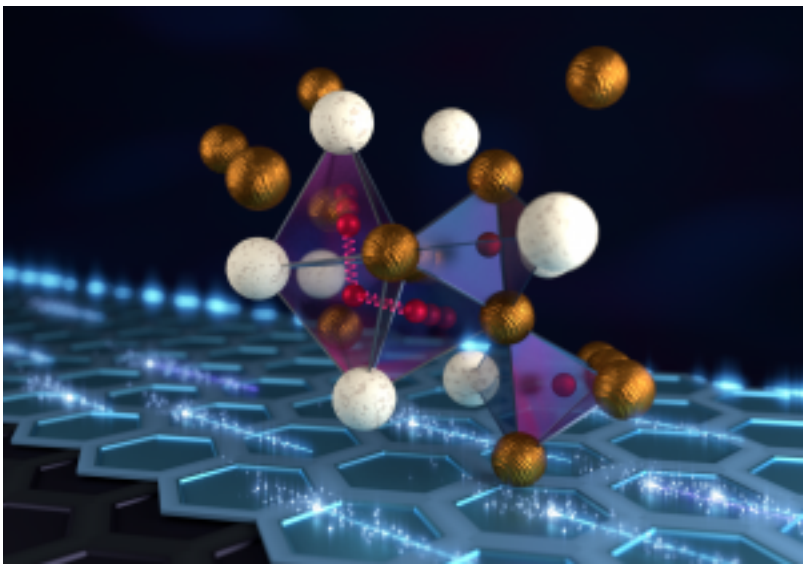


Figure 8.1. Illustration of a zirconium vanadium hydride atomic structure at near-ambient conditions as determined by using neutron vibrational spectroscopy and the Titan supercomputer at ORNL. The lattice comprises vanadium atoms (gold) and zirconium atoms (white) enclosing hydrogen atoms (red). Three hydrogen atoms are shown interacting at surprisingly small hydrogen-hydrogen atomic distances, as short as 1.6 Å. These smaller spacings between the atoms might allow packing significantly more hydrogen into the material to a point where it begins to superconduct. Credit: ORNL/Jill Hemman.

Publication: Andreas Borgschulte, Jasmin Terreni, Emanuel Billeter, Luke Daemen, Yongqiang Cheng, Anup Pandey, Zbigniew Łodziana, Russell J. Hemley, and Anibal J. Ramirez-Cuesta, "Inelastic neutron scattering evidence for anomalous H–H distances in metal hydrides," *Proceedings of the National Academy of Sciences* 117, no. 8 (2020): 4,021–4,026. doi: 10.1073/pnas.1912900117.

Related Links

- "Closely Spaced Hydrogen Atoms Could Facilitate Superconductivity in Ambient Condition," OLCF News (February 3, 2020).

- "Closely Spaced Hydrogen Atoms Could Facilitate Superconductivity in Ambient Condition," ORNL Neutron News (February 3, 2020).

- "Room Temperature Superconductor Breakthough at Oak Ridge National Laboratory," SciTech Daily (February 3, 2020).

8.2.2 ORNL Team Enlists World's Fastest Supercomputer to Combat the Coronavirus

PI: Jeremy Smith, University of Tennessee, Knoxville, and ORNL
Allocation Program: DD

The SARS-CoV-2 coronavirus, which is responsible for the COVID-19 pandemic, infects the body by one of the same mechanisms as the severe acute respiratory syndrome (SARS) virus that spread to 26 countries during the SARS epidemic in 2003. The similarity between the two virus structures inspired a research team to consider that the two viruses might even "dock" to the cell in the same way. They built a model of the coronavirus' spike protein, also called the *S-protein*, based on early studies of the structure. Using Summit, they performed MD simulations, which analyze the movements of atoms and particles in the protein. They simulated different compounds docking to the S-protein spike of the coronavirus to determine whether any might prevent the spike from sticking to human cells.

In the simulations, the compounds bind to regions of the spike that are important for entry into the human cell and thus might interfere with the infection process. Computational screening allows researchers to quickly find promising candidates for experimental studies, which are essential for verifying that certain chemicals will combat the virus. The team hopes that the computational results might inform future studies and provide a framework that experimentalists might use to further investigate the 77 small-molecule compounds.

Researchers at ORNL and the University of Tennessee, Knoxville, have used the world's most powerful and smartest supercomputer, the IBM AC922 Summit, to identify small-molecule drug compounds that might warrant further study in the fight against the SARS-CoV-2 coronavirus, which is responsible for the COVID-19 pandemic. The researchers performed simulations on Summit of more than 8,000 compounds to screen for those that are most likely to bind to the main spike protein of the coronavirus, rendering it unable to infect host cells. They found 77 compounds of interest that they believe might have value in experimental studies of the virus. They published their results on *ChemRxiv* in February.

This research was funded by the LDRD program and used resources of the OLCF.

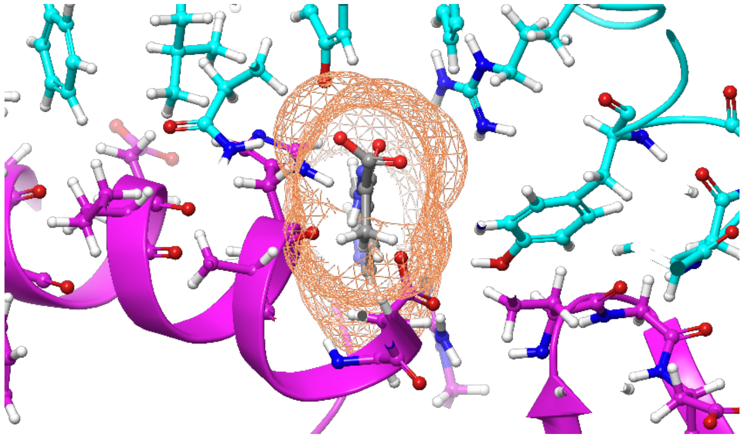


Figure 8.2. A compound (gray) was calculated to bind to the SARS-CoV-2 spike protein (cyan) to prevent it from docking to the human angiotensin-converting enzyme 2 (ACE2) receptor (purple). Credit: Micholas Smith/ ORNL, DOE.

Publication: Micholas Smith and Jeremy C. Smith, "Repurposing Therapeutics for COVID-19: Supercomputer-Based Docking to the SARS-CoV-2 Viral Spike Protein and Viral Spike Protein-Human ACE2 Interface," *ChemRxiv* (2020). doi: 10.26434/chemrxiv.11871402.v3.

Related Links:

- "ORNL Team Enlists World's Fastest Supercomputer to Combat the Coronavirus," OLCF News (March 5, 2020).
- "How Some Molecular Biophysicists Are Trying to Find a Cure for COVID-19," Rolling Stone (March 10, 2020).
- "Researchers at Oak Ridge National Lab Tap into Supercomputing to Help Combat Coronavirus," NextGov (March 11, 2020).
- "The world's fastest supercomputer identified chemicals that could stop coronavirus from spreading, a crucial step toward a vaccine," CNN (March 20, 2020).
- "The World's Fastest Supercomputer is taking on coronavirus," The Hill (March 20, 2020).

8.2.3 ORNL Scientists Tap into AI to Put a New Spin on Neutron Experiments

PI: Alan Tennant and Marcus Eisenbach, ORNL
Allocation Program: INCITE

Quantum materials—those that have correlated order at the subatomic level—have potential applications in electronic devices, quantum computers, and superconductors. Until now, researchers had not used artificial intelligence (AI) to find patterns in neutron scattering data to understand the physics inside quantum materials or complex magnetic materials. Using data from neutron scattering simulations performed on systems at the OLCF, the team trained an artificial neural network (ANN) to successfully interpret data from a neutron scattering experiment performed at ORNL's SNS.

In more than 50 billion calculations on the OLCF's Titan, Eos, and Rhea systems, a team performed scattering simulations that they then used to train an ANN. The ANN was ultimately able to interpret data recorded from the CORELLI instrument at SNS by comparing it with the simulated data, capturing the properties of $Dy_2Ti_2O_7$, a spin ice that is glass-like at low temperatures. The network can reveal new information about current scattering experiments and provide insight into which experiments would be most beneficial to run in the future. The team is now training deeper neural networks on Summit to further understand glass-like quantum materials.

For the first time, a team at ORNL is using AI to find patterns in neutron scattering data that can lead to an understanding of the physics inside quantum or complex magnetic materials. Led by Alan Tennant, initiative lead for quantum materials at ORNL, the team recently trained an ANN to successfully interpret data from a neutron scattering experiment performed at ORNL's SNS. The team trained the network by feeding it data from neutron scattering simulations performed on systems at the OLCF, including the center's decommissioned Cray XK7 Titan. One of the most powerful machines of its time, Titan continues to supply the scientific community with new discoveries even after its retirement last fall. After training, the network was able to interpret data from an actual SNS experiment, giving scientists a new way to analyze scattering experiments. The team used the Compute and Data Environment for Science at ORNL in combination with the systems at the OLCF for further analysis of the simulations.

This research was funded by the LDRD program and used resources of the OLCF.

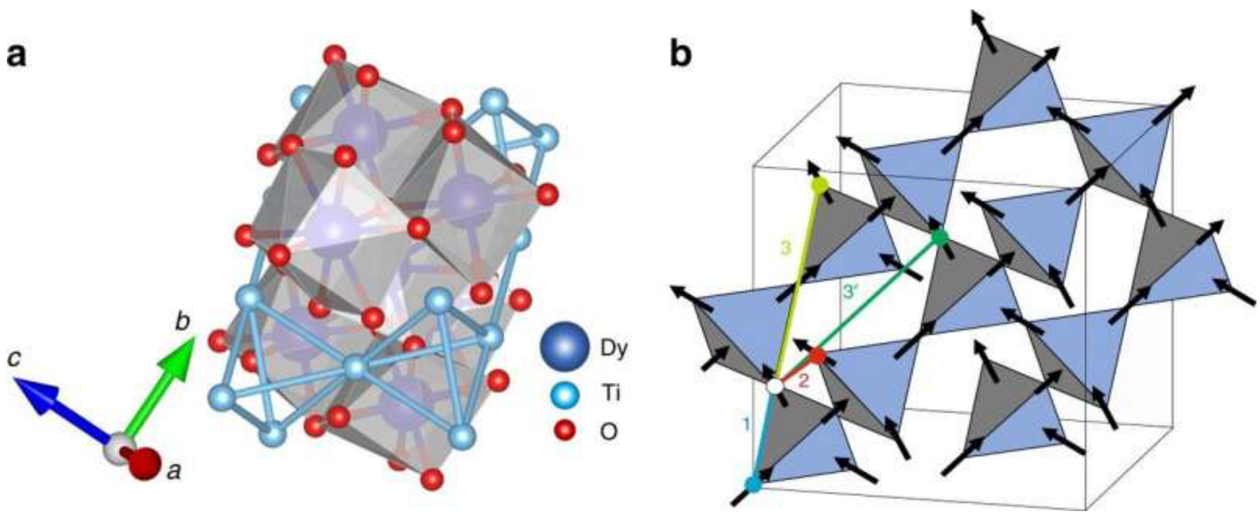


Figure 8.3. Atomic structure of $Dy_2Ti_2O_7$ is composed of tetrahedra of magnetic Dy ions (blue) and nonmagnetic octahedra of oxygen ions (red) surrounding titanium ions (cyan). Image Credit: ORNL.

Publication: Anjana M. Samarakoon et al., "Machine-Learning-Assisted Insight into Spin Ice $Dy_2Ti_2O_7$." *Nature Communications* 11 (2020): 892. doi:10.1038/s41467-020-14660-y.

Related Links:

- "ORNL Scientists Tap into AI to Put a New Spin on Neutron Experiments," OLCF News (March 27, 2020).
- "Scientists tap into AI to put a new spin on neutron experiments," *PhysOrg* (April 2, 2020).

8.2.4 Computational Gene Study Suggests New Pathway for COVID-19 Inflammatory Response

PI: Daniel Jacobson, ORNL
Allocation Program: ALCC, DD

Using the supercomputing resources of the OLCF, a team of researchers compared the genes of COVID-19 patients against a control group and analyzed population-scale gene expression data to see which genes were normally co-expressed, or turned on or off at the same time. Computational analyses suggest that genes related to one of the body's systems responsible for lowering blood pressure—the bradykinin system—appear to be excessively "turned on" in the lung fluid cells of those with the virus. Based on their analyses, the research team posits that bradykinin—the compound that dilates blood vessels and makes them permeable—is overproduced in the body of COVID-19 patients; related systems either contribute to overproduction or cannot slow the process. Excessive bradykinin leads to leaky blood vessels, allowing fluid to build up in the body's soft tissues.

In the lung fluid cells of COVID-19 patients, researchers found an increased expression of enzymes that can trigger the production of bradykinin and a decreased expression of enzymes that would break it down. They also found that some genes in the lung cells increased the production of hyaluronic acid, a substance that can trap around 1,000 times its own weight in water to form a hydrogel. The research team believes a bradykinin storm could be responsible for much of the viral pathogenesis. If the team's disease mechanism model is accurate and substantiated by experimental analysis, it might point to new drug targets worth exploration.

A team led by ORNL's Dan Jacobson used Summit to analyze genes from cells in the lung fluid of nine COVID-19 patients compared with 40 control patients and found that genes related to one of the systems responsible for lowering blood pressure—the bradykinin system—appear to be excessively "turned on" in the lung fluid cells of those with the virus. Using the Summit and Rhea supercomputers at the OLCF, the team compared the genes of COVID-19 patients against a control group and analyzed population-scale gene expression data—17,000 samples from uninfected individuals—to see which genes were normally co-expressed, or turned on or off at the same time. The Jacobson team required the power of Summit to run 2.5 billion correlation calculations that helped them understand the normal regulatory circuits and relationships for the genes of interest.

The team's research was supported by ORNL's LDRD Program; the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the CARES Act; and the National Institutes of Health.

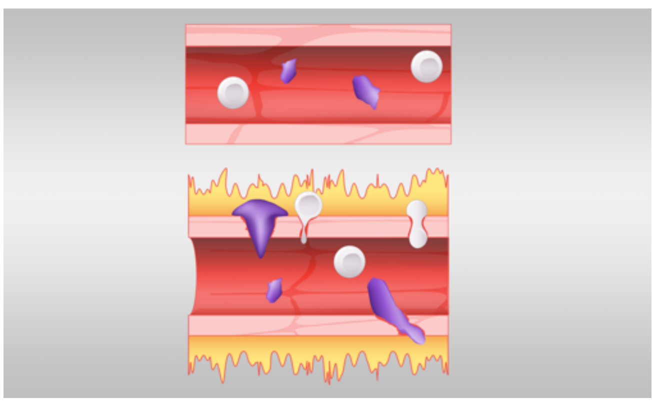


Figure 8.4. A normal blood vessel, shown at top, is compared with a blood vessel affected by excess bradykinin. A hyperactive bradykinin system permits fluid (yellow) to leak out and allows immune cells (purple) to squeeze their way out of blood vessels. Image Credit: Jason Smith/ORNL, DOE.

Publication: M. R. Garvin et al., "A Mechanistic Model and Therapeutic Interventions for COVID-19 Involving a RAS-Mediated Bradykinin Storm," *eLife* 9 (2020): e59177. doi:10.7554/eLife.59177.

Related Links:

- "Computational Gene Study Suggests New Pathway for COVID-19 Inflammatory Response," OLCF News (July 28, 2020).

- "Has the Summit Supercomputer Cracked COVID's Code?" IEEE Spectrum (August 3, 2020).

- "A Supercomputer Analyzed Covid-19—and an Interesting New Theory Has Emerged" Elemental (September 1, 2020).

8.2.5 Reaching New Heights in Weather Forecasting's Exascale Future

PI: Nils Wedi, ECMWF
Allocation Program: INCITE

Established in 1975, the European Centre for Medium-Range Weather Forecasts (ECMWF) is an intergovernmental organization composed of 34 member and cooperating countries that provides numerical weather predictions to its members while also selling forecast data to commercial weather services. The original Integrated Forecasting System (IFS) code, also known as the "European Model," was written about 30 years ago and has been one of the leading global weather-forecast systems in the world, continually updated and optimized for use on many generations of supercomputers at ECMWF. However, to advance its performance and scalability, ECMWF researchers also test their code on external HPC systems, including the OLCF's decommissioned Titan and now on Summit. Using OLCF resources in 2020, the team achieved a computational first: a global simulation of the Earth's atmosphere at a 1 km^2 average grid-spacing for a full 4 month season.

With the higher resolution of 1 km grid-spacing—enabled partly by special adaptations of the I/O scheme exploiting Summit's memory hierarchy and network and further accelerated by recoding IFS to use Summit's GPU accelerators—the team's simulations are able to represent variations in topography with finer detail. For example, IFS's current operational modeling simulates the Himalayan mountains at heights of about 6,000 m; its elevations are averaged out over the 9 km^2 grids, resulting in "smoothed out" peaks. By zooming into smaller areas with 1 km^2 grids, the averaged topography is much closer to reality—the Himalayas' peaks are now presented near their true heights of 8,000 m with much better resolved gradients representing the slopes. This increased detail results in better airflow characteristics that help determine global circulation patterns.

Using Summit, a team of researchers from ECMWF and ORNL achieved a computational first: a global simulation of the Earth's atmosphere at a 1 km^2 average grid-spacing for a full 4 month season. Completed in June 2020, the milestone marks a big improvement in resolution for the "European Model," which currently operates at 9 km grid-spacing for routine weather forecast operations. It also serves as the first step in an effort to create multi-season atmospheric simulations at high resolution, pointing toward the future of weather forecasting powered by exascale supercomputers. The team has made the simulation's data available to the international science community. By eliminating some of the fundamental modeling assumptions prevalent in conventional simulations, the high-resolution data could help improve model simulations at coarser resolutions.

The team's research was supported the European Union's Horizon 2020 future and emerging technologies and the research and innovation programs.

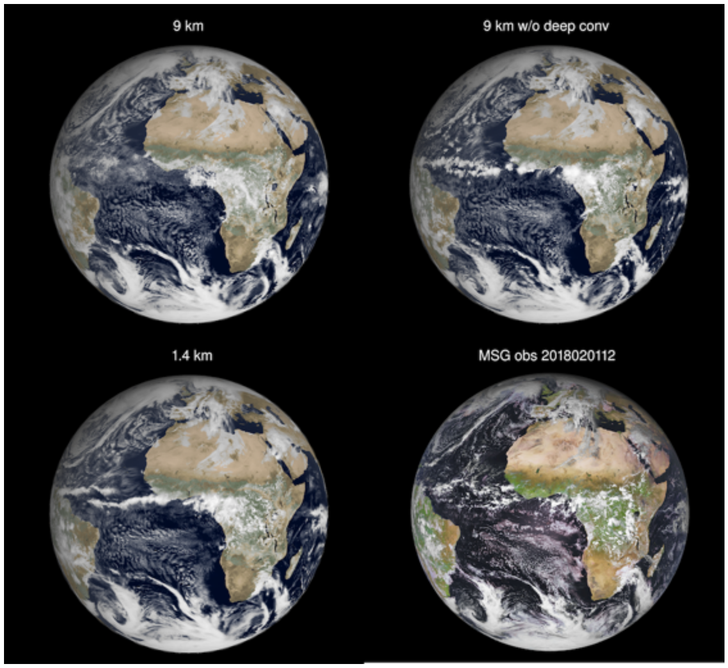


Figure 8.5. These simulated satellite images of the Earth show the improvement in resolution of the ECMWF IFS from 9 km grid-spacing with parametrized deep convection (top left), to 9 km grid-spacing (top right), and 1 km grid-spacing (bottom left). On the bottom right is a Meteosat Second Generation satellite image at the same verifying time. Image courtesy ECMWF.

Publication: N. P. Wedi, I. Polichtchouk, P. Dueben, V. G. Anantharaj, P. Bauer, S. Boussetta, et al., "A Baseline For Global Weather And Climate Simulations At 1 Km Resolution," *Journal of Advances in Modeling Earth Systems* 12 (2020): e2020MS002192. https://doi.org/10.1029/2020MS002192.

Related Links:

- "Reaching New Heights in Weather Forecasting's Exascale Future," OLCF News (September 28, 2020).
- "A baseline for global weather and climate simulations at 1 km resolution," ECMWF Science Blog (June 22, 2020).

8.2.6 DeePMD-kit: A New Paradigm for Molecular Dynamics Modeling

PI: Lin Lin, UC Berkeley
Allocation Program: DD

A UC Berkeley and LBNL-led team developed the DeePMD-kit—an "HPC + AI + Physical model" that combines HPC, AI, and physical principles to achieve speed and accuracy for MD simulations. In their project, researchers pushed the limit of molecular dynamics with ab initio accuracy to 100 million atoms with ML. Ab initio MD (AIMD) allows for greater accuracy over standard MD; however, the approach requires more computation and has thus been limited to the study of small systems (i.e., systems that have a maximum size of thousands of atoms).

The team tested DeePMD-kit on Summit by simulating a block of copper atoms with the goal of seeing how far the simulation's size and timescales could be pushed beyond AIMD's accepted limitations. The team was able to simulate a system of millions of atoms and achieved a record-breaking time-to-solution mark. The team introduced Deep Potential MD (DPMD). DPMD is a new ML-based protocol that can simulate a more than a 1 ns trajectory of over 100 million atoms per day. Although other ML-based protocols have been introduced for MD simulations in recent years, the team contends that its protocol achieves the first efficient MD simulation of 100 million atoms with ab initio accuracy.

The ACM Gordon Bell Prize, which recognizes outstanding achievement in HPC, was presented to an LBNL-led team for its successful testing of the DeePMD-kit software package, named for deep potential MD. This software package bridges classical MD and AIMD to produce complex simulations that are both large and accurate for the first time. The team tested DeePMD-kit on Summit by simulating a block of copper atoms with the goal of seeing how far the simulation's size and timescales could be pushed beyond AIMD's accepted limitations. The team was able to simulate a system of 127.4 million atoms—more than 100 times larger than the current state of the art. Furthermore, the simulation achieved a time-to-solution mark of at least 1,000 times faster at 2.5 ns per day for mixed-half precision with a peak performance of 275 petaflops for mixed-half precision. The great accomplishment of this work is that it opens the door to simulating unprecedented size and timescales with ab initio accuracy. It also poses new challenges to the next-generation supercomputer for a better integration of ML and physical modeling.

The team's research was partially supported by the National Science Foundation and DOE.

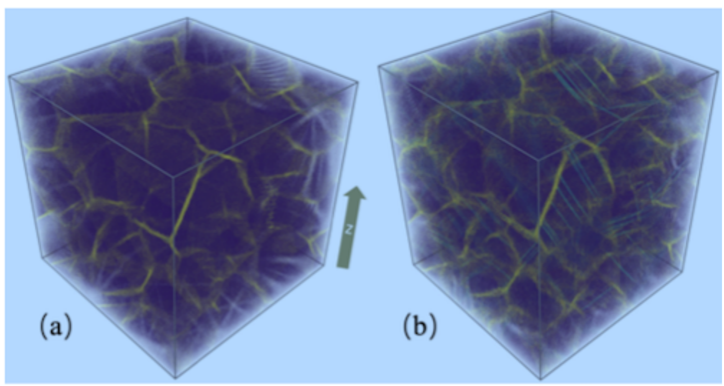


Figure 8.6. DeePMD-kit simulated a block of copper with a system of 127.4 million atoms—more than 100 times larger than the current state of the art. Credit: DeePMD Team

Publication: Weile Jia, Han Wang, Mohan Chen, Denghui Lu, Lin, Roberto Car, Weinan E, and Linfeng Zhang, "Pushing The Limit Of Molecular Dynamics With Ab Initio Accuracy To 100 Million Atoms With Machine Learning," *International Journal of High Performance Computing Applications* (2020).

Related Links:

- "Four Teams Using ORNL's Summit Supercomputer Named Finalists in 2020 Gordon Bell Prize," OLCF News (November 10, 2020).
- "Distinct ACM Gordon Bell Prizes Awarded to Two Teams Using the Summit Supercomputer," OLCF News (November 19, 2020).
- "2020 ACM Gordon Bell Prize Awarded to Team for Machine Learning Method that Achieves Record Molecular Dynamics Simulation," ACM News (November 19, 2020).

8.2.7 Gordon Bell Special Prize Winning Team Reveals AI Workflow for Molecular Systems in the Era of COVID-19

PI: Rommie Amaro, UC San Diego
Allocation Program: DD

Imaging techniques, such as x-ray imaging and cryogenic electron microscopy, can provide snapshots of viruses, such as SARS-CoV-2, but these fall short of capturing the dynamic movements of viral proteins. However, computer simulations can help scientists capture the movements of these structures virtually. Now, a team led by Rommie Amaro at UC San Diego and Arvind Ramanathan at ANL have built a first-of-its-kind workflow based on AI and have run it on the OLCF's Summit supercomputer to simulate the virus's spike protein in numerous environments.

The team was able to successfully scale Nanoscale MD (NAMD) to 24,576 of Summit's NVIDIA V100 GPUs. The results of the team's initial runs on Summit have led to discoveries of one of the mechanisms that the virus uses to evade detection and a characterization of interactions between the spike protein and the protein that the virus takes advantage of in human cells to gain entrance into them—the ACE2 receptor. The team is now integrating its scientific code NAMD into the workflow pipeline to fully automate the transition from simulation to AI for data processing without gaps.

The ACM Gordon Bell Special Prize for High-Performance Computing-Based COVID-19 Research was presented to a UC San Diego-led team for a first-of-its-kind workflow based on AI, and the team ran the code on the OLCF's Summit supercomputer to simulate the virus's spike protein in numerous environments, including within the SARS-CoV-2 viral envelope comprising 305 million atoms—the most comprehensive simulation of the virus performed to date. The team first optimized the NAMD and the Visual MD codes, which model the movements of atoms in time and space, on multiple smaller cluster systems: the Frontera supercomputer at the Texas Advanced Computing Center, the CoMet system at the San Diego Supercomputer Center, and the ThetaGPU supercomputer at ALCF. After code optimizations, the team was able to successfully scale NAMD to 24,576 of Summit's NVIDIA V100 GPUs. The results of the team's initial runs on Summit have led to discoveries of one of the mechanisms that the virus uses to evade detection and a characterization of interactions between the spike protein and the protein that the virus takes advantage of in human cells to gain entrance into them—the ACE2 receptor. By layering the experimental data and the simulation data and combining them with their AI-based approach, the researchers were able to capture the virus and its mechanisms in unprecedented detail.

This research was supported by the ECP; the DOE National Virtual Biotechnology Laboratory, with funding provided by the CARES Act; and the COVID-19 HPC Consortium.

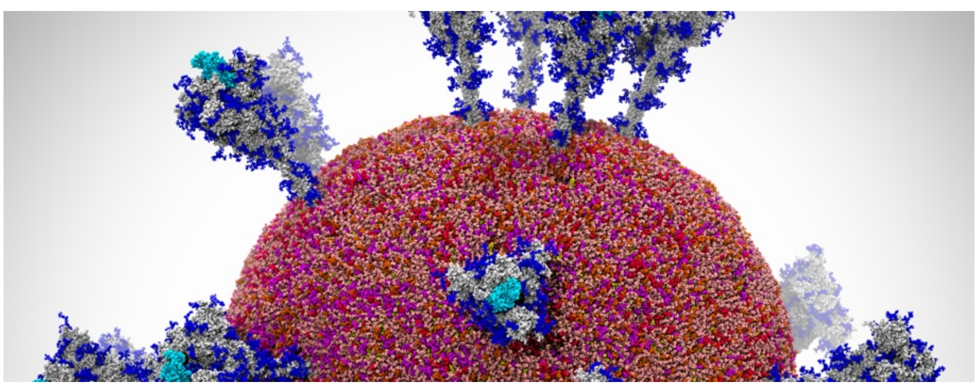


Figure 8.7. A visualization of the SARS-CoV-2 viral envelope comprising 305 million atoms.
Image Credit: Rommie Amaro, UC San Diego; Arvind Ramanathan, ANL.

Publication: Lorenzo Casalino, Abigail Dommer, Zied Gaieb, Emilia P. Barros, Terra Sztain, Surl-Hee Ahn, Anda Trifan, Alexander Brace, Anthony Bogetti, Heng Ma, Hyungro Lee, Matteo Turilli, Syma Khalid, Lillian Chong, Carlos Simmerling, David J. Hardy, Julio D. C. Maia, James C. Phillips, Thorsten Kurth, Abraham Stern, Lei Huang, John McCalpin, Mahidhar Tatineni, Tom Gibbs, John E. Stone, Shantenu Jha, Arvind Ramanathan, and Rommie E. Amaro, "AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics," *International Journal of High Performance Computing Applications*, accepted 2020.

Related Links:

- "Distinct ACM Gordon Bell Prizes Awarded to Two Teams Using the Summit Supercomputer," OLCF News (November 19, 2020).

- "UC San Diego Leads Research that Earns Gordon Bell Special Prize," UCSD News (November 19, 2020).

- "First ACM Gordon Bell Special Prize for HPC-Based COVID-19 Research Awarded," ACM News (November 19, 2020).

8.2.8 INCITE 2020 Allocation/Utilization

The INCITE allocation year is January 1–December 30. In 2020, all INCITE projects were allocated on Summit. Usage is shown in Figure 8.8. Note that when comparing these allocations and usage statistics, the allocation units changed fundamentally as INCITE projects moved from Titan to Summit in CY 2019 (i.e., Titan core-hours vs. Summit node-hours).

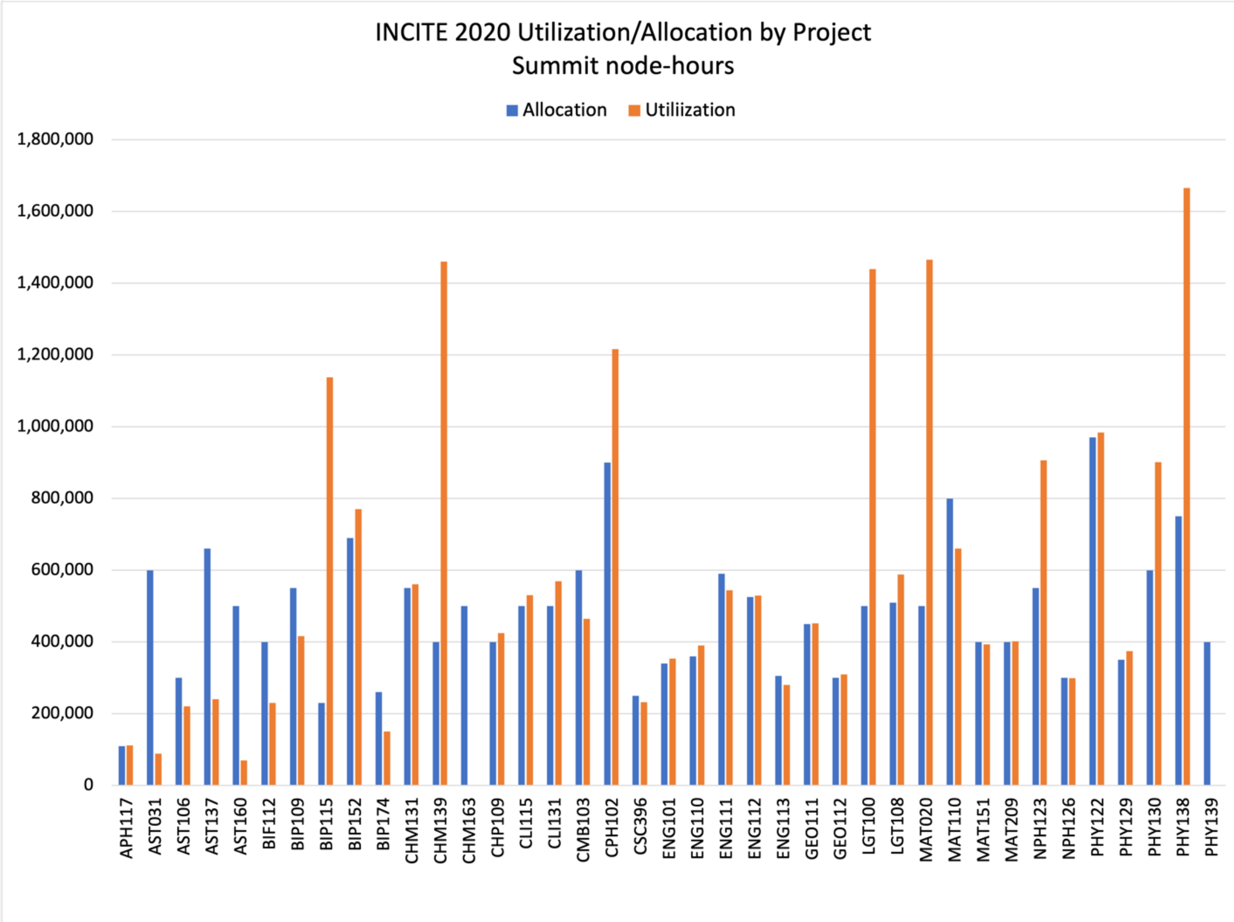


Figure 8.8. INCITE allocation by project on Summit.

8.2.9 ALCC Allocation/Utilization for CY 2018

The ALCC allocation year is July 1–June 30.Usage for both programs is reported against the full allocation amount for each allocation year in Figure 8.9. As with INCITE, the allocation units differ in nature from the historical values on Titan.

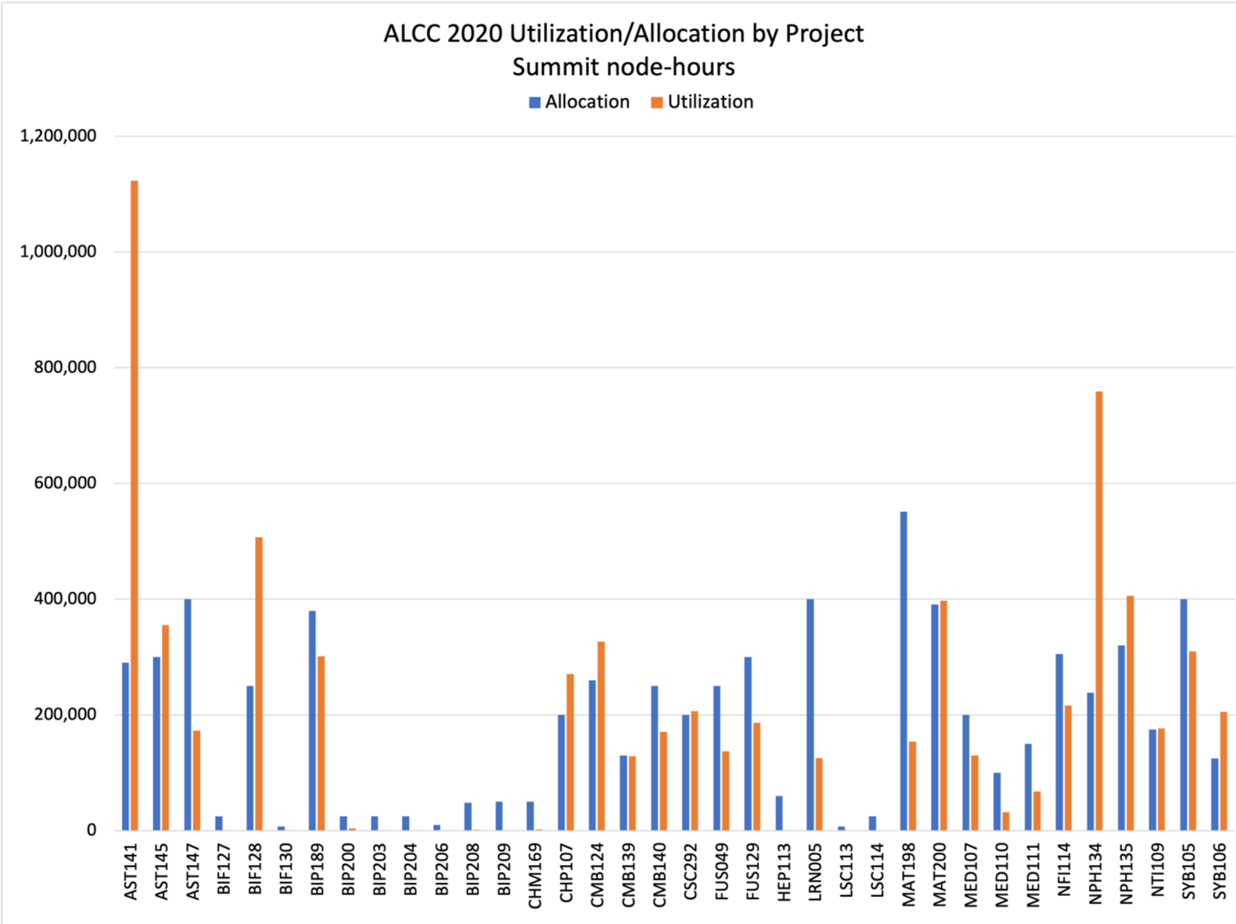


Figure 8.9. ALCC allocation by project on Summit.

8.3 STAKEHOLDER ENGAGEMENT AND OUTREACH

8.3.1 Community Engagement

A primary and natural place for community engagement has been DOE's ECP, whose goal 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 to produce usable software and applications and influence the development of hardware technology for the exascale systems in the 2021–2022 time frame. These investments are very timely and will significantly help the OLCF deliver 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 as follows, as well as many pairwise and other interactions (e.g., staff involvement in ECP AD and Software Technology [ST] projects).

8.3.1.1 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). The OLCF 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 chose 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 so 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 the OLCF Scientific Computing liaison are listed in Table 8.3.

Table 8.3. OLCF ECP engagement applications, the ECP AD PI, and the OLCF Scientific Computing liaison.

Work Breakdown Structure/ ECP AD project	Project PI	OLCF liaison
2.2.1.01 LatticeQCD	Andreas Kronfeld (Fermilab)	TBD
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)	Ronnie Chatterjee
2.2.2.03 ExaSMR	Steven Hamilton (ORNL)	Mark Berrill
2.2.2.05 WDMApp	Amitava Bhattacharjee (Princeton Plasma Physics Laboratory)	Ed D’Azevedo
2.2.3.01 ExaStar	Dan Kasen (LBNL)	Austin Harris
2.2.3.02 ExaSky	Salman Habib (ANL)	Bronson Messer
2.2.3.05 E3SM-MMF	Mark Taylor (Sandia)	Matt Norman
2.2.4.02 ExaSGD	Slaven Peles (Pacific Northwest National Laboratory)	Philip Roth
2.2.4.04 ExaBiome	Kathy Yelick (LBNL)	Philip Roth

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 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. Additionally, because the AD project depends 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 the 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 was further diversified by including these additional projects in the CAAR program.

8.3.1.2 OLCF-ECP Training Program

The OLCF continued its training engagements with ECP in 2020. The OLCF staff are active participants in the ECP Training Advisory Group, which meets monthly to discuss training activities under development at each of the six core ECP laboratories and to identify possible activities for collaboration. The OLCF worked with the ECP to cohost and facilitate training events in 2020. One such example was the OLCF/ECP Kokkos Training bootcamp held in April 2019. The OLCF worked with the ECP Kokkos team to host this bootcamp attended by 68 people. On September 3, the OLCF in conjunction with the Frontier COE hosted a Frontier COE Workshop for CAAR and ECP personnel who were current users of the COE’s testbed systems Poplar (CAAR) and Tulip (ECP). During the workshop, members of the

Frontier COE gave updates on the latest system hardware and software, and selected CAAR and ECP teams gave lessons learned talks regarding their experiences porting/running on the systems. The workshop was restricted to participants covered under existing CORAL-2 NDAs. In total, 192 participants attended the event.

8.3.1.3 OLCF-ECP Continuous Integration

In 2020 while working with the ECP, OLCF explored the Extreme Scale Scientific Software Stack (E4S) as a vehicle for widespread build and smoke testing of ECP ST projects. E4S is a curated set of software packages for developing, deploying, and running scientific applications on HPC platforms, and it contains almost all of the ECP ST software libraries. CI pipelines were deployed on Ascent to build and test this stack regularly to send pass/fail results to a dashboard accessible to ECP software development teams. In this way, basic tests can be performed regularly by facility staff after changes are made to the system as a "sanity check" for the latest software build recipes available to E4S. In addition to E4S software testing, 25 individual ECP AD and ST projects—such as LLVM, VTK-m, HDF5, GAMESS, and ExaSGD—have developed and run individual bespoke CI pipelines on top of the GitLab infrastructure at the OLCF for more advanced and customized testing at the individual project level. These CI pipelines use Ascent, which represents the same hardware and software environment as Summit but in a more easily accessible environment.

8.3.2 Industry Engagement

ACCEL continues to attract large-scale industrial problems that require access to leadership-scale systems to make progress. ACCEL had a strong year despite the impact of the global pandemic on businesses. Thirty-four industry projects were underway during the year: 21 new projects and 13 projects that began before 2020 and continued throughout the year.

- The 34 industrial projects underway in 2020 represented 8% of the total number of projects provided to external user programs, INCITE, ALCC, and DD (including ECP).
- These projects used 2,502,705 Summit node-hours, representing approximately 7% of the total Summit hours.
- 39% of the total industrial project hours on Summit were allocated through INCITE, 47% were through ALCC, and 14% were through DD.
- Of the 21 new projects, three received their awards via INCITE, five received their awards via ALCC, and 13 received their awards via DD.

8.3.2.1 Observations about the Industrial Projects

- Industry received more INCITE awards this year (three) than any year since ACCEL began.
- Numeca, one of the INCITE awardees, is the first independent software vendor to receive an INCITE award at the OLCF.
- Two companies that are also small businesses were new to ACCEL: Berg Health and Entos Inc.
- General Motors and ORNL collaborated to receive their second ALCC award to further develop simulation environments for autonomous vehicles.

- Three companies received DD allocations to accelerate drug discovery research to treat COVID-19: (1) Berg Health, a clinical-stage biotech company that uses AI to research diseases and develop innovative treatments; (2) Entos Inc., a Pasadena-based startup that provides tools for molecular and materials discovery through the use of physics-based ML; and (3) IBM.

- General Electric (GE), through its INCITE project, realized breakthrough insights into the key physics limiting performance in high-pressure turbines, prompting the firm to send a letter to the director of DOE's Office of Science, thanking DOE for access to Summit. Arjan Hegeman, the general manager of advanced technology operation at GE Aviation, said that the new understandings they gained will help the firm "…deliver more fuel-efficient, durable engines for the aircraft industry to reduce costs for airlines and their customers and lessen the environmental impact from emissions. Furthermore, your support is particularly appreciated in these difficult times for the aviation business when we are trying to stem the economic impact of COVID-19 without retreating from advancing technology."

8.3.2.2 GE Uses Summit to Perform Unprecedented Gas Turbine Simulations

GE researchers used Summit to perform unprecedented large eddy simulations (LESs) to predict the flow and temperature of gases in the high-pressure gas turbines that form the core of airplane engines. The stakes are high, competition is fierce, and customers demand the most efficient and durable engines that produce the fewest emissions possible. Accurately predicting the temperature and motion of the fluids in these high-pressure turbines is critical. Internal turbine temperatures become so high during operation that turbine components will melt if they are not cooled properly.

Adding cooling flow on the surface of these components can solve this melting problem, but knowing where to add that flow is complex and requires a deep understanding of the physics involved. Since the 1990s, scientists have relied on computer models to simulate the fluid flows by using techniques called Reynolds-Averaged Navier-Stokes (RANS) and LES. RANS is less computationally intensive, but simulation results are less detailed. LES produces much higher fidelity but is very computationally intensive. Because of the time required to do these calculations, even on large supercomputers, engineers have been constrained to applying LES on representative problems, such as a small part of a turbine blade.

Through DOE's INCITE program, GE was awarded a grant of time on Summit to run these LES calculations. The team used GE's GPU-based GENESIS code to simulate a rotor/stator combination of the high-pressure turbine that is directly downstream of the engine's combustor. The components experience extreme heat from the burnt fuel in the combustor. These heat surges result in significant differences in temperature among the different components of the turbine. Cool fluid is injected to mitigate the temperature variance before cracking occurs.

Michal Osusky, a senior engineer in CFD and methods at GE Research who performed the simulations on Summit, explained that "this is the first time we were able to look at larger 3D flow interactions with this level of detail. Summit allows us to perform LES with extraordinarily fast turnaround, which in turn enables us to solve the underlying equations of motion with far less approximations than we've done before. This particular problem was solved with hardly any approximations at all. We are getting closer and closer to reality with this computation."

The GE Research team performed these LESs of turbine components so quickly that the technique could influence the design process instead of being used solely in a research environment. "Calculations that previously took weeks or months were completed in design-cycle time on Summit. This opens opportunities for LES to become a design tool for GE's engineering work," said Sriram Shankaran, a consulting engineer at GE Aviation. Figure 8.10 compares visualizations of a lower-resolution model of turbulent flow and a higher-resolution model of turbulent flow.

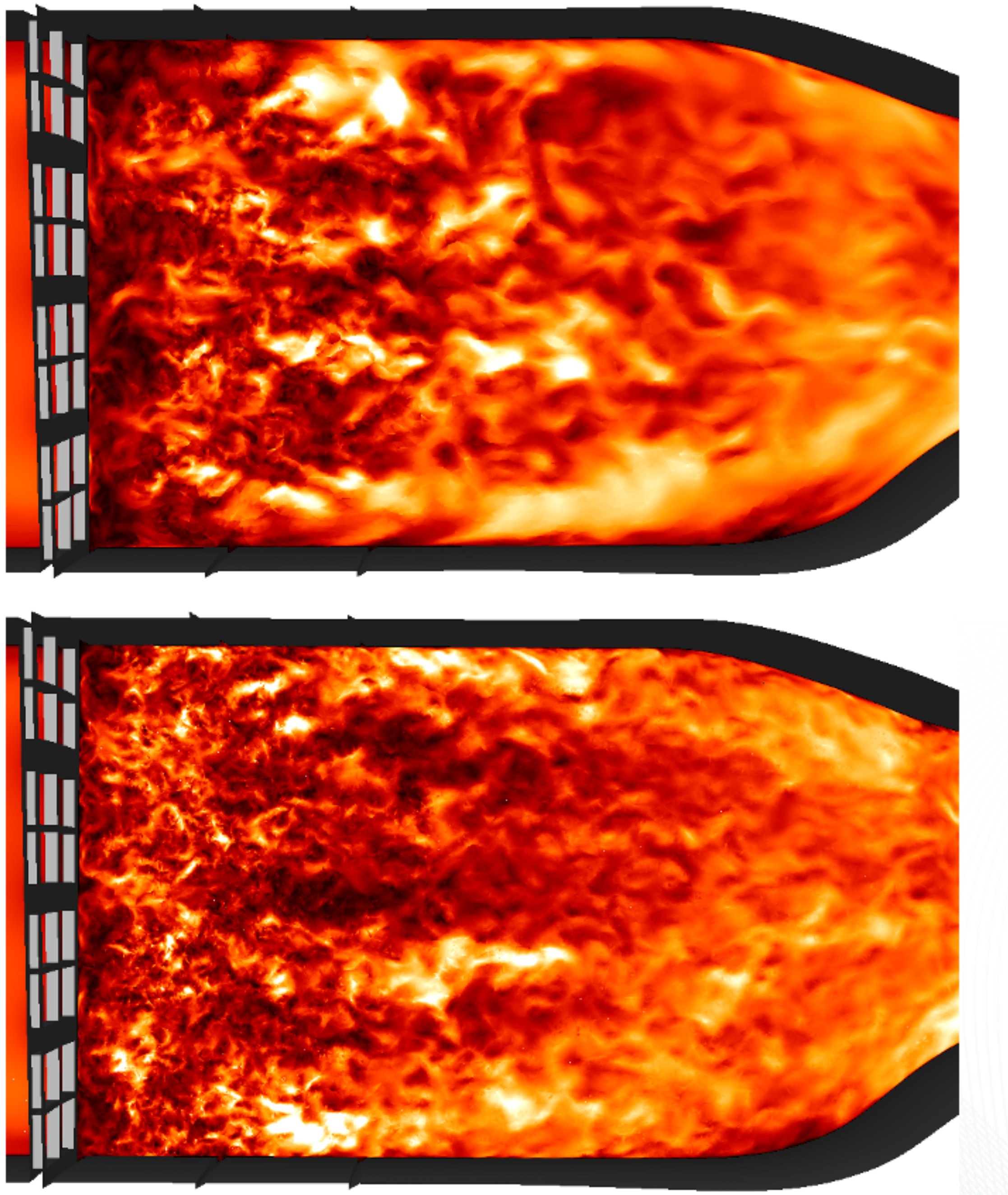


Figure 8.10. GE visualizations of a lower-resolution model of turbulent flow (top) compared with a higher-resolution model of turbulent flow (bottom) only made possible by using the Summit supercomputer. Image Credit: GE Research.

APPENDIX A. RESPONSES TO RECOMMENDATIONS FROM THE 2019 OPERATIONAL ASSESSMENT REVIEW

In April 2020, the operational activities of the OLCF for CY 2019 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 2020 OLCF training and outreach events.

Table B.1. 2020 OLCF training and outreach events.

Event type	Event title	Date	Participants
Monthly user conference call	2020 OLCF User Conference Call: myOLCF	1/29/20	46
Monthly user conference call	2020 OLCF User Conference Call: OLCF Overview for New Users	2/26/20	33
Monthly user conference call	2020 OLCF User Conference Call: Job Step Visualizer	3/25/20	53
Monthly user conference call	2020 OLCF User Conference Call: AI-Driven Multi-Modal HPC Workflow for Resolving Conformational Features of Phase Shifting	4/29/20	71
Monthly user conference call	2020 OLCF User Conference Call: SLATE Container Orchestration Service	6/24/20	95
Monthly user conference call	2020 OLCF User Conference Call: Kokkos	7/29/20	69
Monthly user conference call	2020 OLCF User Conference Call: Archival Storage Survey	8/26/20	59
Monthly user conference call	2020 OLCF User Conference Call: Constellation	9/30/20	26
Monthly user conference call	2020 OLCF User Conference Call: CUDA 11 Features	10/21/20	26
Monthly user conference call	2020 OLCF User Conference Call: Open CE	12/9/20	24
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Refactoring EXAALT MD for Emerging Architectures	1/15/20	49
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Introduction to Kokkos	2/19/20	109
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Testing: Strategies When Learning Programming Models and Using High-Performance Libraries	3/18/20	59
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Best Practices for Using Proxy Applications as Benchmarks	4/15/20	72
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Multi-Precision Algorithms	5/13/20	143
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: SYCL-Introduction and Best Practices	6/17/20	335
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: What's New in Spack?	7/15/20	265
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Colormapping Strategies for Large Multivariate Data in Scientific Applications	8/12/20	202
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Testing and Code Review Practices in Research Software Development	9/9/20	270
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Scalable Precision Tuning of Numerical Software	10/14/20	141

Event type	Event title	Date	Participants
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Reducing Technical Debt with Reproducible Containers	11/4/20	179
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Software Design for Longevity with Performance Portability	12/9/20	171
NVIDIA CUDA Training Series	Introduction to CUDA C++	1/15/20	191
NVIDIA CUDA Training Series	CUDA Shared Memory	2/19/20	88
NVIDIA CUDA Training Series	Fundamental CUDA Optimization (Part 1)	3/18/20	95
NVIDIA CUDA Training Series	Fundamental CUDA Optimization (Part 2)	4/16/20	91
NVIDIA CUDA Training Series	Atomics, Reductions, and Warp Shuffle	5/13/19	83
NVIDIA CUDA Training Series	CUDA Managed Memory	6/18/20	71
NVIDIA CUDA Training Series	CUDA Concurrency	7/21/20	94
NVIDIA CUDA Training Series	GPU Performance Analysis	8/18/20	56
NVIDIA CUDA Training Series	CUDA Cooperative Groups	9/17/20	53
Workshop/training meeting	Scaling Up Deep Learning Applications on Summit	2/10/20	113
Workshop/training meeting	JSRUN Tutorial	2/18/20	54
Workshop/training meeting	NVIDIA Profiling Tools—NSIGHT Systems	3/9/20	102
Workshop/training meeting	NVIDIA Profiling Tools—NSIGHT Compute	3/10/20	87
Workshop/training meeting	OpenACC Training Series: Introduction to OpenACC	4/17/20	43
Workshop/training meeting	KOKKOS Training Bootcamp	4/21–24/2020	68
Workshop/training meeting	Quantum Computer User Forum	4/21–23/2020	243
Workshop/training meeting	2020 Summit New User Training	5/3/20	181
Workshop/training meeting	2020 OLCF User Meeting	5/4/20	199
Workshop/training meeting	2021 INCITE Proposal Writing Webinar	5/5/20	74
Workshop/training meeting	2020 OpenMP Hackathon	8/3–7/2020	56
Workshop/training meeting	OpenACC Training Series: OpenACC Data Management	5/28/20	51

Event type	Event title	Date	Participants
Workshop/training meeting	2021 INCITE Proposal Writing Webinar	6/5/20	94
Workshop/training meeting	KOKKOS Training Webinar	6/16–18/2020	80
Workshop/training meeting	OpenACC Training Series: Loop Optimizations with OpenACC	6/23/20	54
Workshop/training meeting	PITCH IT Workshop for ORNL Summer Interns	7/27/20	N/A
Workshop/training meeting	TAU Performance Analysis Training	7/28/20	46
Workshop/training meeting	Frontier COE Workshop	9/03/20	192
Workshop/training meeting	TAPIA 2020: Easy-Peasy GPU Programming Using OpenMP	9/16/20	8
Workshop/training meeting	CARLA 2020—GPU Accelerated Computing with OpenACC	9/17/20	12
Workshop/training meeting	Introduction to Deep Learning for Scientific Applications	10/7–8/2020	100
Workshop/training meeting	Workshop on GPU Accelerated Data Analytics Virtual Lab on Summit: From Drug Discovery to Radiation Physics	10/15/20	54
Workshop/training meeting	SC20 Hands-On With Summit	11/5/20	60
Workshop/training meeting	Using the New SPEC HC2021 Scientific Application Benchmark Suite	11/9/20	50
Workshop/training meeting	NVIDIA HPC SDK—OpenMP Target Offload Training	12/8–10/2020	260
Hackathons	San Diego Supercomputing Center Hackathon	5/11–13/2020	79
Hackathons	Princeton Hackathon	6/8–10/2020	84
Hackathons	NERSC Hackathon	7/13–15/2020	78
Hackathons	Brookhaven National Laboratory GPU Hackathon	8/17–19/2020	73
Hackathons	OLCF Hackathon	10/26–30/2020	54
Seminar series	Research Scientist in HPC and Learning for Discrete Computing Systems	2/20/20	N/A
Seminar series	Applying Statistical Mechanics to Improve Computational Sampling Algorithms and Interatomic Potentials	2/24/20	N/A
Seminar series	From Computational Methods to Materials: Insights from the Electronic Structure	2/27/20	N/A
Seminar series	Providing Insight into the Performance of Distributed Applications Through Low-Level Metrics	3/5/20	N/A
Seminar series	Exploring the Landscape of Big-Data Analytics through Domain-Aware Incremental and Approximate Algorithm Design	3/26/20	N/A

Event type	Event title	Date	Participants
Seminar series	Science as the Mission of the OLCF: Where We've Been and Where We Are (Perhaps) Going	4/7/20	N/A
Seminar series	Lessons Learned of Directive-Based Programming Models for Accelerators	4/14/20	N/A
Seminar series	Cross-Facility Science: Accelerating Scientific Discovery at NERSC	4/14/20	N/A
Seminar series	Describing Quadruple Collective Excitations of Nuclei within Self-Consistent Methods	4/16/20	N/A
Seminar series	Complex-Energy Description of Open Atomic and Nuclear Quantum Systems	4/17/20	N/A
Seminar series	Study of Halo Nuclei with Covariant Densities Functional Theories and Beyond Mean Field Approach	4/24/20	N/A
Seminar series	Combining Data and Dynamics: Successes and Open Challenges	5/14/20	N/A
Seminar series	Algorithm Challenges and Improvements in the CompFUSE SciDAC-4 Project	6/11/20	N/A
Seminar series	Explicit and Implicit Data Analysis for Recommendation	6/12/20	N/A
Seminar series	Sales Elasticity of Emotional Displays: Large Scale Evidence for Selling with a Straight Face	6/26/20	N/A
Seminar series	SNAPSKETCH: Graph Representation Approach for Intrusion Detection in a Streaming Graph	6/29/20	N/A
Seminar series	Data-driven Concrete Damage Diagnosis with Thermal Imaging and Vibration Testing	7/6/20	N/A
Seminar series	Accelerating Multiscale Materials Modeling with Machine Learning	7/10/20	N/A
Seminar series	Polynomial-Time Estimation of a Gromo-Hausdroff Distance	7/14/20	N/A
Seminar series	HPC Storage Analytics—Methods and Tools	7/14/20	N/A
Seminar series	Data Science in Machine Learning to Predict Degradation and Power of Photovoltaic Systems	7/16/20	N/A
Seminar series	Preference-based "Nested" Network Mode	8/31/20	N/A
Seminar series	Update on Arm in HPC: Business and Technical POVs	10/30/20	N/A
Seminar series	Developing and Using Code for HPC: Towards Efficient Binary Neutron Star Simulations in Spherical Coordinates with SphericalNR	12/17/20	N/A

APPENDIX C. OUTREACH PRODUCTS

Date	Product type	Title
1/27/20	Poster	ECP Annual Meeting (SLATE)
1/2/20	Highlight	New Year Brings New Directory Structure for OLCF’s High-Performance Storage System
1/2/20	Highlight	Speeding Toward the Future of Fusion
1/2/20	Highlight	The Last User of Titan
1/2/20	Highlight	Big Iron Afterlife: How ORNL’s Titan Supercomputer was Recycled
1/2/20	PPT Slide	New Year Brings New Directory Structure for OLCF’s High-Performance Storage System
1/2/20	PPT Slide	Speeding Toward the Future of Fusion
1/2/20	PPT Slide	The Last User of Titan
1/2/20	PPT Slide	Big Iron Afterlife: How ORNL’s Titan Supercomputer Was Recycled
1/17/20	Highlight	Metal Alloys for Industrial Applications Benefit from Supercomputer Modeling
1/17/20	Highlight	OLCF is Looking into the Present and the Future of Large-Scale Data Science
1/17/20	Highlight	OLCF Staff Involved in Khronos Effort to Create New Visualization Standard
1/17/20	Highlight	Eight Teams Successfully Running on First NVIDIA and ARM Test Bed, Wombat at the OLCF
1/17/20	PPT Slide	Metal Alloys for Industrial Applications Benefit from Supercomputer Modeling
1/17/20	PPT Slide	OLCF is Looking into the Present and the Future of Large-Scale Data Science
1/17/20	PPT Slide	OLCF Staff Involved in Khronos Effort to Create New Visualization Standard
1/17/20	PPT Slide	Eight Teams Successfully Running on First NVIDIA and ARM Test Bed, Wombat at the OLCF
1/21/20	Highlight	ORNL Project Demonstrates Viability of “Smart” Traffic Cameras to Save Fuel
1/21/20	PPT Slide	ORNL Project Demonstrates Viability of “Smart” Traffic Cameras to Save Fuel
1/26/20	Highlight	Making Room for Frontier
1/26/20	PPT Slide	Making Room for Frontier
2/3/20	Highlight	Closely Spaced Hydrogen Atoms Could Facilitate Superconductivity in Ambient Condition
2/3/20	PPT Slide	Closely Spaced Hydrogen Atoms Could Facilitate Superconductivity in Ambient Condition
2/5/20	Fact Sheet	Frontier Fact Sheet
2/21/20	Highlight	Machine Learning for Better Drug Design
2/21/20	PPT Slide	Machine Learning for Better Drug Design
3/5/20	Highlight	Simulating the Stars at Exascale Requires Hip Solutions
3/5/20	Highlight	Rethinking a Century of Fluid Flows
3/5/20	Highlight	“Bubbles” Demonstrates the Power of GPUS, Immersion Cooling
3/5/20	Highlight	ORNL Team Enlists World’s Fastest Supercomputer to Combat the Coronavirus
3/5/20	PPT Slide	Simulating the Stars at Exascale Requires Hip Solutions
3/5/20	PPT Slide	Rethinking a Century of Fluid Flows
3/5/20	PPT Slide	“Bubbles” Demonstrates the Power of GPUS, Immersion Cooling
3/5/20	PPT Slide	ORNL Team Enlists World’s Fastest Supercomputer to Combat the Coronavirus
3/18/20	Highlight	COVID-19: OLCF Continues Normal Operations

Date	Product type	Title
3/18/20	PPT Slide	COVID-19: OLCF Continues Normal Operations
3/27/20	Highlight	New Job Step Viewer Tool Captures Job Launching on Summit
3/27/20	Highlight	ORNL Scientists Tap Into AI to Put a New Spin on Neutron Experiments
3/27/20	Highlight	OLCF Offers New Workload Capabilities with SLATE Service
3/27/20	PPT Slide	New Job Step Viewer Tool Captures Job Launching on Summit
3/27/20	PPT Slide	ORNL Scientists Tap Into AI to Put a New Spin on Neutron Experiments
3/27/20	PPT Slide	OLCF Offers New Workload Capabilities with SLATE Service
3/30/20	Highlight	Summit Joins the COVID-19 High Performance Computing Consortium
3/30/20	PPT Slide	Summit Joins the COVID-19 High Performance Computing Consortium
4/15/20	Highlight	U.S. Department of Energy's INCITE Program Seeks Proposals for 2021
4/15/20	PPT Slide	U.S. Department of Energy's INCITE Program Seeks Proposals for 2021
4/29/20	Highlight	OLCF Staff and Researchers Make a Smooth Transition to Working Remotely
4/29/20	Highlight	Summit Ranks on Graph500 List Using Only a Fraction of its Computing Power
4/29/20	Highlight	New Data Analysis Software Tools Aim at Addressing the Convergence of HPC and AI
4/29/20	PPT Slide	OLCF Staff and Researchers Make a Smooth Transition to Working Remotely
4/29/20	PPT Slide	Summit Ranks on Graph500 List Using Only a Fraction of its Computing Power
4/29/20	PPT Slide	New Data Analysis Software Tools Aim at Addressing the Convergence of HPC and AI
5/5/20	Highlight	Simulations Forecast Nationwide Increase in Human Exposure to Extreme Climate Events
5/5/20	PPT Slide	Simulations Forecast Nationwide Increase in Human Exposure to Extreme Climate Events
5/18/20	Highlight	Mining for COVID-19 Connections
5/18/20	Highlight	A New Method for Unraveling Complex Gene Interactions
5/18/20	Highlight	2020 OLCF User Meeting: Remote Together
5/18/20	Highlight	Supercomputing Aids Scientists Seeking Therapies for Deadly Bacterial Disease
5/18/20	PPT Slide	Mining for COVID-19 Connections
5/18/20	PPT Slide	A New Method for Unraveling Complex Gene Interactions
5/18/20	PPT Slide	2020 OLCF User Meeting: Remote Together
5/18/20	PPT Slide	Supercomputing Aids Scientists Seeking Therapies for Deadly Bacterial Disease
6/4/20	Highlight	Knocking Out Drug Side Effects with Supercomputing
6/4/20	Highlight	OLCF Names New Director of Science
6/4/20	Highlight	Developing a Quantum Community
6/4/20	PPT Slide	Knocking Out Drug Side Effects with Supercomputing
6/4/20	PPT Slide	OLCF Names New Director of Science
6/4/20	PPT Slide	Developing a Quantum Community
6/24/20	Highlight	Amidst Pandemic, OLCF Hosts First Remote User Meeting
6/24/20	Highlight	Cracking the Linux Code with the OLCF's Ketan Maheshwari
6/24/20	Highlight	Summit Helps Predict Molecular Breakups
6/24/20	PPT Slide	Amidst Pandemic, OLCF Hosts First Remote User Meeting
6/24/20	PPT Slide	Cracking the Linux Code with the OLCF's Ketan Maheshwari

Date	Product type	Title
6/24/20	PPT Slide	Summit Helps Predict Molecular Breakups
6/25/20	Highlight	X-Rays Size Up Protein Structure at the "Heart" of COVID-19 Virus
6/25/20	PPT Slide	X-Rays Size Up Protein Structure at the "Heart" of COVID-19 Virus
7/8/20	Highlight	Love-Hate Relationship of Solvent and Water Leads to Better Biomass Breakup
7/8/20	PPT Slide	Love-Hate Relationship of Solvent and Water Leads to Better Biomass Breakup
7/28/20	Highlight	Computational Gene Study Suggests New Pathway for COVID-19 Inflammatory Response
7/28/20	Highlight	Weighing Up Plasma Particles
7/28/20	Highlight	OLCF's Constellation DOI Offers a Permanent Home for OLCF Users' Data Sets
7/28/20	PPT Slide	Computational Gene Study Suggests New Pathway for COVID-19 Inflammatory Response
7/28/20	PPT Slide	Weighing Up Plasma Particles
7/28/20	PPT Slide	OLCF's Constellation DOI Offers a Permanent Home for OLCF Users' Data Sets
8/5/20	Highlight	ALCC Program Awards Nearly 6 Million Summit Node Hours Across 31 Projects
8/5/20	PPT Slide	ALCC Program Awards Nearly 6 Million Summit Node Hours Across 31 Projects
8/10/20	Highlight	Sparks Fly in Marriage of GE's Genesis Code and the Summit Supercomputer
8/10/20	Highlight	Take a Virtual Tour of ORNL's Supercomputer Center
8/10/20	Highlight	Laying the Groundwork for a New "Green Revolution'"
8/10/20	PPT Slide	Sparks Fly in Marriage of GE's Genesis Code and the Summit Supercomputer
8/10/20	PPT Slide	Take a Virtual Tour of ORNL's Supercomputer Center
8/10/20	PPT Slide	Laying the Groundwork for a New "Green Revolution"
8/26/20	Highlight	NCCS Summer Interns Completed Successful Season Despite COVID-19
8/26/20	Highlight	Realizing the Dream of Rotating Detonation Engines through an OLCF, GE, and University of Michigan Collaboration
8/26/20	PPT Slide	NCCS Summer Interns Completed Successful Season Despite COVID-19
8/26/20	PPT Slide	Realizing the Dream of Rotating Detonation Engines through an OLCF, GE, and University of Michigan Collaboration
8/28/20	Highlight	Reaching New Heights in Weather Forecasting's Exascale Future
8/28/20	PPT Slide	Reaching New Heights in Weather Forecasting's Exascale Future
9/3/20	Highlight	Oak Ridge Leadership Computing Facility Fosters GCC Compiler Development with Mentor Contract
9/3/20	PPT Slide	Oak Ridge Leadership Computing Facility Fosters GCC Compiler Development with Mentor Contract
9/23/20	Highlight	DOE and NCSA Supercomputing Speed Studies into Black Hole Origins
9/23/20	Highlight	Powering Frontier
9/23/20	Highlight	Runs on Summit Providing Data for Cray's New Networking Benchmark
9/23/20	Highlight	NASA Team Releases Mars Landing Simulation Data to Encourage New Research into Spacecraft Descent Technologies
9/23/20	Highlight	NCCS' James Simmons Recognized as Top Lustre Contributor of the Decade
9/23/20	Highlight	Summit on Summit, Sierra, and Perlmutter' Hits 2nd Anniversary of Team-Effort Problem-Solving
9/23/20	PPT Slide	DOE and NCSA Supercomputing Speed Studies into Black Hole Origins
9/23/20	PPT Slide	Powering Frontier

Date	Product type	Title
9/23/20	PPT Slide	Runs on Summit Providing Data for Cray's New Networking Benchmark
9/23/20	PPT Slide	NASA Team Releases Mars Landing Simulation Data to Encourage New Research into Spacecraft Descent Technologies
9/23/20	PPT Slide	NCCS' James Simmons Recognized as Top Lustre Contributor of the Decade
9/23/20	PPT Slide	Summit on Summit, Sierra, and Perlmutter' Hits 2nd Anniversary of Team-Effort Problem-Solving
10/12/20	Highlight	Imagine the World of Exascale with Oak Ridge National Laboratory and AMD
10/12/20	Highlight	Frontier: America's Exascale Future
10/12/20	Highlight	The Impact of Exascale Computing on Modeling and Simulation with Big Data and AI
10/12/20	PPT Slide	Imagine the World of Exascale with Oak Ridge National Laboratory and AMD
10/12/20	PPT Slide	Frontier: America's Exascale Future
10/12/20	PPT Slide	The Impact of Exascale Computing on Modeling and Simulation with Big Data and AI
10/28/20	Highlight	In a Groundbreaking Move, Summit Joins Forces with BlazingSQL to Speed Up Data Query Processing on Supercomputers
10/28/20	Highlight	The Greenest Option: Finding New Homes for Old Machines
10/28/20	Highlight	Summit Uncovers Details of Molecular Ferris Wheels Inside Cell Structures
10/28/20	PPT Slide	In a Groundbreaking Move, Summit Joins Forces with BlazingSQL to Speed Up Data Query Processing on Supercomputers
10/28/20	PPT Slide	The Greenest Option: Finding New Homes for Old Machines
10/28/20	PPT Slide	Summit Uncovers Details of Molecular Ferris Wheels Inside Cell Structures
11/10/20	Highlight	OLCF Launches New User Portal, myOLCF
11/10/20	Highlight	Four Teams Using ORNL's Summit Supercomputer Named Finalists in 2020 Gordon Bell Prize
11/10/20	Highlight	OLCF to Trail Long-Awaited Fujitsu Processors in Arm-Based Test Bed System
11/10/20	PPT Slide	OLCF Launches New User Portal, myOLCF
11/10/20	PPT Slide	Four Teams Using ORNL's Summit Supercomputer Named Finalists in 2020 Gordon Bell Prize
11/10/20	PPT Slide	OLCF to Trail Long-Awaited Fujitsu Processors in Arm-Based Test Bed System
11/16/20	Highlight	INCITE Program Awards Supercomputing Time to 51 Computational Research Projects
11/16/20	PPT Slide	INCITE Program Awards Supercomputing Time to 51 Computational Research Projects
11/18/20	Highlight	Two Finalists Nominated for Gordon Bell Special Prize for COVID-19 Work on Summit
11/18/20	Highlight	Gordon Bell Special Prize Finalist Team Reveals AI Workflow for Molecular Systems in the Era of COVID-19
11/18/20	Highlight	Multi-Institutional Team Earns Gordon Bell Special Prize Finalist Nomination for Rapid COVID-19 Molecular Docking Simulations
11/18/20	PPT Slide	Two Finalists Nominated for Gordon Bell Special Prize for COVID-19 Work on Summit
11/18/20	PPT Slide	Gordon Bell Special Prize Finalist Team Reveals AI Workflow for Molecular Systems in the Era of COVID-19
11/18/20	PPT Slide	Multi-Institutional Team Earns Gordon Bell Special Prize Finalist Nomination for Rapid COVID-19 Molecular Docking Simulations

Date	Product type	Title
11/19/20	Highlight	Data Matters: Scientific Data Management Tool Improves Labeling, Information Sharing for OLCF Researchers
11/19/20	Highlight	OLCF Researchers, Staff, Recognized on UT-Battelle Awards Night
11/19/20	Highlight	Distinct ACM Gordon Bell Prizes Awarded to Two Teams Using the Summit Supercomputer
11/19/20	PPT Slide	Data Matters: Scientific Data Management Tool Improves Labeling, Information Sharing for OLCF Researchers
11/19/20	PPT Slide	OLCF Researchers, Staff, Recognized on UT-Battelle Awards Night
11/19/20	PPT Slide	Distinct ACM Gordon Bell Prizes Awarded to Two Teams Using the Summit Supercomputer
11/24/20	Highlight	Gina Tourassi One of Six ORNL Scientists Elected Fellows of American Association for the Advancement of Science
11/24/20	PPT Slide	Gina Tourassi One of Six ORNL Scientists Elected Fellows of American Association for the Advancement of Science
12/8/20	Highlight	VISTA: ORNL’s Visual Informatics for Science and Technology Advances Lab
12/8/20	PPT Slide	VISTA: ORNL’s Visual Informatics for Science and Technology Advances Lab
12/11/20	Highlight	OLCF Reaches New Heights with Andes Cluster
12/11/20	Highlight	Simulations Reveal Nature’s Design for Error Correction During DNA Replication
12/11/20	PPT Slide	OLCF Reaches New Heights with Andes Cluster
12/11/20	PPT Slide	Simulations Reveal Nature’s Design for Error Correction During DNA Replication
12/14/20	Highlight	Building an Exascale-Class Data Center
12/14/20	PPT Slide	Building an Exascale-Class Data Center
12/22/20	Highlight	Jupyter Has Landed at the OLCF
12/22/20	Highlight	Scientists Tap the Power of High-Performance Computing in a Bet to Understand Cancer Growth
12/22/20	PPT Slide	Jupyter Has Landed at the OLCF
12/22/20	PPT Slide	Scientists Tap the Power of High-Performance Computing in a Bet to Understand Cancer Growth

APPENDIX D. BUSINESS RESULTS FORMULAS

2020 Operational Assessment Guidance

Scheduled Availability

Scheduled availability (Eq. [D.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. To be considered a scheduled outage, users must be notified of the maintenance event window no less than 24 h in advance of the outage (emergency fixes). 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.

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

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

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

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. (D.3).

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

Mean Time to Failure

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

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

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. [D.5]).

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

APPENDIX E. DD PROJECTS ENABLED AT ANY POINT IN CY 2020

Table E.1 shows the DD projects enabled at any point in CY 2020.

Table E.1. DD projects enabled at any point in CY 2020.

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
ARD134	Ngoc-Cuong Nguyen	Massachusetts Institute of Technology	20,000	8,935	Massively Parallel Discontinuous Galerkin Methods for Wall-resolved Large Eddy Simulation of Transonic Aeroelasticity
ARD136	Davide Modesti	Delft University of Technology	20,000	10,695	STREAmS: A High-fidelity Compressible Flow Solver for Direct Numerical Simulation on GPUs
ARD137	Eric Nielsen	NASA Langley Research Center	20,000	5,327	FUN3D Scaling for CFD Simulations Requiring General Gas Chemistry
ARD138	Ngoc-Cuong Nguyen	Massachusetts Institute of Technology	20,000	2,914	Direct Numerical Simulation of Transitional Turbulent Flows Past Hypersonic Vehicles
ARD139	Zhi Wang	University of Kansas	15,000	1,464	Large Eddy Simulation of a High-Lift Configuration on GPUs
AST146	Brian O’Shea	MSU	30,000	3,001	Measuring Performance and Scaling of Kokkos-Accelerated Athena++ on Summit
AST149	Brant Robertson	UC Santa Cruz	20,000	8,348	Preparing Cholla for Cosmological Simulations and In Situ Visualization on Summit
AST154	Philipp Moesta	University of Amsterdam	20,000	773	Dynamical Space-Time GRMHD Simulations of Neutron-Star Mergers and Remnants
AST157	Ruonan Wang	ORNL	30,000	6,993	Simulating Full-scale SKA Phase I Dataflow
AST158	Robert Fisher	UMass Dartmouth	20,000	464	Explorations of the D6 Scenario of Type Ia Supernovae on Summit
AST162	Bronson Messer	ORNL	50,000	87,542	Three-Dimensional Stellar Explosions with Detailed Nuclear Kinetics
AST163	Christian Cardall	ORNL	20,000	5,511	INCITE Preparation for 3D+1D core-collapse Supernova Simulations with GenASiS

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
AST164	Eliu Huerta	University of Illinois at Urbana-Champaign	10,000	10,849	Convergence of AI and HPC for Multi-Messenger Astrophysics at extreme scale in Summit
AST165	Francis Halzen	University of Wisconsin	15,000	0	IceCube Simulation at Summit
AST166	Gaurav Khanna	University of Massachusetts Dartmouth	20,000	1,865	Mixed-Precision WENO Method for Hyperbolic PDE Solutions
AST167	Richard Sarmento	United States Naval Academy	20,000	0	Chemical Evolution and Reionization in the Early Universe
ATM112	Wei Zhang	ORNL	40,000	25,828	Aim High: Air Force R&D Collaboration
ATM113	Minsu Joh	Korea Institute of Science and Technology Information	15,000	13	ATMOS—Advancement of Typhoon Prediction Models on Summit
ATM114	Richard Loft	University Corporation for Atmospheric Research	12,000	575	Increasing MPAS-A's GPU Production Readiness at Convection Resolving Scales
ATM115	Stan Posey	NVIDIA Corporation	10,000	373	WRFg
ATM116	Todd Hutchinson	IBM	15,000	0	IBM GRAF
ATM117	Branko Kosovic	NCAR	7,000	62	GPU-Resident Real-Time Large-Eddy Simulations of Atmospheric Boundary Layer
BIE104	Daniel Jacobson	ORNL	100,000	210	CBI Contribution in Kind Allocation
BIE108	Peter St. John	National Renewable Energy Laboratory	20,000	4,769	Enzyme Engineering Directly from Protein Primary Sequence via Natural Language Processing
BIF113	Kjiersten Fagnan	LBNL	0	0	JGI Data Archive
BIF117	Yanling Liu	Frederick National Lab for Cancer Research, Frederick National Lab	20,000	0	Systematic Annotation Creation on H&E WSIs for Automated Digital Pathology Informatics
BIF119	Ali Torkamani	Scripps Research	20,000	923	Genetic Imputation with Deep Learning
BIF120	Burkhard Rost	Technical University of Munich	30,000	252,315	LSTL I: Self-Supervised Deep Learning for Protein Sequences

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
BIF121	Chongle Pan	OU	10,000	682	Acceleration and Parallelization of Bioinformatics Software for Processing Large-Scale Proteogenomics Data
BIF122	Ramu Anandakrishnan	VCOM	15,000	6190	Identifying Multi-Hit Combinations of Genetic Mutations in Cancer
BIF132	Jianlin Cheng	University of Missouri	20,000	4	Improving Deep Learning Prediction of Protein Inter-Residue Distance by High Performance Computing
BIF133	Jeffrey Skolnick	Georgia Tech	20,000	118	Application of a Deep-Learning Based Sequence Alignment Algorithm to Annotate the Proteome of Cable Bacteria
BIP167	Philip Kurian	Howard University	27,000	23,650	Computing Many-body Van Der Waals Dispersion Effects in Biomacromolecules
BIP169	Martin Karplus	Harvard University	20,000	9,033	Computational Design of HIV Vaccination Schedule
BIP179	Arvind Ramanathan	ANL	20,000	25	RL-Fold: Artificial Intelligence Guided Molecular Simulations for Targeting Intrinsically Disordered Proteins
BIP182	Juan Perilla	U. of Delaware, University of Illinois at Urbana-Champaign	10,000	19,358	Revealing the Molecular Mechanisms of the Late Stages of the HIV-1 Infection Cycle Through the Computational Microscope
BIP187	George Karniadakis	Brown University	20,000	574	Supervised Parallel-in-time Algorithm For Long-time Lagrangian Simulations of Stochastic Dynamics
BIP191	Cecilia Clementi	Rice University	25,000	7,613	Extensible and Scalable Adaptive Sampling on Summit
BIP192	Josh Vermaas	National Renewable Energy Laboratory	20,500	34,905	NAMD Benchmarking on Summit for Outer Membrane Permeability
BIP196	Kirk Jordan	IBM	20,000	17,558	Molecular Mechanisms of Resistance to Last Defence Antibiotics
BIP197	Shih-Hsien Liu	ORNL	9,600	5,099	Molecular Modeling of Cellulose Nanofibers in Spray Drying Process
BIP198	Ada Sedova	ORNL	10,000	1,302	GTI: Genes to Interactomes

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
BIP205	Aleksei Aksimentiev	University of Illinois at Urbana-Champaign	20,000	23,169	Benchmark the Nuclear Pore Complex Plus Cytoplasmic Environment on NAMD
BIP207	Darrin York	Rutgers	20,000	1,088	Computational Tools for High-Throughput Lead Optimization
CFD120	Carlos Velez	GE	20,000	22,062	GENESIS GPU Scalability Studies on Summit
CFD124	Bamin Khomami	University of Tennessee, Knoxville	20,000	39,979	Elucidating the Molecular Rheology of Entangled Polymeric Fluids via Direct Comparison of NEMD Simulations and Model Predictions
CFD125	Eric Johnsen	University of Michigan	10,000	0	Cavitation Erosion in the SNS
CFD126	Freddie Witherden	Texas A&M	10,000	0	PyFR and ZEFR Scaling
CFD127	Sanjeeb Bose	Cascade Technologies	15,000	13,202	Porting and Optimization of Large Eddy Simulation Flow Solvers
CFD129	Ramesh Balakrishnan	ANL	10,000	0	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 Tech Lab	20,000	3,917	Combustion Dynamics Predictions in the “Cyber Space,” from “Singing Flames” to a “Roaring Fire”
CFD135	Stephen de Bruyn Kops	University of Massachusetts, University of Massachusetts Amherst	15,000	6,478	Stratified Turbulence at Very Low Froude Number
CFD136	John Gounley	ORNL	15,000	188	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	20,000	13,383	Large Eddy Simulations of SBLI Using High-Order Solver Genesis

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CFD141	Kirk Jordan	IBM	20,000	211	Investigating CFD Models Involving Complex Geometries for ML-Driven Searches at Scale
CFD144	Saumil Patel	ANL	18,000	45	Scaling Internal Combustion Engine Simulations Using nekRS
CHM150	Maria Ramos	University of Porto	4,000	6,728	Local and Collective Motions as a Gateway to the Catalytic Power of Enzymes
CHM151	Pedro Fernandes	University of Porto	4,000	5,044	Enzyme Instantaneous Disorder as a Key Player in the Catalytic Power of Enzymes
CHM152	Stephan Irle	ORNL	20,000	2,561	Benchmarks for the GPU-Accelerated Density-Functional Tight-Binding Code DFTB+ on Summit
CHM154	Coen de Graaf	Catalan Institution for Research and Advanced Studies	20,000	116,614	Energy and Charge Transfer by Nonorthogonal Configuration Interaction
CHM155	Peter Coveney	University College London	25,000	25,029	IMPRESS
CHM156	Remco Havenith	University of Groningen	15,000	7	TURTLE
CHM157	Matthew Ryder	ORNL	10,500	3	Vibrational Dynamics of Organic Materials
CHM158	Lin Lin	UC Berkeley	50,000	65,218	GPU-Accelerated Plane Wave Pseusopotential Density Functional Theory with PWDFT
CHM159	Sotiris Xantheas	PNNL	25,000	231,406	Excited States of DNA/RNA Oligomers
CHM160	Andre Severo Pereira Gomes	French National Centre for Scientific Research	25,500	0	PRECISE: Predictive Electronic Structure Modeling of Heavy Elements
CHM161	Ashley Shields	ORNL	20,000	16,052	Atomistic Bridges to Carbon Defects @ Exascale
CHM166	Chao Yang	LBNL	10,000	3,758	DGDFT
CHM174	Monojoy Goswami	ORNL, University of Tennessee, Knoxville	20,000	0	Multi-Scale/Multi-Physics Molecular Simulations at the Chemical Sciences Division
CHM177	Edward Hohenstein	SLAC National Accelerator Laboratory	20,000	0	Development of Heterogeneous Parallel Electronic Structure Methods
CHM179	Eugen Hruska	Emory University, Rice University	12,000	0	Extensible and Scalable AdaptiveAIMD

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CHP108	Timmy Ramirez-Cuesta	ORNL	20,000	118	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
CLI134	Chris Hill	Massachusetts Institute of Technology	1,000	0	Climate Modeling Alliance (CliMA) - Ocean Modeling
CLI137	Forrest Hoffman	ORNL	20,000	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
CLI141	Jitendra Kumar	ORNL	15,000	0	Scalable Machine Learning for Earth System Science
CLI142	Thomas Robinson	Science Applications International Corporation	1,500	10	GFDL Atmosphere Model 4.0 Porting and Validation
CLI143	Nathaniel Collier	ORNL	15,000	0	Identifying Ecosystems Vulnerable to Climate Change
CLI900	Valentine Anatharaj	ORNL	0	0	Provisioning of Climate Data
CMB124	Ronald Grover	General Motors	0	0	Prediction of Engine Knock in a Gasoline Direct Injection Engine
CMB133	Seung Hyun Kim	Ohio StateUuniversity Research Foundation	0	0	Development of a Physics-Based Combustion Model for Engine Knock Prediction: Phase II
CMB142	Yonduck Sung	Solar Turbines	15,000	10087	Detailed Numerical Modeling of Combustion Oscillation of Low Emission Fuel Injection System For Industrial Gas Turbines
CPH005	Dario Alfe	University College London	20,000	10230	New Frontiers for Material Modeling via Machine Learning Techniques with Quantum
CPH111	Andreas Glatz	ANL	18,000	0	Genetic algorithms for Tailored Superconducting Materials

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CPH121	Anton Kozhevnikov	ETH Zurich	5,000	10	GPU acceleration of Quantum ESPRESSO and CP2K codes
CPH122	Cristian Batista	University of Tennessee, Knoxville	20,000	3071	Transport Properties and Dynamics of Quantum Materials
CPH123	Panchapakesan Ganesh	ORNL	20,000	3	Defects, Interfaces and Disorder in Correlated Quantum Materials
CPH124	Amir Haji-Akbari	Yale University	20,000	0	Computational Investigation of Collective Phenomena in Soft Matter and Biological Systems
CPH125	Jean-Luc Fattebert	ORNL	5,000	1521	Robust Efficient Algorithms for Density Functional Theory (DFT) Calculations
CPH126	Mauro Del Ben	LBNL	10,000	53	BerkeleyGW—Electron Excited-state Properties in Materials
CPH127	Mark Oxley	ORNL	20,000	5814	The Application of Machine Learning to 4D-STEM data
CPH128	Corey Melnick	University of Michigan, BNL	25,000	8170	Application of GPU-Accelerated Quantum Monte-Carlo Impurity Solver to Plutonium Compounds
CPH130	Jihong Ma	University of Vermont	20,000	0	Conductive Polymer
CPH132	Ashley Shields	ORNL	20,000	723	Exascale Lattice Dynamics of Uranium Materials
CSC040	Neena Imam	ORNL	30,000	3106	Durmstrang
CSC143	Norbert Podhorszki	ORNL	70,000	73643	ADIOS—The Adaptable IO System
CSC256	Terry Jones	ORNL	15,000	0	ECP: Simplified Interface for Complex Memories
CSC287	Seung-Hwan Lim	ORNL	0	0	GPU-Accelerated High Dimensional Data Management for ML Workloads in Large Scale GPU-CPU Environments
CSC329	Song Han	Massachusetts Institute of Technology	40,000	4685	Neural Architecture Search for Efficient Visual Understanding
CSC343	Sergey Panitkin	University of Montreal	25,000	10547	Porting the ATLAS Experiment Software and Workload Management System
CSC344	Audris Mockus	University of Tennessee, Knoxville	5,000	0	Behavioral Fingerprinting for Identity Resolution

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC345	Ramakrishnan Kannan	ORNL	20,000	6,589	Parallel Low-Rank Approximation with Nonnegative Constraints (PLANC)
CSC346	David Keyes	King Abdullah University of Science and Technology	20,000	346	Asynchronous High-Performance Task-based Reverse Time Migration
CSC348	Bryan Goodman	Ford Motor Company	10,700	243	ML Performance Evaluation of High Performance Computing System
CSC349	Markus Rampp	Max Planck Computing and Data Facility	10,000	227	Porting, Benchmarking, and Optimization of MPCDF Codes
CSC352	Prasanna Balaprakash	ANL	30,000	0	DeepHyper: Scalable Neural Architecture and Hyperparameter Search for Deep Neural Networks
CSC353	Srikanth Yoginath	ORNL	20,000	0	Accelerating End-to-End Prediction of Physical Phenomena by Interleaving Analytics with Multiscale Simulations
CSC354	Timothy Williams	ANL	40,000	813	Portable Application Development for Next Generation Supercomputer Architectures—Argonne Supplement
CSC355	Ewa Deelman	University of South Carolina	500	490	Pegasus Workflows
CSC356	Edmon Begoli	ORNL	1,000	0	Citadel Pilot
CSC357	Laxmikant Kale	University of Illinois at Urbana-Champaign	10,000	3,816	CharmRTS
CSC358	Rio Yokota	Tokyo Institute of Technology	10,000	0	Large-Batch Data-Parallel Deep Learning
CSC360	Kurt Keutzer	UC Berkeley	15,000	3,888	Large Scale Training of Neural Networks
CSC362	Wen-mei Hwu	University of Illinois at Urbana-Champaign	20,000	9,802	Petascale 3D Image Reconstruction with Ptycho-Tomography
CSC363	Robert Patton	ORNL	20,000	45,011	Machine Learning for Connected Autonomous Vehicles

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC364	Azzam Haidar	NVIDIA Corporation, University of Tennessee, Knoxville	150,000	2	Using NVIDIA Tensor Cores to Drive the Summit Toward Exascale
CSC365	Seung-Hwan Lim	ORNL	20,000	7,253	GPU-Accelerated High Dimensional Data Management for ML Workloads in Large Scale GPU-CPU Environments
CSC366	Rogerio Feris	IBM	20,000	19,736	Large-Scale Neural Architecture Search for Visual Recognition
CSC367	Yasuhiro Idomura	Japan Atomic Energy Research Institute	20,000	12,833	Exascale CFD Simulations at JAEA
CSC368	Robert Patton	ORNL	30,000	15,466	Scalable Machine Learning of Scientific Data
CSC369	Dmitry Pekurovsky	UC San Diego	5,000	12,082	Scalable Software Framework for Multidimensional Fourier Transforms
CSC370	Mark Shephard	Rensselaer Polytechnic Institute	16,500	15,441	Measuring the Scalability of Nvidia Multi-Process Service to Improve GPU Throughput in Multiscale Biological Simulations
CSC377	Allan Grosvenor	Microsurgeonbot	20,000	10	Expertise-as-a-Service via Scalable Hybrid Learning: R&D Supporting Improvements to Hybrid Intelligence Agent Tooling
CSC378	Sudip Seal	ORNL	20,000	10,837	Scaling Studies of Machine Learning Workloads on Summit
CSC380	Bronson Messer	ORNL	150,000	123	CAAR for Frontier
CSC381	Ramakrishnan Kannan	ORNL	20,000	0	Graph500 Algorithms for Summit
CSC382	Catherine Schuman	ORNL	15,000	11,115	Scalable Neuromorphic Simulation and Training
CSC384	Arghya Das	University of Wisconsin	20,000	0	Accelerating Large-Scale Metagenomic Analysis with POWER System
CSC387	Una-May O’Reilly	Massachusetts Institute of Technology	25,000	302	Towards A Robust and Scalable Adversarial Learning Framework
CSC388	Thomas Uram	ANL	1,000	2	Balsam—An HPC Job Campaign Management Framework

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC392	Thomas Deakin	University of Bristol	15,000	0	Scalability of Unstructured Mesh SN Particle Transport on GPUs
CSC393	Shantenu Jha	Rutgers	20,000	17,600	RADICAL-Cybertools: Middleware Building Blocks for Workflows
CSC394	Robert Robey	LANL	0	0	PCSRI—LANL Parallel Computing Summer Research Internships
CSC395	Dali Wang	ORNL	15,000	543	An Extreme Scale Deep Reinforcement Learning Framework for Solving Big Graph Problems
CSC397	Michael Bauer	NVIDIA Corporation	15,000	8	Legate
CSC398	Thomas Naughton	ORNL	5,000	0	Federated Scientific Instruments
CSC399	Douglas Doerfler	LBNL	20,000	12,509	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	State University of New York at Stony Brook	10,000	905	Support for the Deployment of OpenMP in Large-Scale Scientific Applications
CSC416	David Keyes	King Abdullah University of Science and Technology	20,000	65,035	Geostatistical Modeling and Prediction In Three Precisions
CSC418	Tjerk Straatsma	ORNL	60,000	11	ADAC Applications Readiness
CSC422	Luke Olson	University of Illinois at Urbana-Champaign	10,000	82	Node-Aware Communication
CSC423	Edmon Begoli	ORNL	50,000	0	Citadel COVID-19
CSC424	Theodore Papamarkou	ORNL	11,200	0	Uncertainty Quantification By Combining Data Augmentation And Ensemble Learning
CSC425	Terry Jones	ORNL	20,000	0	WEAVABLE
CSC426	Athirai Aravazhi Irissappane	University of Washington	20,000	0	Multi-Modal Brain Tumor Segmentation
CSC427	Michela Taufer	University of Tennessee, Knoxville	15,000	1,909	A4NN: Accelerating Scientific AI Leveraging Open Data and Open Models
CSC436	John Paul Walters	University of South Carolina	8,000	2	CASPER: Compiler Abstractions Supporting High Performance on Extreme-scale Resources

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC443	Seung-Hwan Lim	ORNL	20,000	4,600	Generalized Scalable Parallel Training of Deep Graph Neural Networks
CSC445	Rama Vasudevan	ORNL	8,000	1,005	Variational Scalable Multi-agent Reinforcement Learning
CSC452	Abhinav Bhatele	University of Maryland, LLNL	13,000	0	Performance Analysis and Tuning of HPC and AI Applications
CSC453	Abhinav Bhatele	University of Maryland, LLNL	8,000	0	Analyzing and Optimizing Parallel I/O and Performance Tools
CSC454	John Michalakes	National Renewable Energy Laboratory, University Corporation for Atmospheric Research	100	12	US Navy NEPTUNE Weather Model Development for GPU
CSC456	Benjamin Nebgen	LANL	15,000	775	Distributed Non-Negative Tensor Factorization for Global Seismic Analysis
ENG104	David Gutzwiller	Numeca, Numeca International	25,000	5	Porting and Testing of the NUMECA FINE/Open CFD Solver on OLCF Summit
ENG106	Ramanan Sankaran	ORNL	20,000	10,064	Simulation of Transport Phenomena in Molten Media Reactors
ENG107	Andrew McGough	Newcastle University	20,000	0	NUFEB: Newcastle University Frontiers in Engineering Biology
ENG108	Ryan Viertel	Sandia	15,000	1,095	Optimization Based Design for Manufacturing
ENG114	Charles Moulinec	Science and Technology Facilities Council, UK Research and Innovation-Science and Technology Facilities Council	10,000	11,825	Testing the Multi-Scale CPU-GPU Finite Element Squared Method at Scale
ENG120	George Karniadakis	Brown	10,000	143	PINNs on Multi-GPUs
FUS125	Mark Kostuk	General Atomics	9,000	3,264	ALMA—Plasma Fluid Simulations
FUS128	David Smith	University of Wisconsin	2,000	0	Deep Learning Analysis of GPI Data to Advance Tokamak Edge Physics
FUS131	Randy Churchill	Princeton Plasma Physics Laboratory	20,000	14,754	Advanced Sequence Models and Self-Supervised Learning for Fusion Energy Diagnostics

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
FUS134	Xianzhu Tang	LANL	12,000	0	Tokamak Disruption Simulations
FUS135	Chang Liu	Princeton Plasma Physics Laboratory	20,000	17,542	Kinetic-MHD Hybrid Simulation of Alfvén Instabilities Interacting with Energetic Particles
GEN107	Ashley Barker	ORNL	5,000	8	NCCS UA User Testing Emulation Project
GEN150	Feiyi Wang	ORNL	200,000	0	Non-Staff—AI Analytics Methods at Scales Group
GEO132	Ethan Coon	ORNL, LANL	20,000	32	ExaSheds
GEO133	Yidong Xia	Idaho National Laboratory	20,000	0	Multiphase Flow Simulations in Mesoporous Source Rocks
GEO134	Ralph Dunlap	University Corporation for Atmospheric Research	0	0	Earth System Modeling Framework
GEO135	Kohei Fujita	University of Tokyo	20,000	40	Fast and Scalable Implicit Unstructured Finite-Element Earthquake Simulations
GEO136	James McClure	Virginia Polytechnic Institute and State University	20,000	143,668	Modeling Wetting Phenomena in Porous Media
GEO137	Dalton Lunga	ORNL	20,000	6,990	Accelerated Infrastructure Mapping From Trillion Pixels
HEP117	Alan Stagg	Sandia	25,000	0	Z-Machine Target Physics Simulation
HEP118	Brian Nord	Fermilab	15,000	11,291	Machine Learning for Astrophysics and Cosmology
HEP120	Paolo Calafiura	LBNL	15,000	0	Exa.TrkX, HEP Tracking at Exascale
HEP121	Corey Adams	ANL	20,000	139	Sparse Convolutional Neural Networks for Neutrinos
LGT107	Rajan Gupta	LANL	25,000	57,712	Nucleon Matrix Elements: Probes of New Physics
LRN006	Katherine Riley	ANL	16,000	0	Advanced Materials Characterization with AI-Informed Computation
LRN007	Katherine Riley	ANL	16,000	19,061	Dynamic Compressed Sensing for Real-Time Tomographic Reconstruction
LRN008	Katherine Riley	ANL	3,000	811	Developing High Performance Computing Applications for Liquid Argon Neutrino Detectors (Szelc)
LSC110	Amanda Randles	Duke University	40,000	39,245	Coupled FSI/CFD Simulations

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
LSC111	Daniel Jacobson	University of Tennessee, Knoxville	10,000	0	Wound Healing Machine Learning
MAT193	Trung Nguyen	Northwestern University	15,000	415	Development of New Models and Sampling Techniques for LAMMPS
MAT197	Yangyang Wang	ORNL	20,000	10,090	Exploring New Paradigms for Understanding Ionic Transport in Polymer Electrolytes
MAT198	Ivan Oleynik	University of South Florida	75,000	75,725	Predictive Simulations of Phase Transitions in Dynamically-Compressed Materials
MAT199	Lou Kondic	New Jersey Institute of Technology	5,000	102	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	150	Center for Nanophase Materials Sciences
MAT202	Vikram Gavini	University of Michigan	22,750	891	Large-Scale Discrete Fourier Transform Simulations with DFT-FE
MAT203	Zhiting Tian	Cornell University	15,000	9,409	Thermal Transport Properties of Covalent Organic Frameworks (COFs)
MAT204	Ram Devanathan	PNNL	20,000	0	Simulated Adhesive Response to Aging
MAT206	Robert Szilagyi	Montana State University	20,000	24,533	Quantum Chemical Refinement of Periodic Structural Models for Zeolite-Templated Carbon Materials
MAT208	Volker Blum	Duke University	20,000	18,850	Distributed-Parallel Dense Eigenvalue Solutions for Electronic Structure Calculations in ELSI and ELPA
MAT213	Vikram Gavini	University of Michigan	100,000	85,229	Large-Scale Ab-Initio Molecular Dynamics Simulations of InP Based Photoelectrochemical Cells
MAT215	Lin	UC Berkeley	60,000	58,798	NNMD
MAT220	Feliciano Giustino	University of Texas at Austin	50,000	49,002	EPW Scaling Tests for INCITE Proposal
MAT223	Neil Gershenfeld	Massachusetts Institute of Technology	10,000	357	Performance Scaling of Particle Systems for Discovery of Multiphysics Models

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
MAT224	Dongwon Shin	ORNL	50,000	10,704	Computational Design of High-Temperature Alloys
MAT226	Jan Michael Carrillo	ORNL, University of Tennessee, Knoxville	20,000	707	Molecular Dynamic Simulations of Amphiphilic Oligomer Membranes: Design Rules towards Stable Membranes Capable of Learning and Memory
MED107	John Gounley	ORNL	10,000	4,932	AI-Enabled Computational Cancer Phenotyping for Precision Oncology
MED108	Joel Saltz	State University of New York at Stony Brook	13,000	7,199	Comprehensive Morphology Analysis of Whole Slide Tissue Specimens
MED112	Sumitra Muralidhar	US Department of Veterans Affairs	200,000	12	Genome-Wide Phenome-Wide Association Study in the Million Veteran Program
MPH113	Balakrishnan Naduvalath	University of Nevada, Las Vegas	20,000	11,137	Ultracold Controlled Chemistry
NFI115	Cole Gentry	ORNL	10,000	0	Machine-Learning-Based Multi-Physics Nuclear Reactor Core Simulations of Molten Salt Reactor
NFI120	Justin Watson	University of Florida	20,000	576	Neural Networks for Exa-Scale Reactor Parameter Prediction
NPH133	Eric Church	PNNL	15,000	5,882	NEXT—Neutrino Experiment in a Xenon Time Projection Chamber
NPH136	Rick Archibald	ORNL	20,000	127	Development of Design Optimization Code for the Transformational Challenge Reactor
NRO104	Josh McDermott	Massachusetts Institute of Technology	25,000	2,236	Next-Generation Hearing Aids Via Neural Network Models of the Auditory System
NRO105	Satrajit Ghosh	Massachusetts Institute of Technology	10,000	0	Creating Baseline DNN Models for Biomedical Signals
NRO106	Kristofer Bouchard	LBNL	15,000	9,770	DL4Neurons
PHY149	Dmitry Liakh	ORNL	32,000	1,417	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	386	2020 SMC Data Challenge

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
PHY155	Edwin Pednault	IBM	24,500	0	Leveraging Secondary Storage in Quantum Circuit Simulation
STF018	Feiyi Wang	ORNL	20,0000	0	Staff—AI Analytics Methods at Scales Group
TUR120	Pui-kuen Yeung	Georgia Tech	40,000	7,734	High Resolution Study of Intermittency in Turbulence and Turbulent Mixing
TUR130	Arvind Mohan	LANL	7,000	0	Machine Learning for Turbulence
TUR132	Mujeeb Malik	NASA Langley Research Center	20,000	19,723	Direct Numerical Simulations of Complex Turbulent Flows at High Reynolds Number
TUR135	Sanjiva Lele	Stanford	20,000	1,588	Compressibility and Curvature Effects on Turbulent Mixing
TUR137	Justin Sirignano	Oxford	20,000	4	Deep Learning Closure Models for Large-Eddy Simulation of Unsteady External Aerodynamics