OAK RIDGE
NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR THE DEPARTMENT OF ENERGY

ORNL/TM-2011/314

U.S. Department of Energy, Office of Science

High Performance Computing Facility Operational Assessment, FY11 Oak Ridge Leadership Computing Facility

August 2011

Prepared by

Arthur S. Bland
James J. Hack
Ann E. Baker
Ashley D. Barker
Kathlyn J. Boudwin
Ricky A. Kendall
Bronson Messer
James H. Rogers
Galen M. Shipman
Jack C. Wells
Julia C. White

UT-BATTELLE
ORNL-27 (4-00)

U.S. Department of Energy, Office of Science

HIGH PERFORMANCE COMPUTING FACILITY OPERATIONAL ASSESSMENT, FY11 OAK RIDGE LEADERSHIP COMPUTING FACILITY

Arthur S. Bland
James J. Hack
Ann E. Baker
Ashley D. Barker
Kathlyn J. Boudwin
Ricky A. Kendall
Bronson Messer
James H. Rogers
Galen M. Shipman
Jack C. Wells
Julia C. White

August 2011

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

CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

ACRONYMS

3-D	three-dimensional
ACTS	Academies Creating Teacher Scientists
ADIOS	ADaptable Input/Output System
ALCC	ASCR Leadership Computing Challenge (DOE)
ALCF	Argonne Leadership Computing Facility
ANI	Advanced Networking Initiative
ANL	Argonne National Laboratory
API	application programming interface
ARC	Appalachian Regional Commission
ARRA	American Recovery and Reinvestment Act (of 2009)
ASCAC	Advanced Scientific Computing Advisory Committee (DOE SC)
ASCR	Advanced Scientific Computing Research (DOE program office)
BA	budget authority
C&A	certification and accreditation
CAAR	Center for Accelerated Application Readiness
CAM	Community Atmosphere Model
CCES	Climate-Science Computational End Station (INCITE project)
CCSM	Community Climate System Model
CEA	Commissariat à l’énergie atomique et aux énergies alternatives
CFD	computational fluid dynamics
CFP	call for proposals
CCI	Common Communication Interface
CR	continuing resolution
CSB	Computational Sciences Building (ORNL)
CSSEF	Climate Science for Sustainable Energy Future
CY	calendar year
DD	Director’s Discretionary
DDT	Distributed Debugging Tool (Allinea Software Ltd.)
DDN	DataDirect Networks (data storage infrastructure company)
DME	development, modernization, and enhancement
DOE	Department of Energy
eSimMon	electronic Simulation Monitoring
ESnet	Energy Sciences Network
EOFS	European Open File System consortium
FAQ	frequently asked question
FTE	full-time equivalent
FY	fiscal year
GB	gigabyte
GB/s	GB per second
GPGPU	general purpose GPU

GPU	graphics processing unit
GROMACS	GROningen MAchine for Chemical Simulations
HMPP	hybrid multicore parallel programming (compiler)
HPC	high-performance computing
HPSS	High-Performance Storage System
I/O	input/output
ICMS	Institute for Computational and Molecular Science
INCITE	Innovative and Novel Computational Impact on Theory and Experiment
ISV	independent software vendor
IT	information technology
LAMMPS	Large-Scale Atomic/Molecular Massively Parallel Simulator
LBNL	Lawrence Berkeley National Laboratory
LCF	Leadership Computing Facility
LLNL	Lawrence Livermore National Laboratory
LEED	Leadership in Energy and Environmental Design
LSMS	locally self-consistent multiple scattering
LUG	Lustre User Group
MFiX	Multiphase Flow with Interphase eXchanges
MPI	message passing interface
MTTF	mean time to failure
MTTI	meant time to interrupt
NAMD	Not just Another Molecular Dynamics program
NCAR	National Center for Atmospheric Research
NCCS	National Center for Computational Sciences
NEMO	Nanoelectric Modeling (program)
NERSC	National Energy Research Scientific Computing Center
NETL	National Energy Technology Laboratory
NOAA	National Oceanic and Atmospheric Administration
OA	overall availability
OMB	Office of Management and Budget
OLCF	Oak Ridge Leadership Computing Facility
OpenSFS	Open Scalable File Systems, Inc.
ORISE	Oak Ridge Institute for Science and Education
ORNL	Oak Ridge National Laboratory
PAS	Personnel Access System
PB	Petabyte
PI	principal investigator
PNNL	Pacific Northwest National Laboratory
RMP	risk management plan
RMTAP	Risk Management Techniques and Practice
RT	Request Tracker (ticket tracking software)
RUC	Resource Utilization Council (OLCF)
SA	scheduled availability

SC	Office of Science (DOE)
SC10	Supercomputing 2010
SciComp	Scientific Computing Group (OLCF)
SciDAC	Scientific Discovery through Advanced Computing
SDN	Science Data Network
SMP	symmetric multiprocessing
SNL	Sandia National Laboratories
SSD	solid-state disk
SSM	storage system management (part of HPSS software)
SWC	Software Council (OLCF)
TB	Terabyte
TechInt	Technology Integration Group (OLCF)
UAO	User Assistance and Outreach Group
UME	uncorrectable memory error
UTRC	United Technologies Research Center
Vampir	Visualization and Analysis of MPI Resources (TUD)
VRM	voltage regulator module
WBS	work breakdown structure
YTD	year to date

EXECUTIVE SUMMARY

Oak Ridge National Laboratory's Leadership Computing Facility (OLCF) continues to deliver the most powerful resources in the U.S. for open science. At 2.33 petaflops peak performance, the Cray XT Jaguar delivered more than 1.5 billion core hours in calendar year (CY) 2010 to researchers around the world for computational simulations relevant to national and energy security; advancing the frontiers of knowledge in physical sciences and areas of biological, medical, environmental, and computer sciences; and providing world-class research facilities for the nation's science enterprise.

Scientific achievements by OLCF users range from collaboration with university experimentalists to produce a working supercapacitor that uses atom-thick sheets of carbon materials to finely determining the resolution requirements for simulations of coal gasifiers and their components, thus laying the foundation for development of commercial-scale gasifiers. OLCF users are pushing the boundaries with software applications sustaining more than one petaflop of performance in the quest to illuminate the fundamental nature of electronic devices. Other teams of researchers are working to resolve predictive capabilities of climate models, to refine and validate genome sequencing, and to explore the most fundamental materials in nature – quarks and gluons – and their unique properties. Details of these scientific endeavors – not possible without access to leadership-class computing resources – are detailed in Section 4 of this report and in the *INCITE in Review*, available at http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/INCITE_IR.pdf.

Effective operations of the OLCF play a key role in the scientific missions and accomplishments of its users. This Operational Assessment Report (OAR) will delineate the policies, procedures, and innovations implemented by the OLCF to continue delivering a petaflop-scale resource for cutting-edge research.

2010–2011 highlights of OLCF operational activities include the following.

- Leadership of SciApps meeting in August 2010, bringing together more than 70 computational scientists to share experience, best practices, and knowledge about how to sustain large-scale applications on leading HPC systems while looking toward building a foundation for exascale research.
- Active engagement of the OLCF User Council in Center outreach (User Science Exhibition on Capitol Hill), policy changes, and solicitation of user survey responses (Reference Section 2.1).
- Delivery of operational solutions: Working with Cray, an engineering change related to the input voltage to the voltage regulator modules (VRMs) was identified and implemented (Reference Section 3)

The 2010 operational assessment of the OLCF yielded recommendations that have been addressed (Reference Section 1) and where appropriate, changes in Center metrics were introduced. This report covers CY 2010 and CY 2011 Year to Date (YTD) that unless otherwise specified, denotes January 1, 2011 through June 30, 2011.

User Support remains an important element of the OLCF operations, with the philosophy "whatever it takes" to enable successful research. Impact of this center-wide activity is reflected by the user survey results that show users are "very satisfied." The OLCF continues to aggressively pursue outreach and training activities to promote awareness—and effective use—of U.S. leadership-class resources (Reference Section 2).

The OLCF continues to meet and in many cases exceed DOE metrics for capability usage (35% target in CY 2010, delivered 39%; 40% target in CY 2011, 54% January 1, 2011 through June 30, 2011). The Schedule Availability (SA) and Overall Availability (OA) for Jaguar were exceeded in CY2010. Given the solution to the VRM problem the SA and OA for Jaguar in CY 2011 are expected to exceed the target metrics of 95% and 90%, respectively (Reference Section 3).

Numerous and wide-ranging research accomplishments, scientific support, and technological innovations are more fully described in Sections 4 and 6 and reflect OLCF leadership in enabling high-impact science solutions and vision in creating an exascale-ready center.

Financial Management (Section 5) and Risk Management (Section 7) are carried out using best practices approved of by DOE. The OLCF has a valid cyber security plan and Authority to Operate (Section 8). The proposed metrics for 2012 are reflected in Section 9.

1. RESPONSES TO RECOMMENDATIONS FROM THE 2010 OPERATIONAL ASSESSMENT REVIEW

CHARGE QUESTION (1) ***Are the Facility responses to the recommendations from the previous year's OAR reasonable?***

OLCF RESPONSE The OLCF responses to the recommendations from the previous year's OAR are provided below, with both the intial response from August 2010, and with an updated response where appropriate.

1. Are the processes for supporting the customers, resolving problems, and communicating with key stakeholders effective?

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
Consider evaluating changes in user survey ratings between years to determine whether the changes are statistically significant.	The OLCF already performs this function but would be happy to include comments about the statistical significant of variations in user survey results in the next Operational Assessment (OA) report.	No significant variations were found from 2009 to 2010, the most recent user survey.
OLCF is to be commended for the improvement of its survey scores over the past four years; however it should investigate possible ways to improve the survey response rate.	Thank you for the recommendation. To address this, the center director will send a kick-off email asking users to participate in the survey. This past year, all notifications to the users were handled by the 3rd party contractor who administered the survey. We believe a personal message from the center director will increase the response rate. In the same email, we plan to enumerate a few of the changes made as a result of the 2010 survey feedback. Our belief is that if users understand that their input is used to make effective change, more will participate. Lastly, we plan to engage the OLCF user council in reaching out to users for their participation.	For 2011, the following direct outreach was used to increase participation in the user survey: • The OLCF Project Director, Arthur Bland, sent a notice to all users emphasizing the importance of the survey. • The OLCF User Council Chair, Balint Joo, also sent a notice to all users on behalf of the council. • The UAO Group Leader, Ashley Barker, sent out reminders. • The Center liaisons reached out to the principal investigators to encourage their participation. Each of these efforts demonstrated a measurable and immediate increase in the number of returned surveys.

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
There was a large drop in the percentage of new user respondents for the 2009 user survey as well as the number of respondents who used the user assistance center at least one time. OLCF should investigate and report on the reason for these changes.	We suspect there was a drop in both numbers due to the number of INCITE projects that were renewals rather than new projects. Therefore, we had fewer new users and more returning users than in previous years. Returning users tend to need less user assistance as they form relationships with their scientific liaison and they already have experience using the center resources.	No additional update.
OLCF should consider publishing the survey results and its responses on the OLCF website. This helps users understand that their input has been received and that the center has taken steps to explain or improve the environment.	Thank you, this is a good idea for the reasons stated. The OLCF will publish the results of the 2010 user survey on the center website.	The OLCF has created a web content section accessible from the OLCF home page where users can review the results of all surveys, beginning with the 2010 report. The 2010 report is currently posted and is available at http://www.olcf.ornl.gov/media-center/center-reports/2010-outreach-survey/ .
OLCF should provide separate user survey scores for the INCITE/ALCC projects. This will allow it to assess whether its strategic customers are satisfied. Typically, there are a lot more Discretionary users than INCITE/ALCC users, and the Discretionary users responses could overwhelm the INCITE/ALCC responses.	We don't agree that a separate user survey is required. Responses can be categorized by asking the user to identify their project type(s): INCITE, Discretionary, or ALCC and assessing any variations. Discretionary awards, in particular, are one vehicle for users to gain experience on the OLCF resource in preparation for an INCITE proposal. By participating in the user survey process, they become accustomed to the policies and requirements applied to all users.	The Center asked respondents to the 2010 user survey to self-identify their project's program type(s). Reference Table 2.2 in Section 2.1
Consideration should also be given to surveying projects rather than individuals to prevent many vocal users on a single project from skewing the results.	Conversely, only surveying the PI of the project provides limited value since the PI typically has only minimal time on the machine or interactions with staff. We find it more beneficial to have more information, which we can sift through to identify needs and areas where we can and should make improvements, than less information that leaves us guessing as to user problems or concerns.	N/A
OLCF should consider reporting problem ticket statistics based on type of ticket (account, compiler, hardware, etc.).	We collect this information and will include it in next year's report.	Reference Section 2.2 for the results.

2. Is the OLCF maximizing resources consistent with its mission?

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
Try to improve MTTI for Jaguar; it would be good to get this from the current 2 days into the 4 to 6 day range. Hopefully as Jaguar matures it will require fewer scheduled maintenances.	With Spider going into full production, we have decreased the frequency of Lustre testing, which will favorably impact Jaguar MTTI. We concur that as Jaguar matures, scheduled maintenance will be less frequent.	OLCF systems administrators implemented a software patch to CLE 2.2UP03 that significantly reduced the impact of portals errors and their contribution to SeaStar interconnect failure (HT_Lockup). OLCF and Cray implemented an engineering change that significantly reduces the instances of voltage regulator module (VRM) failure. Early (60-day) analysis has been very positive. Reference Section 3 for details.
The OLCF Resource Utilization Council (RUC) initiated a study of queuing on the OLCF. Based on the results, RUC suggested a new policy which has been implemented. • The OLCF should report on the impact of the new policy in the next OA. • The OLCF should consider adding questions to the 2010 user survey to gather user feedback on the policy change	The change to the scheduling policy was implemented in response to the machine's increased expansion factor in late 2009. In order to continue to give leadership class jobs priority, the OLCF adjusted the queue policy to reflect the change in definition of a leadership class job. The impact of the scheduling policy can be measured by the OLCF's success in meeting the leadership metric after such a dramatic increase in compute resources as the site has experienced over the past 18 months. The OLCF also surveys users every year regarding queue policies and will continue to track user satisfaction in this area and use the feedback as basis for further adjustments as needed. The leadership metrics and user survey responses reported this year will continue to be given in future OA reports.	The impact of queuing policy is reflected in the capability metric. Reference Figure 3.5 and Table 3.6.

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
OLCF should provide details about how it calculates its scheduled and overall availability for the different resources. For example, when does it consider the full system down? If the file system or network is down, is the full system considered down? If a majority of nodes are down is full system down? If the scheduler is down, but existing jobs continue to run, is the system down? If a tape drive or redundant file serve are down, is there any fractional lost of availability? If hardware failures cause performance impacts that make it difficult for users to recover data from tape or access the file system at reasonable speeds are the resources considered down?	The three sites are currently working together to define a common set of formulas and definitions for these metrics.	OLCF participated in the discussions about SA and OA with NERSC, (F. Verdier) and ANL (S. Coughlan), led by Betsy Riley. The results of that discussion were provided to the Program Office for their consideration. OLCF has an extensive monitoring system that collects sensor data (availability/health) from multiple system components, and reports an aggregated high-level status to users through a web-based dashboard. This monitoring system takes into account the loss of a system component, and whether the loss of that component should contribute to the reporting of a degraded or down state. System administrators assess the impact of a system or component failure on the availability of the larger resource. In general, a degraded/down state of a redundant component does not constitute "down."
DOE metric calculations should be standardized across all facilities. Targets for the metrics can, and should, differ between the facilities based on their missions, but the definitions, and calculations, of MTTI, MTTF, Scheduled and Overall Availability should be the same.	This recommendation has already been addressed by HQ, in its initial gathering of data from each site. We are happy to participate in discussions about metrics and their standardization.	OLCF management joined NERSC and the ALCF in discussions about metric definitions. The results of that discussion were provided to the Program Office for their consideration.
OLCF should report actual utilization numbers instead of the percentage of INCITE allocations used where utilization means: $\frac{\textit{Core hours consumed by jobs}}{\textit{Core hours overall available}}$. A graph similar to the capability graph, with better resolution (such as weekly average), should be provided.	The OLCF will provide this information in future OA reports.	Reference Section 3.2

3. Is the OLCF meeting the Department of Energy strategic goals 3.1 and 3.2?

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
OLCF should provide some measurement of the presentations given by OLCF INCITE/ALCC projects, especially high-profile conference presentations.	OLCF currently collects information on presentations given by project participants as a part of the quarterly report process. We are happy to provide this data in future reports and would be interested in engaging the other sites and HQ in a discussion of the types of information that can best characterize the progress of research projects.	Reference Section 4.1.

4. How well is the program executing to the cost baseline pre-established during the previous year's Budget Deep Dive? Explain major discrepancies.

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
DOE Program Management and OLCF management should review FY11 and FY12 plans once a more reliable estimate is known.	We agree with the recommendation. The current plan is based on best knowledge to date, but funding changes and facility status could alter plans.	The OLCF reviewed FY11 and FY12 plans with DOE Program management several times in FY11 including a budget deep dive in July 2011.
In addition to the chart (Figure 4.1), a table, such as provided in the guidance, showing the FY10 pre-established data, the actual data to date, and the proposed FY11 budget should be provided to facilitate comparison of the data across years.	This data is presented in graphic form but in future OA reports a table will be added as suggested.	Reference Section 5.
Variances details, as well as details on significant changes from one year to the next (e.g., center balance activity jump in FY11) should be provided.	The variance details were provided for the largest variances, but in future OA reports more detail will be provided as requested; details for this year can be provided if requested.	Reference Section 5.
Details about what is in each budget line item should be provided.	We concur and this will be included in future OA reports; details for this year can be provided if requested.	Details about each budget line item are shown in Table 5.1.

5. What innovations have been implemented that have improved OLCF's operations?

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
The OLCF should provide details on the OLCF contribution to innovations that involved other institutions and/or companies, specifically on the topic of the division of responsibilities and work performed.	The center is happy to provide this information in future reports. With regards to the 2010 OA Report, staff involvement is summarized below. <u>Center-Wide File System</u>: The OLCF's Spider parallel file system was a collaborative effort between OLCF staff, Cray, DDN, and Oracle (Formerly SUN, Formerly Cluster File Systems). OLCF staff members led virtually all aspects of prototyping and early deployment of systems prior to the production deployment of the Spider file system. This included adding support for the InfiniBand software stack on the Cray XT SIO node, followed by early prototyping of the Lustre LNET router on the Cray XT SIO node. Evaluation of hardware components from the DDN and LSI storage arrays to InfiniBand optical cabling was performed by OLCF staff. Scalability testing and tuning was conducted by OLCF staff in collaboration with Cray and Oracle. Lustre engineers at Oracle were contracted to develop the Lustre networking router component, a critical technology allowing of high-performance network transfers between heterogeneous networks. Oracle and Cray provided expertise in improving the scalability of the Lustre file system while Oracle and DDN provided expertise in improving the performance of Lustre on the DDN storage systems. In many cases Oracle and Cray leveraged prototypes developed by OLCF staff in adding support for features required for the successful deployment of the Spider parallel file system. <u>Tool development at the OLCF</u>: MDSTrace, DDNTool, Monitoring GUIs, System log analysis, and parallel data tools are developed exclusively by OLCF staff.	Details are provided in Section 6, and include OLCF collaborations with, for example, OpenSFS, CCI, HPSS, Allinea Software, and Vampir. In addition to assisting in the evaluation of the scalable debugger and administration of Allinea contract deliverables to the OLCF, OLCF staff define the new technical features and performance requirements.

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
	ADIOS and eSimMon are collaborative research and development projects with lead development conducted by OLCF staff. ADIOS is a collaborative effort between the College of Computing at Georgia Tech and the OLCF. eSimMon is a collaborative development effort between the OLCF, University of Utah, and the University of North Carolina. Primary development Is led by OLCF staff members for both areas of the project. The OLCF's centralized software maintenance system known as SWTools is a product of the OLCF in collaboration with the National Institute for Computational Sciences. OLCF staff members conduct primary development and management of this system. <u>HPSS development</u> is conducted by the HPSS collaboration that includes IBM, LANL, LBNL, LLNL, ORNL, and SNL. OLCF staff are the primary developer on a number of HPSS components including the bitfile server, the logging and accounting systems and the administrator interface (the storage system manager). <u>Improved Operating System Scalability:</u> Efforts to improve the scalability of the Cray XT Linux platform were led primarily by Cray with testing and design critique conducted by OLCF staff. Results of this work were published at the Cray Users Group—2010. <u>Scalable Debugging and Performance Tools</u>: The development and demonstration of a scalable debugger is led by Allinea with most aspects of the work conducted by Allinea on a contractual basis. OLCF staff assists in the evaluation of the scalable debugger and administration of contract deliverables to the OLCF. <u>Industry Partnerships</u>: The Industrial Partnership program is led by the OLCF with participation from a number of industrial partners. The OLCF provides expertise in application scalability and the use of the OLCF resources.	

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
OLCF should clarify efforts that are leveraged from other funded sources such as ADIOS and eSimMon funding from the DOE SciDAC program	All projects described within the OLCF innovations section are funded exclusively by the OLCF project except the eSimMon and ADIOS projects. eSimMon leverages funding from OFES and OASCR through the CPES Fusion SciDAC project and SDM OASCR funding in addition to OLCF funding. ADIOS leverages funding from NSF (HECURA), FES, and the SDM OASCR project.	The Earth Systems Grid is funded by BER SciDAC and the innovation described in Section 6 is its deployment through the OLCF. eSimMon also leverages other funding sources, as previously described.

6. Is the OLCF effectively managing risk?

Recommendation	August 2010 ORNL Action/Comments	Updated (June 30, 2011)
OLCF should follow the DOE guidance document when developing its OA report, in particular, it should report current top level operating and technical risks and CY 11 projected risks.	The OLCF will provide this information in future OA reports.	Reference Section 7.
OLCF should explain the rationale for the ranges of risk likelihood used for risk assessment. The <30%, 30%, –80%, >80% ranges appear skewed towards Low and Medium scores and differ from those used by both NERSC and Argonne.	The combination of OLCF's likelihood and impact thresholds produces risk ratings that experienced project team members believe are appropriate to manage project and operational risks successfully. For example, using a High likelihood threshold of 75% produced too many High risk results that didn't seem sufficiently critical to warrant that rating. Given the reviewer's comments, however, we will re-evaluate our thresholds and rating definitions and adjust if appropriate.	The OLCF re-evaluated the rationale for the range of risk likelihood. We believe that these ranges provide the accuracy needed for effective risk management.
To ensure that adequate reserves are in place, the OLCF should consider performing a more detailed cost impact/exposure estimate for—at a minimum—the three high-level risks (i.e., Funding uncertainties, Lustre support model change, Metadata performance). The intent is to ensure operations are not impacted should all three be realized, or at a minimum, have a plan in place to minimize impacts to operations.	The OLCF will perform more detailed analyses as recommended.	Cost/impact estimates for funding uncertainties have been assessed by laying out possible budget scenarios, including a conservative estimate. This is described in more detail in Sections 5 and 7. Cost/impact estimates for Lustre support model change and Metadata performance are described in Section 7.

7. Does the OLCF have a valid authority to operate?

Recommendation	August 2010 ORNL Action/Comments	Updated
The OLCF should consider a brief summary of reportable incidents and the corresponding resolutions for the past year.	This information is provided to HQ through our standard weekly reports, for example, during the IPT conference calls. We don’t publicize our methods of response to security incidents and, since the OA report is a public document, it would not be appropriate for us to include this type of information.	N/A

8. Are the performance metrics for the next year proposed by the OLCF reasonable?

Recommendation	August 2010 ORNL Action/Comments	Updated
This format should be used by all 3 centers. It is very clear.	The OLCF would be happy to work with the sites and HQ on the format of future OA reports.	OLCF management provided input to the Program Office as part of a three-site collaboration.
Add to Customer Metric 1: The OLCF will report on the survey results to DOE by March of the following year and will include a breakdown of the results by INCITE, ALCC and Discretionary projects	The OLCF will work with the Program Manager to determine the desired user survey reporting intervals and format.	Initial survey information was provided earlier in the year to the IT Project Manager for inclusion in a report to DOE. The breakdown of the results by program type are provided in Section 2.
Add to Customer Metric 2: The numbers will be reported for each quarter to DOE in the quarterly Customer Results report, and annually in the Operational Assessment in August.	The OLCF will work with the Program Manager to determine the desired problem-ticket-resolution reporting intervals and format.	This information is provided to the IT Project Manager on a monthly basis, for inclusion in reports to DOE.
Add to Customer Metric 3: The OLCF will track its workshops, tutorials, monthly user teleconferences and application support provided to users and will provide quarterly reports to DOE.	The OLCF currently tracks this information and will provide it quarterly to DOE.	N/A
Additional metric: Business Results Metric 4: Resource Utilization and Failure Tracking – Utilization, mean time to interrupt (MTTI), and mean time to failure (MTTF) will be tracked and reported for OLCF resources.	The OLCF currently tracks and reports utilization, MTTI, and MTTF.	N/A
Additional metric for Cyber Security Metrics (V11): The OLCF will report their “reportable” cyber security incidents (providing a brief summary of the incident and the resolution for each reportable incident for the past year.)	This information is provided to HQ through our standard weekly reports. We don’t publicize our methods of response to security incidents and, since the OA report is a public document, it would not be appropriate for us to include this type of information.	N/A

Recommendation	August 2010 ORNL Action/Comments	Updated
Add to existing metric for Risk Management (VI): The OLCF will provide information about the development, evaluation, and management of the top five to seven operating and technical risks encountered during the previous year. It will also provide projections for the top operating and technical risks that it expects to encounter in the next FY.	We are currently working to include more explicit risk cost analyses in our risk management efforts and will include this in next year's report. We will also extend our reporting to include expectations of out-year risks as well.	More explicit analyses are now included as part of the risk management process and are documented in the risk register (e.g., residual exposure analysis).
Replace Financial Performance metric with new metric: The OLCF will provide monthly reports on steady-state (SS) and Development, Modernization and Enhancement (DME) costs to compare against plans as described in the OMB300. Reporting will include the following: • How well the program is executing to the cost baseline established during the previous year's Budget Deep Dive, with an explanation of any major discrepancies. • Results and projects generated using methodology developed with the concurrence of the Program Manager, demonstrating operational cost effectiveness. • A financial sheet which delineates effort, lease, operations and DME; the sum will add up to the facility total budget • Lines showing staffing levels (in FTEs) for both DME and SS.	The Financial Performance metrics that have been requested by DOE are already provided in a monthly report to HQ.	N/A

2. USER RESULTS

CHARGE QUESTION 2: ***Are the processes for supporting the customers, resolving problems, and communicating with key stakeholders and Outreach effective?***

OLCF RESPONSE: The OLCF has a dynamic user support model based on continuous improvement and a strong customer focus. A key element of the program is an annual user survey developed with input from qualified survey specialists and the DOE program manager. OLCF users have consistently stated they have been very satisfied with the facility and its services. In keeping with goals for continuous improvement, four metrics perceived as indicative of good customer support (see below) have been extracted from the survey and are reported to DOE as indicators of user satisfaction. The OLCF continues to implement and maintain operational activities designed to provide technical support, training, and communication to current users and the next-generation of researchers. 402 users responded to the 2010 user survey (36 percent of individuals who were contacted by the OLCF). Reference Table 2.1 for response rates.

2011 Operational Assessment Guidance – User Results

For each of the following metrics, the Facility reports the results and provides projections using methodology developed with the concurrence of the DOE Program manager. The following categories have data that come from the user survey:

- User Satisfaction - reports the results of the Facility's yearly user survey, which provides feedback on the quality of its services and computational resources; and
- Problem Resolution - summarizes user requests for assistance and their resolution.

In addition, the Facility reports on the following categories that give the Center staff the opportunity to share their experiences with their users and stakeholders:

- User support and outreach - highlights and appropriate number of user support stories; documentation of training and workshops (this category may or may not be part of the user survey); and
- Communications with key stakeholders - summarizes efforts in these areas.

The Facility conducts an annual survey using the following methods:

- The survey shall be developed in conjunction with survey experts and have questions that cover the applicable categories of User Results.
- The survey is open to all users, with the explicit exception of a) vendors that have user accounts as part of a service agreement with the facility, and b) those who could be viewed as having a staff role;
- The Facility will negotiate the target response rates for the user survey with their DOE Program Manager. The Leadership Computing Facilities will include sufficient demographic information such that the report can describe results by INCITE, ALCC, and Discretionary allocations.
- Satisfaction questions on the survey are reported on a scale agreed to with the Facility's DOE Program Manager. The Facility also has an agreement with its Program Manager as to what constitutes a satisfactory score;

- The Facility will report metrics for the previous year where similar measures were gathered; and
- The Facility will include statistical analysis of the results. This shall include basic measurements such as mean, and an assessment of the quality of the sample using, e.g., the variance, standard deviation, or result of a t-test.

2011 Approved OLCF Metrics – User Results

Customer Metric 1: Customer Satisfaction

Overall OLCF score on the user survey will be satisfactory (3.5/5.0) based on a statistically meaningful sample.

The 2010 OLCF survey overall satisfaction rating was 4.3 out of a possible 5.0. This rating of "very satisfied" mirrors the results of the 2009 survey.

Annual user survey results will show improvement in at least ½ of questions that scored below satisfactory (3.5) in previous period.

None of the user responses in the previous period (2009 user survey) were below the 3.5 satisfaction level.

Customer Metric 2: Problem Resolution

80% of OLCF user problems will be addressed within three working days, by either resolving the problem or informing the user how the problem will be addressed.

In CY 2010, 91.2% of user tickets were either resolved or information about how the problem would be resolved was provided in 3 working days—a 5% improvement over the previous result (2009) In CY 2011 YTD, 89.5% of queries were addressed within 3 working days (Reference Section 2.2).

Customer Metric 3: User Support

OLCF will report on survey results related to user support.

The OLCF does not have a survey question specifically targeted at the full range of user support from OLCF staff members, and instead solicits an overall user satisfaction rating and comments about support, services, and resources. Representative comments and descriptions of user support and outreach and communications with key stakeholders from July 1, 2010, through June 30, 2011, are described below.

The OLCF has developed and implemented a dynamic, integrated customer support model. It comprises various customer support interfaces, including user satisfaction surveys, formal problem resolution mechanisms, user assistance analysts, and scientific liaisons; multiple channels for communication with users, including the OLCF User Council; comprehensive training programs and user workshops; and tools to reach and train the next generation of computer scientists.

Through a team of communications specialists and writers, the OLCF produces a steady flow of reports and highlights for potential users, the public, and sponsoring agencies. The Oak Ridge facility has expanded this mode of outreach through an internship program for science writers: by working alongside senior science writers at the facility and with computational researchers, these interns gain a more thorough understanding of the impact of leadership computing, and this is translated into more insightful news stories as these students transition to other media outlets. The OLCF communication infrastructure has been identified by ORNL as a best practice and other ORNL facilities (for example, the Spallation Neutron Source) are currently exploring ways to implement similar groups.

The OLCF recognized early on that users of HPC facilities have a range of needs requiring a range of solutions, from immediate, short-term, “trouble-ticket-oriented” support such as assistance with debugging and optimizing code to more in-depth support requiring total immersion in and collaboration on projects. The OLCF responded with two complementary OLCF user support vehicles: the User Assistance and Outreach Group (UAO) and the Scientific Computing Group (SciComp), which includes the scientific and visualization liaisons. Scientific liaisons are a unique OLCF response to high-performance scientific computing problems faced by users (examples of their support are provided in Sections 4.2 and 4.3).

The OLCF offers many training and educational opportunities throughout the year for both current facility users and the next generation of HPC users (Section 2.3). This year, the OLCF’s contributions in this area were recognized with several awards, discussed in Section 2.4.2.

As discussed above, the OLCF uses a variety of methods to reach out to our customers and measure user satisfaction throughout the year, but the annual user survey is by far the most comprehensive feedback mechanism used. The survey consists of 50 questions, comprising a mixture of ratings and open-ended questions. While the ratings are important, a high value is placed on the specific comments made by the users in the open-ended questions, which often provide very good suggestions for improving the user experience or identify issues staff members were not aware of until they were identified in the survey. To this end, UAO staff members comb through the survey each year to identify items to follow up on. Section 2.1 describes the survey results in detail, including some of the more dynamic examples of this proactive approach to user suggestions. Further input is also solicited by, and provided to, OLCF staff members through direct interactions, scientific support, tickets, and so on.

2.1 EFFECTIVE USER SUPPORT

2011 Operational Assessment Guidance – User Support

The OA metrics for High Performance Computing (HPC) Facilities user support as assessed by the annual user survey are:

- Overall satisfaction rating for the Facility is satisfactory;
- Average of all user support questions on user surveys is satisfactory; and
- Improvement on past year unsatisfactory ratings as agreed upon with the Facility’s DOE Program Manager

A multifaceted approach is used to measure the effectiveness of the OLCF customer support model.

- A yearly survey measures customer satisfaction in key areas;
- A ticket management system ensures all queries are responded to in a timely manner; and
- OLCF staff members solicit feedback directly from stakeholders through various formal and informal interactions.

The OLCF User Survey

The OLCF conducts an annual survey of all users to solicit feedback on the quality of our customer service and computational resources. The survey is conducted by an independent third party, the Oak Ridge Institute for Science and Education (ORISE), using questions developed by the OLCF in collaboration with the DOE OLCF program manager and with input provided by ORISE. The surveys, which contain 50 questions, are sent electronically to all individuals with active accounts (1,116 this year, excluding OLCF staff and vendors); periodic reminders are sent to nonresponders. Survey results are validated using a streamlined version of the Delphi Technique, a set of guidelines for remote gathering of information from experts.

For 2010, the last survey conducted, 402 out of a total of 1,116 users responded to the survey for a response rate of 36%. While this is slightly lower than last year's response rate (37%), it is well above the average for such surveys[1], and the total number of responders actually increased from 261 to 402 due to a higher number of users. Reference Table 2.1 for the 2008-2010 User Survey Participation results.

Table 2.1 User Survey Participation

	2008 Survey	2009 Survey	2010 Survey
Total Number of Respondants (Total percentage responding to survey)	226 (48%)	261 (37%)	402 (36%)
New Users (OLCF User < 1 Year)	41%	29%	31%
OLCF User for 1–2 Years	27%	36%	29%
OLCF User > 2 Years	32%	35%	40%
Used User Assistance Center at least 1 time	82%	74%	80%

The OLCF took a number of measures to encourage good participation. The project director, Arthur (Buddy) Bland, sent a notice out to all of the users emphasizing the importance of the survey. OLCF User Council chair, Balint Joo of the Thomas Jefferson National Accelerator Facility, also sent a notice to all users on behalf of the council. In his note he said,

> *"No doubt, you have received messages asking you to participate in the 2010 OLCF User Survey. We, the members of the OLCF User Council, would like to add our voice, and urge you to participate.*
>
> *Taking part in the survey is really a service to yourself: it is an important opportunity for you to express your views and feelings about the services provided to you by the Oak Ridge Leadership Computing Facility. OLCF Center staff truly value your feedback and*

[1] Response rates to surveys are difficult to predict as they are based on various factors; however, the average response rate for similar surveys appears to range from 10% to 30% [see for example, "Survey Response Rates" (http://www.peoplepulse.com.au/Survey-Response-Rates.htm); Hamilton, white paper, 2009 (http://www.supersurvey.com/papers/supersurvey_white_paper_response_rates.pdf); Baruch and Holtom, Human Relations, 61(8), August 2008, pp. 1139–60].

> *strive to improve their service to you based on, amongst other things, the results of this survey. While you can always get help from the helpdesk for technical problems, if there are big picture issues you would like resolved, or if there are additional ideas which may benefit other users, the survey is the place where you can make these known."*

In addition, Ashley Barker, the UAO Group Leader, and a member of the ORISE survey team also sent out reminders. Lastly, the scientific liaisons reached out to the principal investigators (PIs) to encourage their participation. Each of these efforts demonstrated a measurable and immediate increase in the number of returned surveys.

Survey respondents were asked to classify the program types with which they were affiliated. Reference Table 2.2.

Table 2.2 User Survey Responders by Program Type

Program	Percentage[a]
INCITE[1]	62
Director's Discretionary	25
Other[2]	25
ALCC[3]	2

2.1.1 Overall Satisfaction Rating for the Facility

Users were asked to rate satisfaction on a 5-point scale, where a score of 5 indicates a rating of very satisfied and a score of 1 indicates a rating of very dissatisfied. The metrics agreed upon by the DOE OLCF Program Manager define **3.5** to be **satisfactory**.

There was an explicit question on the 2010 survey about the overall satisfaction rating for the Facility. From the 402 responses, the calculated mean was 4.3 out of 5.0, well above the stated metric of 3.5. Key indicators from that survey, including overall satisfaction are shown in Table 2.3. These are summarized, and broken out by Program.

Table 2.3 Satisfaction Rates by Program Type for Key Indicators

Indicator	Mean	Program[a]		
		INCITE	**ALCC**	**Director's Discretionary**
Overall Satisfaction with the OLCF	4.3	4.3	4.0	4.3
Helpfulness of User Assistance Staff	4.3	4.3	4.2	4.4
Overall System Performance of the XT5	4.0	3.9	3.7	4.0

[1] Innovative and Novel Computational Impact on Theory and Experiment
[2] Reflects uncertainty about program type
[3] Advanced Scientific Computing Research Leadership Computing Challenge

2.1.2 Average Rating Across All User Support Questions

The calculated mean of all answers to all user support questions on the 2010 survey was 4.27 out of 5.0, indicating that OLCF exceeded the 2010 user support metric. Sample comments, shown in Table 2.4, indicate that users are very satisfied with OLCF customer service and computational resources.

Table 2.4 Sample User Comments from the 2010 Survey

"At the human (support) and technical (software, admin) level, OLCF is a first-rate institution."
"Project staff experiences when contacting OLCF support have been very positive. Support staff seems to be very customer oriented and works hard to maximize the customer experience. I appreciate the comments provided by subject matter experts and the proactive approach of reaching out to users via telephone conference calls and on-site meetings."
"The help services provided by OLCF are the best I have ever experienced in over a decade of interaction with multiple supercomputer centers."
"The facilities at OLCF are world class."
"The overall size of the system and the correspondingly larger allocations of CPU time have continued to enable us to push the boundaries of what is possible in the field of turbulent combustion science."
"Machines are excellent to compute on, good allocation and accessibility."
"Excellent support."
"The user support I received over the telephone was outstanding."
"This is such an extreme edge area where everyone is learning together. The amount of help that User Assistance can provide is really quite excellent given these conditions."
"User Assistance is doing an excellent job."
"I feel the website is pretty good and useful."
"Help desk is excellent; System status page is extremely valuable, Large computer system with relatively good turnaround and ability to run both moderate and huge jobs."

Statistical Analysis of the Results

Statistical analysis of four key survey areas is shown in Table 2.5. These reflect overall Facility satisfaction, services, and computational resources.

Table 2.5 Statistical Analysis of Key Results

	Overall Satisfaction	Helpfulness of User Assistance Staff	Effectiveness of problem resolution	Overall System Performance of the XT5
Number Surveyed	1116	1116	1116	1116
Number of Respondents	375	333	336	323
Mean	4.3	4.3	4.2	4.0
Variance	0.6	0.9	0.9	0.7
Standard Deviation	0.8	0.9	0.9	0.9

2.1.3 Improvement on Past Year Unsatisfactory Ratings

Each year the OLCF works to show improvement in at least half of the questions that scored below satisfactory (3.5) in the previous year's survey. All questions scored above 3.5 in both 2008 and 2009, and only one item scored below 3.5 in 2010. This item was related to the frequency of unscheduled outages on the XT5. (Reference Section 3 for the OLCF response to unscheduled outages.)

Soliciting Feedback for Areas for Improvement

Because the surveys are one of the tools we use to continually improve operations, users are also asked a few open-ended questions to solicit feedback on our strengths and specific areas for improvement. In response to an open-ended question about the best qualities of the OLCF, thematic analysis of user responses identified the following as the top three.

- Great staff and support (37% of responses)
- Powerful/fast machines (33% of responses)
- Large computational capacity [17% of responses (overlap with "powerful/fast machines")]

In the 2010 survey, the following areas for improvement were cited the most frequently.

- Reliability/Stability (23%)
- Data Transfer Rate (15%)
- Queuing Policies (13%)

The response to these requests for improvement from our user community are summarized as follows:

Reliability/Stability

The OLCF reviewed the specific comments made related to reliability/stability. The following comments are representative of the majority of comments on this issue. Reference Section 3 for a discussion of the actions taken to address these concerns.

> *"Jaguar XT5 was very stable in Spring 2010, but then was quickly aged, by the time of reaching fall, the system had too many unscheduled outages due to node issues and/or file system issues, which made it very difficult to run full machine scale job for more than 2-hours (our full machine 24-hour job crashed 9×)"*
>
> *"Reduce unscheduled outages"*

Data Transfer Rate

The OLCF reviewed the specific comments made related to data transfer rates. Most of the comments centered around the performance of the Lustre file system, including the comments below. "

> *"Improve performance of Lustre file system"*
>
> *"My biggest headache this year has been I/O performance"*

Several initiatives to improve I/O performance were undertaken this year. The OLCF worked with application teams to improve the scalability of their application inputs/outputs (I/O). The Center also installed two additional file systems to reduce shared resource contention, increasing both aggregate metadata performance and bandwidth.

Beginning in May 2011, the OLCF began delivering substantially improved I/O performance of the Spider parallel file system after implementing a congestion control mechanism for the Spider parallel file system known as *fine-grained routing*. These performance improvements are illustrated in Figure 2.1.

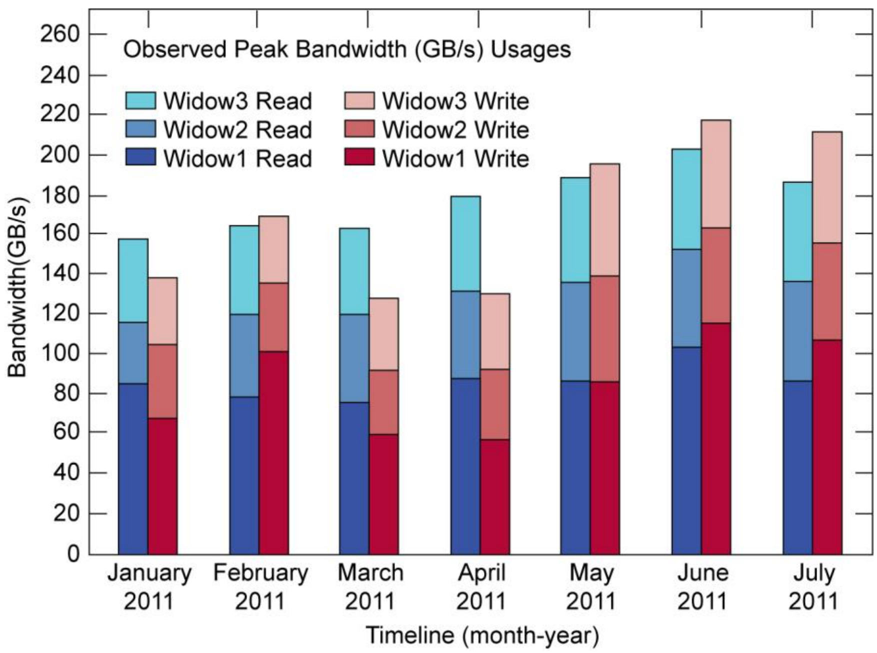

Figure 2.1 The Effect of Fine-grained Routing on I/O Performance.

The results demonstrate substantial improvements in file system write performance with targeted scientific simulations achieving over 89% of best-case write performance. Fine-grained routing provides a mechanism to control the path of file system related network I/O, providing an optimized path for these I/O flows on the Cray SeaStar2+ and InfiniBand networking infrastructure. Further information is available as an ORNL technical report via http://info.ornl.gov/sites/publications/Files/Pub30140.pdf.

Last, the OLCF entered into a subcontract with Whamcloud to improve metadata performance in Lustre. While the results are not fully ready for production, the Center has seen substantial performance improvements during testing on Jaguar XT5. The goal is to introduce these metadata performance enhancements into production by the end of the year 2011. For additional descriptions of Lustre-related activities, Reference Section 6.2.

Queuing policies

The OLCF reviewed the specific comments made related to the queue policies. The following comments are representative of the majority of these.

> *"I would first like to remark positively on the queuing policy, which prioritizes very large runs, is an excellent and unique feature of the OLCF that enables calculations that are unthinkable elsewhere. Typically before we get to the stage of being ready to compute at this scale we need to run many smaller runs with much lower core count, but we still need these to turn around quickly to enable eventually running the larger runs. Another similar issue is runs for post-processing. Although these runs are relatively short, again we must do many of them because we develop new conceptual approaches and tools to essentially every run-set we do, and this development occurs iteratively as ideas are solidified. (We do not apply a standard analysis to each run-set.) Some way of prioritizing these types of pre- and post- processing steps, which are essential to the overall scientific goals, could be useful, though I am not sure how to implement it without compromising the ability to perform huge runs requiring a large fraction of the machine."*

> *"Great service. It would be useful to have a benchmark queue which would allow for running longer on smaller number of cores (scaling studies often run in the 2h limit)."*

> *"Sometimes, I want to run a small job using several hundreds of cores without a long queue time."*

The queuing policy and its effect on smaller jobs has been an ongoing issue. Because DOE's goal is to enable high-impact, grand-challenge research that could not otherwise be performed without access to the leadership-class systems, to ensure that its leadership facilities are meeting the objectives of this goal, DOE has established certain usage targets for leadership-class jobs on these systems. To meet these targets, the OLCF has adopted queuing policies that heavily favor large jobs. It is a delicate balance that must be constantly monitored to ensure that the needs of all users are met, along with the national goals for a leadership computing facility.

The Center recognizes that there is often a need for smaller jobs, such as pre- and post-processing for large runs. For that reason, small jobs are not prohibited from using the system. They are, however, limited to prevent them from impacting larger leadership and national goal runs. Additionally, in some cases, small jobs have higher per-processor memory requirements than larger-scale jobs. These are often ideal for smaller cluster-based systems as the workload (both the smaller jobs on clusters and larger jobs on massively parallel resources) makes more efficient use of the resource by more accurately matching its capabilities. This type of input from users that are running small- to medium-sized jobs is essential to optimal planning and use of leadership-class machines as it can be used to better understand how those computing needs can be maximally met on a leadership machine, maximizing the potential of all leadership-class machines. In addition to communicating the DOE OLCF goals and how they impact small runs to users, the OLCF is currently investigating options to ensure this issue is addressed optimally through queuing policies.

Other User Comments and OLCF Actions

The OLCF takes all survey suggestions, as well as feedback received through other channels (e.g. tickets, the User Council, interactions with OLCF staff members, etc), very seriously. The following additional actions were taken this past year by the OLCF based on other survey suggestions and feedback received from users.

1. A few survey respondents indicated some dissatisfaction with the turnaround time for getting an account on the system.

   There are several steps involved before a user can gain access to OLCF resources. The OLCF recognizes these requirements can take a while and it can be frustrating when users encounter delays in getting access to the system. This year the Center reevaluated the access procedures and policies and worked with the relevant support groups at ORNL to streamline the Personnel Access System (PAS) processes for creation of user accounts. Previously carried out for all foreign national users AND users on data-sensitive projects, PAS entries will now be focused on foreign national users from sensitive countries, as well as foreign national users from non-sensitive countries working on data-sensitive projects. If the user is employed by a US national laboratory, an exception will be made. This has been approved by the relevant ORNL support groups, including the OLCF cyber security team, and should cut down significantly on the time-to-access. The Center will continue to monitor the access procedures to improve the time it takes to gain access to a project.

2. A few survey respondents requested more information on getting started.

   A “getting started” page has been created for new users (or as a refresher for existing users). The page can be found from the OLCF Home Page at http://www.olcf.ornl.gov/support/getting-started/. The page covers the general steps to use the OLCF systems from connecting to running batch jobs and the steps a user should take to request an allocated project and/or join an allocated project.

3. A few survey respondents requested more information on batch scripts for the XT5.

   A knowledge base article containing example XT5 batch scripts has been created for the OLCF support site: http://www.olcf.ornl.gov/kb_articles/xt-batch-script-examples/. The article covers a number of basic scenarios and is meant to provide basic building blocks for actual cases that may be more complicated.

2.2 PROBLEM RESOLUTION

2011 Operational Assessment Guidance – Problem Resolution

The OA Metrics for Problem Resolution are:

- Average satisfaction ratings for Problem Resolution related questions on the user survey are satisfactory or better; and
- At least 80% of user problems are addressed (the problem is resolved or the user is told how the problem will be handled) within three working days.

The OLCF uses Request Tracker software (RT) to track queries and ensure that response goals are not missed. In addition, the software collates statistics on tickets issued, turnaround times, etc., to produce weekly reports, allowing the OLCF staff to track patterns and address anomalous behaviors before they have an impact on additional users. The OLCF issued more than 2,800 tickets in response to user queries for CY 2010 (Figure 2.2). The team exceeded the resolution time metric:

- 94.9% of queries were addressed within 3 working days (target metric is 80%),
- the average response time for a query was 24 minutes.

The CY 2011 YTD problem resolution metric is also on track to exceed the targeted 80% response:

- 89.5% of queries were addressed within 3 working days,

- the average response time for a query was 27 minutes.

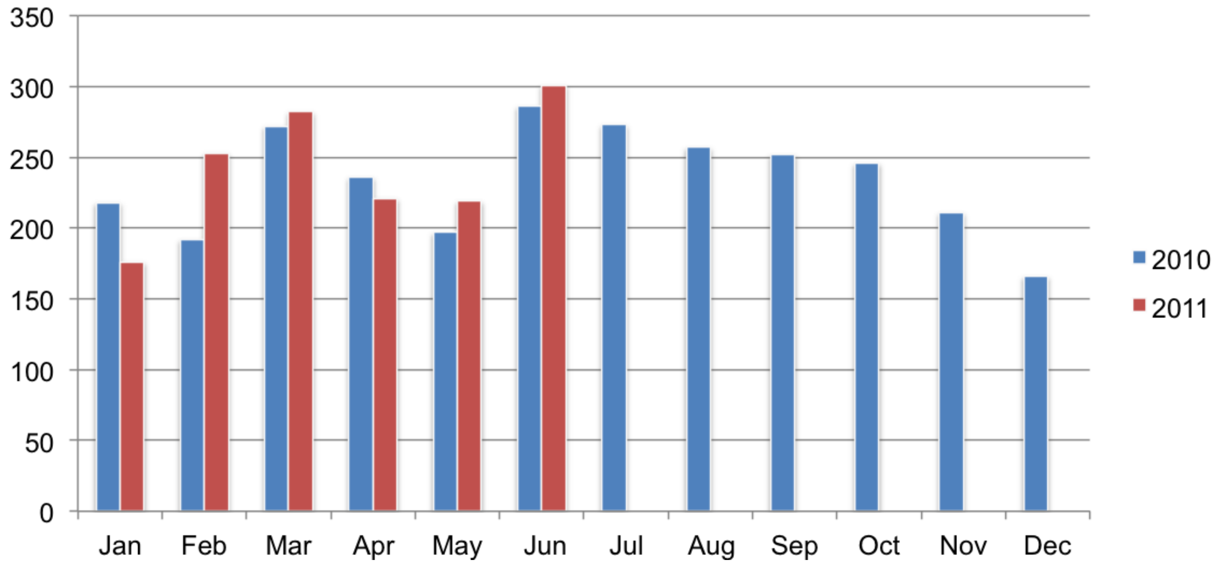

Figure 2.2 Number of Helpdesk Tickets Issued per Month.

Each query is assigned to one user assistance or account analyst, who establishes customer contact and tracks the query from first report to final resolution, providing not just fast service, but service tailored to each customer's needs. While UAO is dedicated to addressing queries promptly, user assistance and account analysts consistently strive to reach the "right" or best solution rather than merely a quick turnaround. Tickets are categorized by the most common types (Figure 2.3).

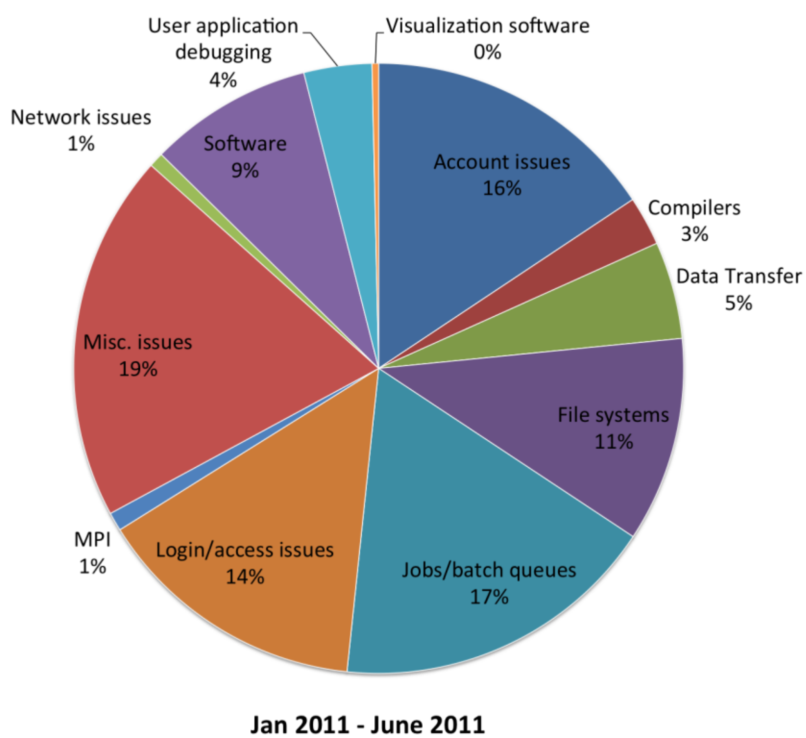

Figure 2.3 Categorization of Helpdesk Tickets

UAO's regular ticket report meetings, discussed in last year's report, are another OLCF innovation that has paid huge dividends in efficient customer service. Because of the information shared in these meetings the OLCF has maximized the impact of the staff far beyond their numbers. One outcome from ticket meetings this past year was the creation of new mobile phone apps for users that show the status of the machines. UAO analysts developed the apps for both the Android and iPhone platforms. In addition, UAO developed "opt-in" notice lists that provide automated notices about the status of OLCF systems, as well as more detailed updates from the OLCF staff as needed. Users can subscribe to receive notifications about particular systems short- or long-term (e.g., for as little as 1 week or for an entire calendar year). Thus, users now have numerous ways to check the status of the machines including checking the website, via email or Tweets, and/or on their mobile phone devices.

In addition, UAO members also routinely provide the following types of support to OLCF users.

- Establishing accounts and responding to account issues.
- Helping users compile and debug large science and engineering applications.
- Identifying and resolving system-level bugs in conjunction with other technical staff and vendors.
- Installing third-party applications and providing documentation for usage.
- Engaging center staff and/or users to ensure all users have up-to-date information about OLCF resources and to solicit feedback.
- Researching, developing, and maintaining reference and training materials for users.

2.3 USER SUPPORT AND OUTREACH

2011 Operational Assessment Guidance – User Support and Outreach

The OA data for User Support include:

- Summary of training events, including number of attendees, and success results where possible.

The OLCF provided the following specific training- and outreach-related workshops and seminars since the last operational review. A summary of these events is shown in Table 2.6.

Table 2.6 User Training and Workshop Event Summary

Event Description	Date	Participants
SciApps10 - Challenges and Opportunities for Scientific Applications	Aug 3-6, 2010	63
Introduction to CUDA	Jan 20, 2011	15
Exascale Workshop	Feb 22-23, 2011	58
OLCF Spring Training	Mar 7–10, 2011	80
OLCF User Meeting	Mar 11, 2011	43
INCITE Proposal Writing Seminar	Mar 21	38
Lustre User Group Meeting	Apr 12–14, 2011	163
Vampir Training Class	May 17, 2011	25
HPC Fundamentals	Summer 2011	44
Visualization with VisIt 2011	Jun 14, 2011	44
Crash Course in Supercomputing	Jun 16, 2011	112
Introduction to OLCF-3 Webinar	Jul 26, 2011	74
LCF Seminar Series: Femtoscale on Petascale: Nuclear physics in HPC, Hai Ah Nam, ORNL	Sept 21, 2010	32
LCF Seminar Series: Massively parallel simulations for industrial applications—multiphase injection, Anne Birbaud, GE Global Research	Oct 29, 2010	38
LCF Seminar Series: Temporal Debugging via Flexible Checkpointing: Changing the Cost Model, Gary Cooperman, Northeastern University	Jan 25, 2011	40

The following sections focus on significant highlights from the OLCF communications, outreach, and training programs over the past year.

Award Winning Science Communication

—Nothing in life is more important than the ability to communicate effectively. (Gerald Ford)

The OLCF recognizes that it is not just in the computing business, but also the communication business. An important aspect of its mission is the communication of results. To this end, the OLCF provides a wide range of communications products to current and potential users, the general public, and sponsoring agencies, including the annual report, ASCR (Advanced Scientific Computing Research) News Roundup highlights, articles for popular/generalist journals, and brochures.

Since 2006, the OLCF has employed science writers to communicate the facility's scientific and technical accomplishments to general audiences, which include the public, whose taxes make the research possible; journalists, who broadcast the OLCF's messages more widely; policy makers, who want good information on which to base recommendations; DOE program managers, who serve as guardians of the public investment in science; students, who fill the pipeline to provide the next generation of scientists and engineers; and partners in industry, academia, and government. The writers mainly produce news and feature articles, press releases, annual reports, newsletters, and video scripts. Their work appears on DOE websites such as those of ORNL, NCCS, OLCF, ASCR, and DOE headquarters; in trade or specialized publications such as *HPCwire* and *ORNL Review*; and in mainstream venues such as newspapers, magazines, and exhibits at museums and trade expos. More than 28 science stories, 19 releases to external media, and 21 write ups on Center activities, events and awards were prepared and released in CY 2010. The Center writers also produced the *INCITE in Review* and 2009/2010 OLCF Annual Report and well as contributed to the monthly ASCR News Roundup and biweekly OLCF Snapshots to DOE. As in past years, each of the OLCF's science writers was once again received the prestigious Magnum Opus award. The article, "Jaguar Pounces on Child Predators" won silver and both "Earthquake Simulation Rocks Southern California" and "Exploring the Magnetic Personalities of Stars" won honorable mentions in the category, Electronic Publications or Website, Best Feature Article. This year there were more than 550 entries, with winners coming from organizations such as Walt Disney, American Airlines, and Proctor & Gamble.

OLCF User Council

The OLCF User Council provides a forum for the exchange of ideas and development of recommendations to the OLCF regarding the Center's current and future operation and usage policies. The User Council is made up of researchers who have active accounts on the leadership computing facility compute resources. The council meets via a teleconference on a monthly basis. The current User Council is chaired by Balint Joo. The council has been very engaged and provided valuable input to OLCF management this past year. Following are some of the items discussed in and contributions of the User Council this past year.

- Balint Joo joined Arthur Bland in representing the OLCF at the NUFO User Science Exhibition on Capitol Hill. The event was organized to highlight the significant and important role that scientific user facilities play in science education, economic competitiveness, fundamental knowledge, and scientific achievements. The Center contributed a poster that highlighted both the science and the Center resources and provided video images of the facility. Attendees at this public exhibit included Congressional leaders and their staff members; management from the DOE Office of Science (SC); four national laboratory directors, including ORNL Director Thom Mason; a representative from the National Science Foundation; and representatives from a number of science agencies or societies such as the American Physical Society, the American

Institute of Physics, the Federation of American Societies for Experimental Biology, *Physics Today*, the Coalition for National Science Funding, ASTRA, and the American Astronomical Society.

- The User Council reviewed the 2010 survey and provided suggestions on how to increase user participation. One suggestion included sending an email on behalf of the User Council asking for participation. We received 96 additional responses immediately following the email from the User Council.
- The User Council reviewed the updated OLCF website before it was released to general users and provided input.
- The User Council provided input into the curriculum for the Spring Training class. Specifically, the council recommended that the Center provide more hands on training. The OLCF's Bobby Whitten organized breakout sessions during the day to meet this request. The survey results from the training class indicated that users liked the breakout sessions.
- The User Council volunteered to be early testers of the WebEx software and began using it for council meetings, helping to work out the bugs before it was put into production for general users.
- The User Council tested the opt-in email lists before they were released to general users. No issues were found, but the council provided positive feedback on the lists to UAO.
- The User Council provided input into the queue policy change made at the end of 2010.
- The User Council requested that the Center add the status of Data Transfer Nodes to the online status page. This request has been completed.
- The User Council requested an extension on the amount of time before an RSA fob becomes disabled for nonuse. The Center extended the time from 6 months to 1 year. This reduced administrative loads to reactivate accounts for principal investigators that log in infrequently.
- The User Council asked the Center to consider adding a logon message on system status for users entering their passkey information so that if the system is down or having issues, the user can attribute a failed login to the machine rather than an incorrect password. Implementation of this request is under way.
- In September 2010, two additional file systems were added to Spider, the OLCF's centerwide file system. The new file systems require regular preventative maintenance during which parts or all of the system are unavailable. The Center presented the User Council with downtime options to help determine which would be more favorable for users. With their guidance, the Center was able to come up with a schedule for preventative maintenance for the new file systems that was more favorable to users.

Web Resources

This past year UAO deployed a dynamic new website (http://olcf.ornl.gov) to highlight the science, technology, people, and activities of the OLCF and provide enhanced access, information, and services, including system information and statistics, OLCF project details, an online newsletter, and videos. In addition, the OLCF site provides users with allocation and account assistance, education and training modules, and a robust knowledge base. UAO also designed and implemented a new training guide for Jaguar to help users find information more quickly.

A few of the survey respondents indicated that they would like more visibility of the system status pages on the OLCF website. To provide more visibility, the system status page can now be found in multiple places.

Some of the users also indicated that the site where they can check their project usage is hard to find. UAO added links to this page from multiple places on the OLCF site to make it easier for users to get to the other site (the sites have to stay separate because the project usage site requires a login).

User Workshops and Related Outreach Activities

Workshops and seminar series are another important component of the customer support model. They provide an additional opportunity to communicate and act as a vehicle to reach out to the next generation of computer scientists. OLCF outreach to train current and future scientists and engineers is summarized in Table 2.6. The OLCF also provides tours to groups throughout the year for visitors that range from middle-school students through senior-level government officials. The OLCF provided tours for 953 distinct groups in CY 2010 and 395 groups in CY 2011 (YTD).

The OLCF began live webcasting of workshops and seminars earlier this year to broaden participation. These webinars are recorded and will be published on the enhanced OLCF website. A survey is conducted immediately following each event, and the OLCF will begin querying participants and users about types of webcasts they would find most effective and valuable.

In addition, a more comprehensive education program has been initiated, including the 10-minute tutorials series, HPC fundamentals series, graphics processing unit (GPU) series, and advanced-topic series. The 10-minute tutorials are recorded screencasts of common technical tasks that OLCF users perform (e.g. the top ticket topics). The OLCF will solicit feedback in the coming year from the User Council as well as the users about the 10-minute tutorial series. The HPC fundamental series will target new users who wish to expand their knowledge about common HPC topics. The GPU series is designed to support the Titan project and prepare users for successfully using hybrid architectures. The advanced-topic series targets users who have a need to understand advanced programming models, debugging strategies, or optimization techniques.

Content generated for these and other education series will be combined into online training materials that will be made available on the enhanced OLCF website in the coming year.

INCITE Proposal Writing Webinar

The OLCF and Argonne Leadership Computing Facility (ALCF) cohosted a series of webinars, guiding researchers through the proposal process for earning time on the two facilities' leadership-class supercomputers. The webinars provide researchers with necessary information for writing a competitive proposal and using leadership-class systems and an opportunity to ask questions of the computing facilities' staffs.

Lustre User Group Meeting

As a leader in parallel file systems, the OLCF led the organization of the 2011 Lustre User Group (LUG) meeting. This was the first user-led LUG meeting, previously hosted by Oracle, and marked the transition of leadership to the broader user community. LUG 2011 provided a unique opportunity for Lustre users, developers, and system vendors to share knowledge and best practices related to the Lustre file system. With more than 160 attendees from more than 60 organizations, LUG 2011 was a tremendous success. Bull, DataDirect Networks (DDN), Dell, HP, LSI, Oracle, SGI, Terascala, Whamcloud, and Xyratex contributed to this collaborative event. The organizing committee was made up of representatives from Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Indiana University, Lawrence Livermore National Laboratory (LLNL), Naval Research Laboratory, Oak Ridge National Laboratory, Sandia National Laboratories, and the Texas Advanced Computing Center. "LUG 2011 is the first LUG that is completely community driven. It opens a promising new area in the Lustre community" said Jacques-Charles Lafoucrière, Chef de Service at CEA. The LUG offers participants opportunities to share knowledge, ideas, and achievements with a diverse audience." said Stephen Simms, Data Capacitor project lead at Indiana University.

Training the Next Generation

The OLCF maintains a broad program of collaborations, internships, and fellowships for young researchers. From July 1, 2010, through December 31, 2010, the OLCF supported more than 22 faculty, student interns, and postdoctoral researchers. Twenty-three faculty, student interns, and postdoctoral researchers were supported from January 1, 2011, through June 15, 2011. Of these, six were funded with ARRA funds. Six additional researchers will be funded with ARRA funds in the second half of 2011. OLCF interns and postdoctoral employees have contributed in a tangible way to OLCF projects and objectives, further demonstrating the quality of the learning environment provided. OLCF staff are engaged in many activities, both internally and around the country, to help reach the next generation of computer scientists and computational researchers.

DOE Recognizes OLCF Outstanding Mentors

The Department of Energy (DOE) recently awarded the Oak Ridge Leadership Computing Facility (OLCF) staff members Jim Rogers and Bobby Whitten with Outstanding Mentor Awards. Coordinated by the SC Workforce Development for Teachers and Scientists, the award recognizes mentors for their personal dedication to preparing students for careers in science and science education through well-developed research projects. Winners are nominated by their mentees.

Rogers, who is the director of operations for the OLCF, most recently mentored Nathan Livesey, a graduate of Oak Ridge High School and rising junior in the department of chemical engineering at Tennessee Technological University. Under Rogers' tutelage for two consecutive summers and a short stint during the winter of 2010, Livesey worked on facilities-related projects including the design of a database that captured the end-to-end design of the electrical systems supporting high-performance computers including the OLCF's Cray XT5 Jaguar. Rogers provided Livesey with space in his own office so that questions could be addressed without delay. Working with other divisions of the laboratory and different groups within the OLCF, Livesey deployed his system on a virtual machine for use by facilities and operational staff.

Whitten, a member of the OLCF UAO, acts as a mentor in two specific programs, one aimed at educators and the other at students. The DOE-sponsored ACTS (Academies Creating Teacher Scientists) program helps high school teachers grow as leaders of science, technology, engineering, and mathematics education by pairing them with mentors at national laboratories. Mentors provide these teachers with one-on-one training on how to better integrate the practice of science into their curricula. Whitten was paired with Rosalie Wolfe, a Network Systems teacher at Vinton County High School in McArthur, Ohio, who helped Whitten create a course in which students build a small supercomputer. Students in the ARC (Appalachian Regional Commission) program—also mentored by Whitten—tested this supercomputing course, gaining insight into how supercomputers work and how they are programmed. Since 2008, Whitten has mentored 22 students in both the ACTS and the ARC programs.

"Bobby is a great teacher, and I have learned so much from working with him this summer," said Wolfe of her experience with the ACTS program. "Bobby has provided me with a project that is within my capabilities, and yet at the same time challenging. [He] encouraged me to do research to learn programming languages I didn't even know existed, and yet when there was something I didn't understand or a problem that I couldn't solve, Bobby was there to provide 'hints' and encouragement that kept me from giving up."

High School Students Build Their Own Supercomputer—Almost—at OLCF

For the third straight year, students and teachers from around Appalachia gathered at ORNL this past summer for interactive training from some of the world's leading computing experts. The summer camp, a partnership between ORNL and the ARC Institute for Science and Mathematics, took place July 12–23, 2010. The OLCF hosted 10 students from various backgrounds and parts of the region.

The course was titled "Build a Supercomputer—Well Almost." And that they did. With the help of ORNL staff, collaborators, and interns from universities, the high school students went to work building a computer cluster, or group of computers communicating with one another to operate as a single machine, out of Mac mini CPUs. The students' cluster did not compute nearly as fast as the beefed-up machine right down the hall—ORNL's Jaguar—but successfully ran the high-performance software installed. Through the program students received a foundation in many of the things that make a supercomputer work.

"They get to learn HPC basics, and it's a chance for them to live on their own for a couple of weeks," said Bobby Whitten, an HPC specialist at ORNL and facilitator of the OLCF program. ORNL first partnered with ARC on a program of this type in 2008. Whitten happily notes that one of his students from that year is heading off to Cornell University in the fall to study biomechanical engineering.

Award Winning Science—Even at the Middle School Level

ORNL staff helped National Geographic's award-winning middle school science education program "The JASON Project" capture a prestigious CODiE award in early 2011 for the geology curriculum "Operation Tectonic Fury," described in the 2010 OLCF Operational Assessment report. This is a highly competitive, juried award for online educational publishers, game developers, and software programmers, presented annually by the Software and Information Industry Association. Operation Tectonic Fury won the Best Science or Health Curriculum category. JASON uses real world "explorers" to excite students and teach science curriculum: Oak Ridge researchers along with OLCF staff have provided time and expertise as "explorers." In Operation Tectonic Fury, ORNL host researcher Virginia Dale led the "mission" on weathering, erosion and soils. In addition to taking the students to Mount St. Helen's, Dr. Dale and team members also hosted students and teachers at ORNL to study soils under switchgrass in fields near Vonore. Students then visited the OLCF and EVEREST to learn how modeling and simulation with leadership systems is an important part of the process to study and understand the sustainability implications of energy crops. James J. Hack, director of the National Center for Computational Sciences also hosted JASON students and helped them gain a better understanding of the role of climate on our earth's ecosystem.

Ready, Set, Go!

On Monday, November 15, 2010 at Supercomputing 2010 (SC10), the starting gun was fired, and students began feverishly computing. For 47 hours, sleep was out of the question, caffeinated beverages were consumed like water, and the power of supercomputers was laid at the fingertips of eight teams vying to be known as the best of the next-generation of HPC. "We're having [students] run a high-performance cluster on the power it takes to run three coffee makers," said OLCF's Hai Ah Nam, computational scientist and technical chair of the SC10 Student Cluster Competition (SCC). Students had to build a computer cluster capable of running open-source software and meeting HPC Center benchmarks.

The competition had OLCF staff organizing, judging, interviewing, and getting to know the students throughout the week. “An organization like ours is unique because we address every aspect of HPC and span many science domains, which means we can provide these students 360 degrees of support,” Nam said. OLCF’s Jeff Kuehn, Bronson Messer, Arnold Tharrington, Rebecca Hartman-Baker, and Ilene Carpenter all served as scientific application judges for this year’s competition. The competition truly was international, with teams from National Tsing Hua University in Taiwan, Nizhni Novgoroad State University in Russia, Florida A&M University, Louisiana State University, the University of Colorado, the University of Texas at Austin, Purdue University in Indiana, and Stony Brook University in New York. Students were aided in their preparation for the competition by teaming with experts from the HPC industry. When the closing bell rang, National Tsing Hua University was declared the winner. In addition to the valuable experience that the students gain in the program, Nam said the competition is “building a computationally aware workforce,” and is a driving force for academia to develop and improve HPC curricula in the classroom.

2.4 COMMUNICATIONS WITH KEY STAKEHOLDERS

2011 Operational Assessment Guidance – Communications with key Stakeholders

The Facility summarizes the way it communicates with its Program managers, its users, and its vendors.

2.4.1 Communication with the Program Office

The OLCF communicates regularly with the Program Office through a series of established events. These include weekly IPT calls with the local DOE Oak Ridge office (DOE ORO) and the Program Office, monthly highlight reports, quarterly reports, the annual Operational Assessment, an annual Budget Deep Dive and the annual report. In addition, the DOE ORO and Program Office have access to tailored web pages that provide system status and other reporting information at any time.

2.4.2 Communication with the User Community

The role of communications in everything the OLCF does cannot be overstated, whether it is communicating science results to the larger community or communicating tips to users on using OLCF systems more efficiently and effectively. The OLCF uses various avenues, both formal and informal, for communicating with users. Formal mechanisms include the following:

- UAO and SciComp support services;
- weekly messages to all users on events;
- monthly OLCF User Council calls;
- quarterly user conference calls;
- annual users meeting;
- workshops and training events; and
- web resources such as system status and update pages, project account summaries, online tutorials and workshop notes, and other documentation such as “frequently asked questions” (FAQs).

2.4.3 Communication with the Vendors

OLCF conducts formal quarterly reviews of current and emerging hardware and software products with Cray Research. This includes specific meetings with the Product and Program managers, correlation of development schedules across hardware and software products, and field demonstrations of emerging equipment. Early involvement is key to driving design considerations that positively affect emerging products. Supplementing these formal events, OLCF meets weekly with their Cray Site Advocate, and Cray Hardware and Systems Analysts to ensure that there is frequent and consistent communication about known issues, bug tracking, and near-term product development.

OLCF maintains a robust vendor briefing schedule with other product manfacuturers as well, making sure that emerging products that are targeted to this program are well suited to the high performance, high capability, high capacity needs of the Center.

3. BUSINESS RESULTS

CHARGE QUESTION 3: ***Is the facility maximizing the use of its HPC systems and other resources consistent with its mission?***

OLCF RESPONSE: Users continue to make effective, maximum use of the resources available through the OLCF, carrying out production simulations that could not be done without leadership-class computing systems.

2011 Operational Assessment Guidance – Business Results

In this section, the Facility summarizes and reports its HPC and other resources usage:

- Resource Availability for appropriate computational and storage systems. The individual Facility and Program manager shall agree to specific metrics for resource availability as appropriate.
- Resource Utilization for appropriate computational and storage systems; and
- Capability Usage for appropriate HPC systems. The individual Facility and the Program manager shall agree to specific metrics for capability utilization as appropriate.

2011 Approved OLCF Metrics – Business Results

Business Metric 1: ***System Availability (includes XT4, XT5, HPSS, and Spider):***

- ***Scheduled availability: 95%***
- ***Overall availability 90%.***

(For a period of one year following a major system upgrade, the targeted Scheduled availability is 85% and Overall availability is 80%).

OLCF computational resources' scheduled availability (SA) and overall availability (OA) for CY 2010 and CY 2011 YTD are summarized in Tables 3.2 and 3.3 for the OLCF XT5, XT4, HPSS and Spider.

The scheduled availability (SA) metric was exceeded in CY 2010 for the OLCF XT5 (target 85%, achieved 94.1%) as well as for the XT4, HPSS, and Spider systems (target 95%). SA is projected to exceed the target metric in 2011.

The overall availability (OA) metric was exceeded in CY 2010 for the OLCF XT5 (target 80%, achieved 89.2%) as well as for the XT4, HPSS, and Spider systems (target 90%). OA is projected to exceed the target metric in 2011.

Business Metric 2: ***Resource Utilization: OLCF will report on INCITE allocations and usage.***

Total system utilization for the Cray XT5 for the period January 1, 2011-June 30, 2011 was 85.98%.

CY 2010 allocations: Total 1,268 million hours (950 million INCITE, 215 million ALCC, 103 million Director's Discretionary)

CY 2011 allocation through June 30,2011: Total 1,408 million hours (930 million INCITE, 368 ALCC, 110 million Director's Discretionary)

INCITE usage for CY 2010 was 1,070 million core-hours, 112.6% of the total allocation. INCITE usage in CY 2011 to date (6/30/2011) is 375 million core-hours, or 40.3% of the total allocation. For details about usage, Reference Section 3.2.

Business Metric 3: ***Capability Usage: For the calendar year, at least 40% of the consumed core hours will be from jobs requesting 20% or more the available cores.***

The OLCF XT5 exceeded the capability usage metric in CY 2010 (target 35%, achieved 39%) and is on track to exceed the capability usage metric in CY 2011 (target 40%, achieved 54% YTD; Reference Section 3.2).

Business Results Summary

Business results measure the performance of the OLCF against a series of operational parameters. The operational metrics most relevant to OLCF business results are resource availability and capability usage of the HPC resources.

The OLCF mission is to deliver leadership computing for science and engineering, focus on grand-challenge science and engineering applications, procure largest-scale computer systems (beyond vendor design point), and develop high-end operational and application software in support of the DOE science mission. To ensure that the facility is maximizing the use of its HPC systems and other resources, consistent with this mission, the OLCF closely monitors appropriate business and operational metrics and regularly measures and tunes the effects of operational policy through a series of technical and operations councils. These councils not only maximize efficiency and effectiveness, but also add another dimension to customer communications and support.

Cray XT Compute Partition Summary

The 2010 OA report described the upgrade of the existing Cray XT5 from AMD Opteron quad-core processors to AMD Opteron six-core processors, providing a 50% increase in the resources available for OLCF users (Table 3.1). Since this upgrade, the Cray XT5 hardware configuration is unchanged, with steady-state operation delivering well over 1 billion compute hours per year. The Cray XT5 configuration will remain unchanged until the first quarter of FY12, when another systemwide upgrade will provide 16-core AMD Opterons, an upgrade to 600TB of DDR3 memory, and the Gemini high-speed interconnect and introduce GPU accelerator technology.

Table 3.1 Cray XT Compute Partition Specifications, July 1, 2010–June 30, 2011

System	Type	CPU Type/Speed	Nodes	Memory/Node	Node Interconnect	Cores per Node	Total Cores	Aggregate Memory
Jaguar	Cray XT4	AMD Opteron 1354/2.1 GHz	7,832	8 GB	SeaStar2	4	31,328	62 TB
JaguarPF	Cray XT5	Opteron 2435/2.6 GHz	18,688	16 GB	SeaStar2	12	224,256	300 TB

Cray XT4 Decommissioning and the Role of the XT5 as a Leadership-class System

The Cray XT4, while an exceptionally productive system since its introduction as a 25TF XT3 in 2004, was scheduled to be retired before the end of FY11. The Cray XT5, last upgraded in the first quarter of FY10, was clearly the new ORNL leadership-computing platform with eight times as many cores as the XT4 and twelve-core nodes. The Cray XT4, physically limited to jobs below 31,000 cores reflected less of a leadership and more of a capacity role in FY2011.

The Cray XT4 was officially decommissioned at the end of February 2011. The timing of the decision protected operating dollars during a period of significant budget uncertainty. The impact of this decision to users was estimated at less than 5% of the total cycles to be delivered in the reporting period ending June 30, 2011.

The Cray XT5 is configured to support leadership computing in a single partition, allowing scheduling and execution of jobs of more than 224,000 cores. The operational focus is on delivering stable hardware and software and the tools that allow users to pursue grand-challenge science and engineering applications.

Delivering Production-Quality Computing Hours

In CY 2010, the OLCF projected that 1.55B compute hours would be delivered, distributed among 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) programs. The combination of XT4 and XT5 systems delivered more than 100M core hours above this projection, demonstrating the OLCF commitment to maximizing resource availability for users.

HPC Operations Delivering Results

Hardware failure rates are monitored closely. Cray maintains actual field measurements for failure rates of many system components and compares them frequently against the equipment manufacturer's failure rates and against the failure rates of the same parts in other systems. This ensures that discrepancies can be identified quickly and tracked to root cause.

During this reporting period, Cray and ORNL detected that failures of voltage regulator modules (VRMs) on the ORNL XT5 were statistically higher than at other XT5 sites. A VRM failure can impact a compute blade, take down the system interconnect fabric, and require a reboot to recover. The impact of these higher VRM failure rates can be observed in the metrics for mean time to failure (MTTF), scheduled availability (SA), and overall availability (OA).

Working with Cray, an engineering change related to the input voltage to the module was identified and implemented. This change is expected to increase the MTTF for the VRM and to positively impact the MTTF, SA, and overall availability for the system as a whole.

Governance Contributing to the Efficient and Effective Use of Resources

To ensure that operational metrics are met or exceeded and that resources are used efficiently and effectively, the OLCF regularly measures and tunes the effects of operational policy through a series of technical and operations councils. These councils not only maximize efficiency and effectiveness, they also contribute yet another facet to customer communications and support.

Resource Utilization Council

The Resource Utilization Council (RUC), which includes representation from across the facility, meets weekly, making decisions on things like DD awards (Section 4.4.3), and analyzing operations, including failure rates and resource utilization with a strong user focus to help shape OLCF policies and procedures. This has led to the following service improvements and resource innovations in the past year.

- To promote leadership usage of the OLCF systems, the RUC initiated a study of queuing on OLCF systems last fall. Empirical data in the form of queue simulations and examination of batch system logs were used to formulate a new queuing policy. Based on the results, the RUC suggested a combination policy that gives precedence to high-core-count jobs while lowering the priority of users who have more recently used the system to ensure that all projects get an equitable chance to use system allocations. The new queuing policy was implemented after the OLCF User Council reviewed it. Before implementation of the new queuing policy in November of 2010, the OLCF had experienced a decline in capability usage. Since the policy was implemented, leadership usage has exceeded the metric for 8 straight months.
- All INCITE projects are required to provide quarterly reports. These quarterly reports provide a snapshot of how the projects are progressing and an opportunity to assist or offer suggestions if projects encounter problems affecting the progress of their research. Regular reports from the projects are also very important to show the value of the INCITE program to its sponsors and the public. Because of the importance of quarterly reports, the RUC implemented penalties for late reports in 2011, which has resulted in higher compliance than previously experienced.
- Because the OLCF experienced enormous growth in files stored to HPSS again in 2010–2011, the RUC identified and notified the top 10% of HPSS users and asked for their cooperation in reducing their storage use where possible and appropriate. Within 1 week of notifying the users, HPSS storage declined by 1 PB, approximately 5% of the total data stored.
- To ensure users could access system information in the ways most convenient to them, the RUC requested that UAO consider the use of tweets as another way of notifying users when the state of a system changes. An OLCF twitter status has been established and is being tested before release to the users.

Software Council

Representatives from all OLCF groups serve on the Software Council (SWC). It grew from the desire to make the OLCF user experience as positive as possible by

- ensuring that software decisions are made in an efficient, effective, consistent manner;
- giving users a central place to go with software requests;
- ensuring that user requests are answered in an expeditious manner (1 week); and
- ensuring that new software approved for the system is promptly and efficiently loaded.

The SWC assesses user requests for new or updated versions of software to be installed on OLCF systems and ensures that all software, once loaded, is managed throughout its lifetime. Communication among SWC members is routinely carried out via e-mail, with formal council meetings once each quarter. In the past 12 months, nearly 30 software requests were fielded by the council. In addition to routine software upgrades, about half a dozen new applications were evaluated for potential value to Center users and installed on Center resources.

This activity has grown so much and is such an integral part of the success of the Center that in FY11 the OLCF created a position whose responsibilities will include managing, coordinating, building, installing, and maintaining the third party applications and libraries on the OLCF systems. This software specialist will also contribute to validation testing efforts and work with developers and other members of the Center when incompatibilities with their code bases and third-party software products are identified. In addition, this software specialist will provide documentation and troubleshoot issues with installed third party software.

User Council

The User Council is composed of a group of system users and especially focused on issues, concerns, and suggestions for facility operation and improvements. Members are selected annually at the User Meeting in May, with officers selected biennially.

Balint Joo of the Thomas Jefferson National Accelerator Facility is chair of the 2010–2011 User Council. For details about this past year's activities, Reference Section 2.4.2.

3.1 RESOURCE AVAILABILITY

The OLCF tracks a series of metrics that reflect the performance requirements of DOE and the user community. These metrics assist staff in monitoring system performance, tracking trends, and identifying and correcting problems at scale, all to ensure that OLCF systems meet or exceed DOE and user expectations.

3.1.1 Scheduled Availability

2011 Operational Guidance – Scheduled Availability

$$SA = \left(\frac{time\ in\ period - time\ unavailable\ due\ to\ outages\ in\ the\ period}{time\ in\ period - time\ unavailable\ due\ to\ scheduled\ outages\ in\ the\ period}\right) \times 100$$

Scheduled Availability (SA) measures the effect of *unscheduled* downtimes on system availability. For the SA metric, scheduled maintenance, dedicated testing, and other scheduled downtimes are not included in the calculation. The SA metric is to meet or exceed an 85% scheduled availability in the first year after initial installation or a major upgrade, and to meet or exceed a 95% scheduled availability for systems in operation more than 1 year after initial installation or a major upgrade. Reference Table 3.2.

Table 3.2 OLCF Computational Resources Scheduled Availability (SA) Summary 2010–2011

System	CY 2010		CY 2011 YTD (Jan 1-Jun 30 2011)		
	Target SA	**Achieved SA**	**Target SA**	**Achieved SA through June 30, 2011**	**Projected SA, CY 2011**
Cray XT5	85%	94.1%	95%	93.9%	>95%
Cray XT4	95%	97.1%	95%	97.6%[b]	97.6%[1]
HPSS[2]	95%	99.6%	95%	99.9%	>95%
Spider[2]	95%	99.8%	95%	98.5%	>95%
Spider2[2,3]	N/A	N/A	95%	99.9%	>95%
Spider3[2]	N/A	N/A	95%	99.9%	>95%

3.1.2 Overall Availability

2011 Operational Guidance – Overall Availability

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

Overall Availability (OA) measures the effect of both *scheduled and unscheduled* downtimes on system availability. The OA metric is to meet or exceed an 80% overall availability in the first year after initial installation or a major upgrade, and to meet or exceed a 90% overall availability for systems in operation more than 1 year after initial installation or a major upgrade. Reference Table 3.3.

[1] The Cray XT4 was decommissioned at the end of February 2011. Projected SA values for the XT4 reflect the actual data through the decommissioning date.
[2] A new metric to track HPSS and Spider availability was introduced in 2010.
[3] New filesystem added in 2010

Table 3.3 OLCF Computational Resources Overall Availability (OA) Summary 2010–2011

System	CY 2010		CY 2011[a]		
	Target OA	**Achieved OA**	**Target OA**	**Achieved OA through June 30, 2011**	**Projected OA, CY 2011**
Cray XT5	80%	89.2%	90%	88.7%	>90%
Cray XT4	90%	94.9%	90%	97.1%	97.1% [b]
HPSS[c]	90%	98.6%	90%	98.9%	>90%
Spider[c,e]	90%	99.0%	90%	96.5%	>90%
Spider2[d]	NA	NA	90%	99.1%	~99%
Spider3[d]	NA	NA	90%	99.2%	~99%

Independent Measurement of OLCF File and Archive Systems Availability

Beginning in 2010, the OLCF added tracking and reporting of the HPSS archive system and of the parallel file systems as independent metrics, separate from the compute systems. The associated metrics tracked are both scheduled and overall availability. The Spider file systems are in their second year of production, and are measured against the same second-year availability metrics as the compute systems. These correspond to approved metrics of a 95% scheduled availability and a 90% overall availability. Note that ORNL has chosen to retain the more stringent "second-year" designation for the file system metrics even though the original Spider file system is now maintained as three separate file systems.

2011 Scheduled and Overall Availability Assessment

The Cray XT5 is the only system that is not currently meeting the 2011 SA and OA metrics at the calendar-year mid-point. It will need to achieve an SA slightly greater than 96%, and an OA greater than 91.3% for the second half of the year to meet the full-year metric. However, with the ability to now significantly reduce unscheduled interrupts due to node VRM faults, described in detail below, ORNL expects that the year-end metric will reflect an SA that meets or exceeds the metric. The single-month snapshot of OA and SA for the Cray XT5 for July 2011, which is outside of the guidelines for this report, indicates that the system should exceed the metric for the second half of the year, and for the year overall.

Increasing System Availability – Resolving Critical Portals Issue and Reducing VRM Failure Rates

The SA and OA metrics are predicated on many factors, including the large physical scale of the system, the aggregate calculation of the failure rates of many disparate components, the architecture of the system and its resiliency to interrupt or failure due to a hardware component failure, and the resiliency to interrupt or failure due to a software failure.

In December 2010, ORNL and Cray resolved an open CRITICAL bug against the Portals low-level network programming interface. Resolution included a software patch to CLE 2.2 that was tested extensively at ORNL in Q1 FY11. This patch significantly reduced the number of instances where the Portals software failed to recover correctly from an HT_Lockup hardware event. The Portals patch, first incorporated in to the CLE 2.2 software stack, is now incorporated in to all CLE 3.x and 4.x releases. The distribution of HT_Lockup failures is shown in Figure 3.1. This failure rate and distribution is typical for a machine of this size. However, with the portals patch installed, the Cray XT5 can tolerate these HT hardware failures, riding through them without a system interrupt and the need to reboot.

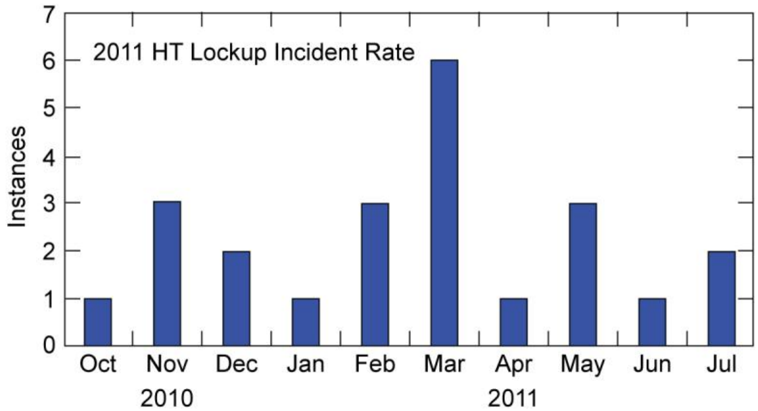

Figure 3.1 Cray XT5 HT_Lockup Incident Rate

As part of standard XT5 operations, ORNL and Cray continually assess the hardware component failure rates in the XT5 system against both the expected component failure rates using original equipment manufacturer and their own qualification data and against the failure rates of the same components at other Cray installations. During this reporting period, ORNL and Cray identified two component failure rates that were statistically anomalous. The first of these was associated with a higher incidence than anticipated of DIMM failures categorized as uncorrectable memory errors (UME). A UME will cause a job running on the associated compute node to fail. This error condition does not cause a system interrupt, and the affected node is removed from the available compute pool until the next scheduled maintenance period. To reduce the impact of UMEs, the onsite Cray hardware staff monitor correctable memory errors on DIMMs to identify potentially failing memory and use scheduled maintenance time to execute rigorous memory diagnostics to identify and drive out suspect parts.

The second anomalous condition was associated with high failure rates for voltage regulator modules (VRM). On the Cray compute blade, each VRM is a step-down DC to DC converter that provides the associated 6-core AMD Opteron (Istanbul) the appropriate supply voltage of +1.3V from the higher voltage (nominally +12V, with ±5% variance) supplied to the compute blade.

VRM failures are associated with compute nodes powering down, heartbeat faults and link-inactive failures. These affect the SeaStar interconnect fabric, and can produce a condition that causes an unscheduled system interrupt. Cray and ORNL investigated multiple engineering solutions to this event and have identified and implemented a solution related to a change to the VRM input voltage that is expected to significantly reduce the failure rate of the VRM. The result is expected to be an increased MTTI, increased MTTF, and better overall availability. The initial implementation of this engineering change was started in mid-June 2011. Since implementation, there have been only two VRM failures; one in the second half of June, and one in July. This represents a reduction, on average, from more than two unscheduled interrupts per week to less than one unscheduled interrupt due to this condition every three weeks. The continued assessment of this change over a longer period is expected to reveal dramatically

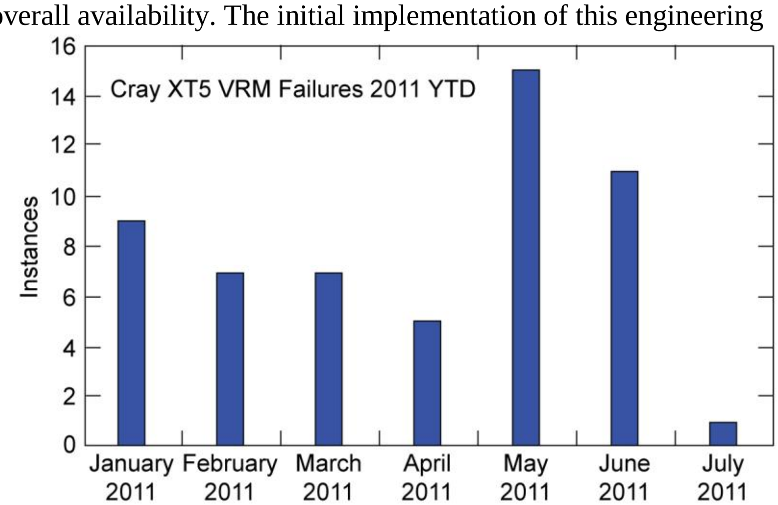

Figure 3.2 Eliminating VRM failures increases system stability.

better stability for the remaining life of the SeaStar-interconnected system. The change to the node VRM failure rates is shown in Figure 3.2.

In all such cases, ORNL works with Cray to identify the root cause for statistically significant deviations in failure rates and to identify and implement solutions to these conditions.

3.1.3 Mean Time to Interrupt

Mean Time to Interrupt (MTTI) measures the impact of both scheduled service interruptions (planned maintenance or dedicated testing) and unscheduled system interruptions from both internal and external sources.

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

Where time in period is start time – end time, start time = end of last outage prior to reporting period, and end time = start of first outage after reporting period (if available) or start of the last outage in the reporting period.

The Mean Time to Interrupt Summary is shown in Table 3.4.

Table 3.4 OLCF Mean Time to Interrupt (MTTI) Summary 2010–2011

System	MTTI, CY 2010 (hours)	MTTI, CY 2011 YTD (hours)
Cray XT5	45.2	35.7
Cray XT4	95.8	78.7
HPSS	291.8	258.6
Spider [a]	481.6	322.5
Spider2 [a]	NA	538.1
Spider3 [a]	NA	538.3

[a] Due to the extremely long uptime of the Spider files systems, the formula for MTTI produces artificially skewed results using the period as defined in the formulas. Values presented here for Spider, Spider2, and Spider3 have been determined based on calendar year periods (January 1 through December 31, 2010 and January 1 through June 30, 2011).

3.1.4 Mean Time to Failure

Mean Time to Failure (MTTF) measures the time to a system interrupt associated with an *unscheduled* event from either an internal or external source.

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

Where time in period is start time – end time, start time = end of last outage prior to reporting period, and end time = start of first outage after reporting period (if available) or start of the last outage in the reporting period.

The Mean Time to Failure Summary is shown in Table 3.5.

Table 3.5 OLCF Mean Time to Failure (MTTF) Summary 2010–2011[a]

System	MTTF, CY 2010 (hours)	MTTF, CY 2011 YTD (hours)
Cray XT5	59.5	45.7
Cray XT4	134.0	87.8
HPSS	501.3	610.6
Spider [a]	623.8	856.1
Spider2 [a]	NA	867.6
Spider3 [a]	NA	868.0
[a] Due to the extremely long uptime of the Spider files systems, the formula for MTTF produces artificially skewed results using the period as defined in the formulas. Values presented here for Spider, Spider2, and Spider3 have been determined based on calendar year periods (January 1 through December 31, 2010 and January 1 through June 30, 2011).		

[a]Overall availability by calendar year (CY). CY 2011 year-to-date (YTD) data in Section 3 were generated from January 1, 2011, through June 30, 2011, unless otherwise noted.
[b]Cray XT4 was decommissioned at the end of February 2011. Projected OA values for the XT4 reflect the actual data through the decommissioning date.
[c]A new metric to track HPSS and Spider availability was introduced in 2010.
[d]Two new file systems were added in CY 2010.
[e]Dedicated Lustre testing was conducted using Spider leaving Spider2 and Spider3 (default scratch) available to users.

Assessing the Cray XT5 MTTI and MTTF

MTTI and MTTF provide a mechanism for measuring system stability. The Cray XT4, decommissioned at the end of February 2011, continued to demonstrate stable MTTI and MTTF through its end-of-life.

The Cray XT5 MTTI reflects higher than expected DIMM failure rates (UMEs). UMEs will impact the job associated with the node, but will not typically affect the remainder of the system. Cray Hardware drives out marginal DIMMs with additional memory diagnostic testing. These tests are executed routinely during scheduled PMs. DIMMs that do not pass the more rigorous testing are returned for additional testing by Cray-Chippewa Falls and the original equipment manufacturer.

The Cray XT5 MTTF reflects both the CRITICAL Portals bug that impacted the system through Q1 FY11, and the higher than expected VRM failure rates that were resolved in June 2011. MTTF is expected to be substantially better in the two remaining quarters of CY11, and to have a corresponding positive impact on the full CY results.

3.2 RESOURCE UTILIZATION

2011 Operational Assessment Guidance

The Facility reports Total System Utilization for each HPC computational system as agreed upon with the Program Manager

For the period January 1 – June 30, 2011, 744,861,807 core-hours were delivered from a scheduled maximum of 866,291,158 core-hours. This resulted in total system utilization for the Cray XT5 of 85.98%.

INCITE Utilization

Allocations to Center systems are made via three programs: INCITE, ALCC, and DD. The majority of the hours are awarded via INCITE and are granted by calendar year.

- CY 2010 allocations: Total 1,268 million hours (950 million INCITE, 215 million ALCC, 103 million DD)
- CY 2011 allocation to date: Total 1,408 million hours (930 million INCITE, 368 ALCC, 110 million DD)

The INCITE allocation is at least 60% of the total allocated hours on the OLCF systems.

INCITE usage for CY 2010 was 1,070 million core-hours, 112.6% of the total allocation. INCITE usage in CY 2011 to date (6/30/2011) is 375 million core-hours, or 40.3% of the total allocation. A logarithmic trend line has been applied to the 2011 weekly chart data to indicate the stabilization of the weekly usage. INCITE usage in the first part of the Calendar Year is typically lower due to the on-ramp of projects and consumption. Utilization in the remainder of the year is traditionally higher and more stable. Reference Figure 3.3 for 2011 INCITE Usage by Week.

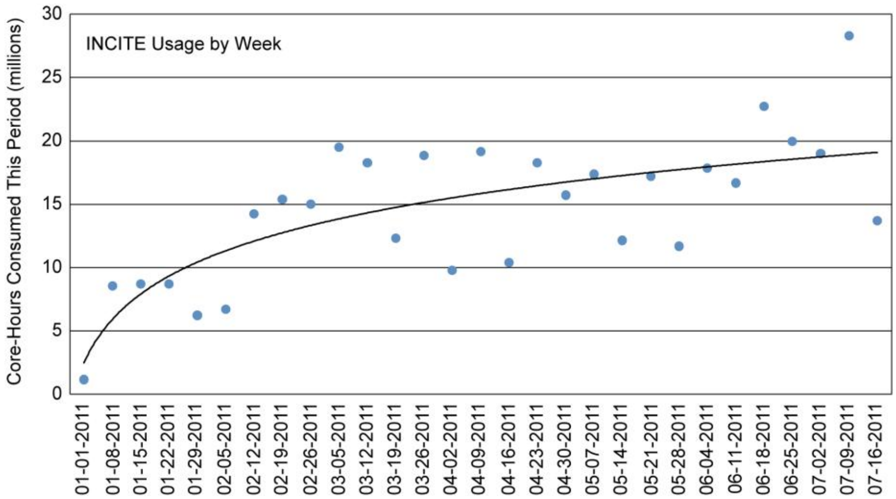

Figure 3.3 2011 INCITE Usage by Week

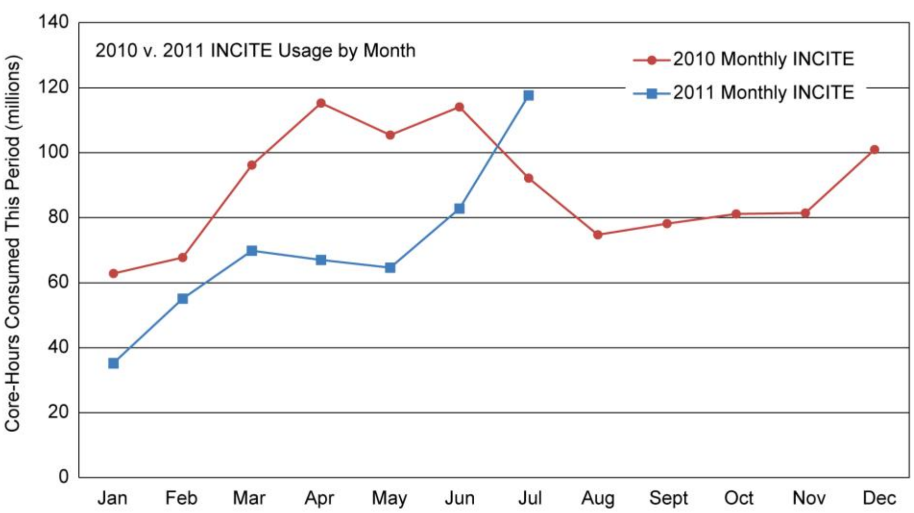

Figure 3.4 Comparing 2010 and 2011 INCITE Usage

A comparison of the 2010 INCITE usage on Jaguar against the 2011 INCITE usage YTD is shown in Figure 3.4. Both 2010 and 2011 figures reflect the typically slower initial consumption rate that reaches a more predictable state in mid-year. Out-year consumption for 2010 remained above 80 million core-hours per month in the second half of the year, a reflection of multiple factors including total system demand among INCITE, ALCC, and DD programs, scheduling policy that favored larger, Leadership Class computing, and system availability. For the first half of 2011, there is one additional factor to be considered, as node VRM failures contributed to a lower OA than anticipated. This situation was corrected in mid-June 2011 as described earlier.

3.3 CAPABILITY UTILIZATION

2011 Operational Assessment Guidance – Capability Utilization

An individual Facility shall maintain an agreement with its DOE Program Manager on the definition of capability utilization, and the HPC systems to which the metric applies (called capability systems). The Facility shall describe the agreed metric, the operational measures that are taken to support the metric, and the results, by capability system.

Leadership usage on the Cray XT5 is defined by the number of cores used by a particular job. For both 2010 and 2011, a leadership-class job must use no less than 20% of the available cores (Figure 3.5). In the current configuration this equates to about 44,800 cores.

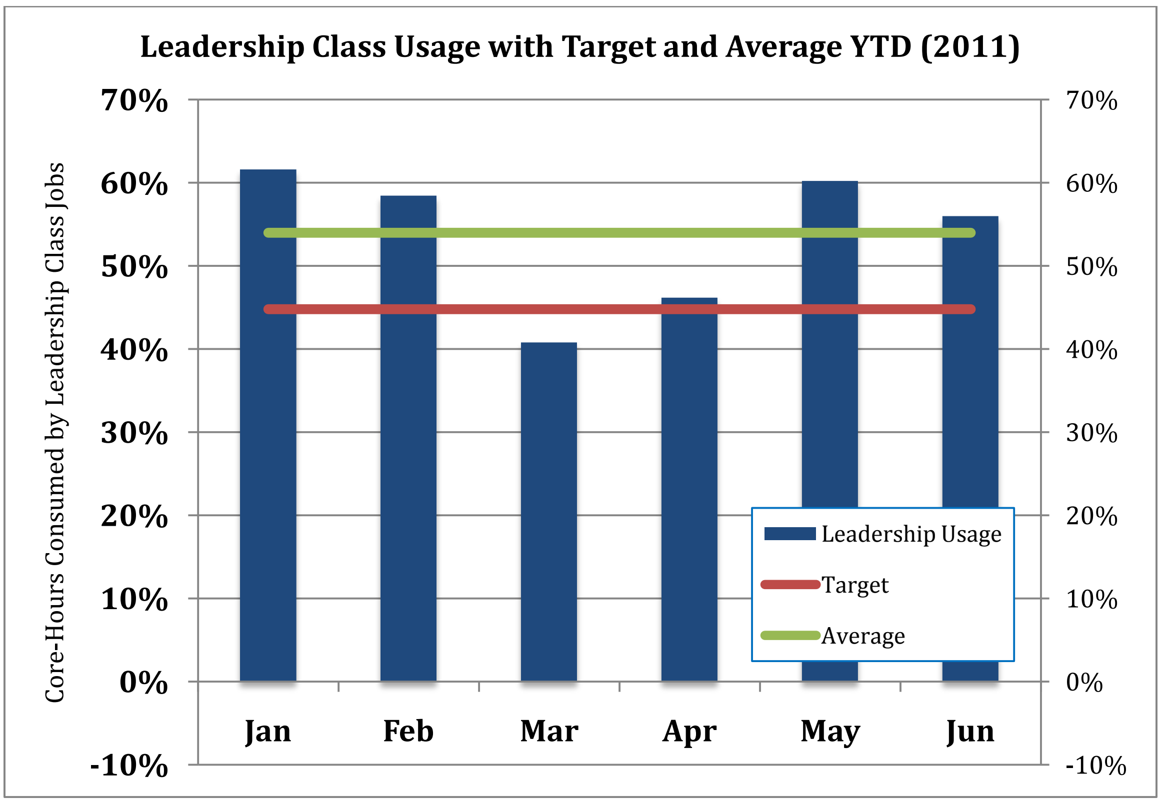

Figure 3.5 Effective Scheduling Policy Enables Leadership-class Usage.

The capability metric is defined by the number of CPU hours that are delivered by leadership-class jobs. For the initial year of production (2010), the Cray XT5 metric stipulated that no less than 35% of the delivered CPU hours would reflect leadership-class jobs. For the second year of production (2011), the Cray XT5 metric stipulates that no less than 40% of the delivered CPU hours reflect leadership-class jobs.

The OLCF continues to meet – and exceed – expectations for capability usage of its HPC resources (Table 3.6). Keys to the growth of leadership usage include the liaison role provided by the SciComp

Group members, who work hand-in-hand with users to port, tune, and scale code, and ORNL support of the Joule metrics, where staff actively engage with code developers to promote application performance.

Table 3.6 OLCF Leadership Usage on JaguarPF

Leadership Usage	CY 2010 Target (%)	Achieved (%)	CY 2011 Target (%)	CY 2011 YTD (%)
≥20% of cores	35.0	39%	40%	54.0%

3.4 INFRASTRUCTURE

3.4.1 Networking

ORNL/OLCF is participating in the Energy Sciences Network (ESnet) Advanced Networking Initiative (ANI) as one of the very large network endpoints. The ANI will provide a 100 Gb/s prototype network, with endpoints at ORNL, NERSC, ANL, and the metropolitan New York area. It will also provide a network test bed facility for users and industry. This ANI network is funded by the American Recovery and Reinvestment Act (ARRA) The goal of the prototype network is to accelerate deployment of 100 Gb/s technologies and build a persistent infrastructure that will transition to the production ESnet network as early as 2012. This is considered a key step toward the DOE vision of a 1 TB network linking DOE supercomputing centers and experimental facilities.

The ANI transport network has an initial delivery and implementation schedule that will have the primary sites up and connected before the end of the calendar year. In the interim, existing ESnet Science Data Network (SDN) circuits are being used for preliminary testing. SDN enables dynamic provisioning of dedicated circuits between connected research facilities, specifying the bandwidth and the amount of time needed for the dedicated circuits. The OLCF connects to the SDN in Nashville at 10 Gb/s, using ORNL dark fiber and optical transport between Oak Ridge and Nashville. This 10 Gb/s circuit is being used for disk-to-disk data transfer testing between ORNL and the National Energy Research Scientific Computing Center (NERSC). This testing will transition to the 100 Gb/s network when it becomes available later this calendar year.

Perimeter and Local Area Network Upgrades

This past year, the OLCF deployed stateful 10 Gb firewalls and is working on migrating networks over to them. These firewalls are deployed in a high availability (HA) configuration, ensuring greater reliability of the OLCF network.

A new core router, which will form the core of the OLCF network, has also been deployed. This router gives the OLCF a path forward to 40 and 100 Gb network connections and, potentially, terabit connections in the future. This upgrade also enables the OLCF to retire aging hardware, saving funds on maintenance and reducing power usage.

The OLCF internal network is being reconfigured to use more low latency, high speed, nonblocking switches. This architecture was deployed for infrastructure services last year and is being further deployed for HPSS this year. This change will facilitate a much more scalable upgrade path for the HPSS network.

3.4.2 Storage

The OLCF is actively involved in several storage-related pursuits including media refresh, data retention policies, and filesystem/archive performance. As storage, network, and computing technologies continue to change, the OLCF is evolving to take advantage of new equipment that is both more capable and more cost-effective.

Storage requirements for both the centerwide file system (Spider) and the high-performance tape archive (HPSS) continue to grow at high rates. In September 2010, two new Lustre file systems were added to the existing centerwide file system. These two file systems increased the amount of available disk space from 5 to 10 PB and will help improve overall availability as scheduled maintenance can be performed on each file system individually. The addition of these file systems provides a 300% increase in aggregate Metadata performance and a 200% increase in aggregate bandwidth. Additional monitoring improvements for the health and performance of the file systems have also been made.

In August 2010, a major software upgrade on the HPSS archive was completed, and staff members began evaluating the next generation of tape hardware. In April 2011, twenty STK/Oracle T10KC tape drives were integrated into the HPSS production environment. This additional hardware is proving to be very valuable to the data archive in two distinct ways. The new drives provide both a 2× read/write performance improvement over the previous model hardware and a 5× increase in the amount of data that can be stored on an individual tape cartridge. Along with improved read/write times to/from these new drives, the OLCF now benefits from being able to store 5 TB on each individual tape cartridge–effectively extending the useful life of the existing tape libraries. This has allowed the OLCF to postpone its next library purchase until the first half of FY12.

The HPSS archive currently houses more than 18 PB of data, up from 12 PB a year ago. The present ingestion rate is between 20–40 TB every day, with occasional periods of high usage approaching 100 TB in a single day. The OLCF has two Sun/STK 9310 automated cartridge systems (ACS) and four Sun/Oracle SL8500 Modular Library Systems. The 9310s have reached the manufacturer end-of-life (EOL) and are being prepared for retirement. Each SL8500 holds up to 10,000 cartridges, and there are plans to add a fifth SL8500 tape library in 2012, bringing the total storage capacity up to 50,000 cartridges. The current SL8500 libraries house a total of 16 T10K-A tape drives (500 GB cartridges, uncompressed), 60 T10K-B tape drives (1 TB cartridges, uncompressed), and 20 T10K-C tape drives (5 TB cartridges, uncompressed). The tape drives can achieve throughput rates of 120–160 MB/s for the T10KA/Bs and up to 240 MB/s for the T10K-Cs.

HPSS Version 7.3 in OLCF Production

The OLCF completed its upgrade to HPSS version 7.3.2 in August of 2010. Implementation of this release has resulted in performance improvements in the following areas.

- Handling small files. For most systems it is easier and more efficient to transfer and store big files; these modifications made improvements in this area for owners of smaller files. This has been a big gain for the OLCF because of the great number of small files stored by our users.
- Tape aggregation. The system is now able to aggregate hundreds of small files to save time when writing to tape. This has been a tremendous gain for the OLCF.
- Multiple streams or queues of what HPSS refers to as “class-of-service changes.” This has enabled the system to process multiple files concurrently and, hence, much faster, another huge time saver for the OLCF and its users.
- Dynamic drive configuration. Configurations for tape and disk devices may now be added and deleted without taking a system down, giving the OLCF tremendously increased flexibility in fielding new equipment, retiring old equipment, and responding to drive failures without affecting user access.

3.5 FOCUSING ON ENERGY SAVINGS

The Computational Sciences Building (CSB) currently houses three very large computer systems. The largest is DOE's Jaguar. The University of Tennessee's Kraken, the world's fastest academic supercomputer, and the National Oceanic and Atmospheric Administration's Gaea, the world's largest dedicated resource for climate prediction, are also installed on the same raised floor. In total, these systems can sustain as much as 2.8 PF. These systems also consume substantial amounts of energy with equally large demands for a robust cooling and support infrastructure.

The CSB adheres to rigorous engineering management practices and is LEED (Leadership in Energy and Environmental Design) certified, the only rating available at the time of construction. As a result of these careful engineering practices, the CSB produces a power usage effectiveness (PUE) of less than 1.25 compared to an average of about 1.8 among other large-scale data centers. In practical terms, this means that within the CSB facility each 1 MW used to power the machines, requires just 0.25 MW for supporting functions, including the removal of waste heat, lighting, and other ancillary facility services. ORNL has a second computing center that was built shortly after CSB. This facility is LEED-Gold certified.

Since completion of the facility in 2004, the OLCF has aggressively pursued methods for reducing its resource footprint even more, harnessing energy savings wherever possible.

Mechanical system improvements continue to yield good savings. After completing a number of changes in 2010, including the installation of high volume pumps in the Central Energy Plant and variable frequency drives (VFDs) in the computer room air conditioning units (CRUs), ORNL targeted a number of smaller improvements that will cumulatively improve the capability of the chilled-water delivery system. The most substantial change was the installation of a centralized set of humidity sensors and reconfiguration of the CRUs to use this single input. This reduced the tendency of units in separate areas of the room to independently operate in conflict with other units. This single change reduced energy consumption and stabilized the relative humidity, dew point, and temperature in the room. A number of other changes were also made to the CRUs, increasing their efficiency, including installation of flow limiting valves, calibrating CRU sensors, installing shut off valves for inactive heat exchangers, optimizing humidification controls, and enabling night setback for variable air volume supply air.

Within the equipment in the computer room, raised floor openings were sealed, and air flow management was improved through the use of blanking panels and other devices, improving the air flow from the forced-air distribution system under the floor to the inlet/supply side of the air-cooled systems. Another example of the air flow management, a simple metal "top-hat" on the 30-ton CRUs in the computing facility is also being evaluated, with significant results to-date.

The Effects of CRU Top Hats on Air Flow

The ORNL Computer Facilities Manager and Facilities & Operations continue to evaluate various methods for improving the airflow within the data center, especially in high-density areas, and in constrained-supply areas. The target goals include increasing the capacity or effectiveness of an air handler, providing greater control over the air-distribution process, and providing more optimal inlet air temperatures to high-density air-cooled equipment.

In July 2011, ORNL installed air handler top hats on two 30-ton units. These top hats are simple ducting extensions that pull return air from a higher stratification in the data center. With the top hats installed, ORNL measured an increased return air temperature of 6 degrees Fahrenheit. According to the ASHRAE psychometric chart for mechanical cooling performance, a rise from 70°F to 75°F at 50% Relative

Humidity is equivalent to a 45% increase in cooling capacity at identical motor kW. Given the relatively low material cost for the top hats, and the high performance increase, ORNL is extending the installation of these top hats to the remaining air handlers in the Computational Sciences Building.

The results of this experiment are shown in Figure 3.6. Two CRUs, labeled CRU 39 and CRU 40 were sampled before and after top hat installation. These two units reside in a very dense air-cooled equipment area that has traditionally demonstrated mechanical challenges with both control of inlet temperatures, and control of exhaust heat. The summary of the impact of the top hats on the CRU on the return-air temperatures is shown in Table 3.7.

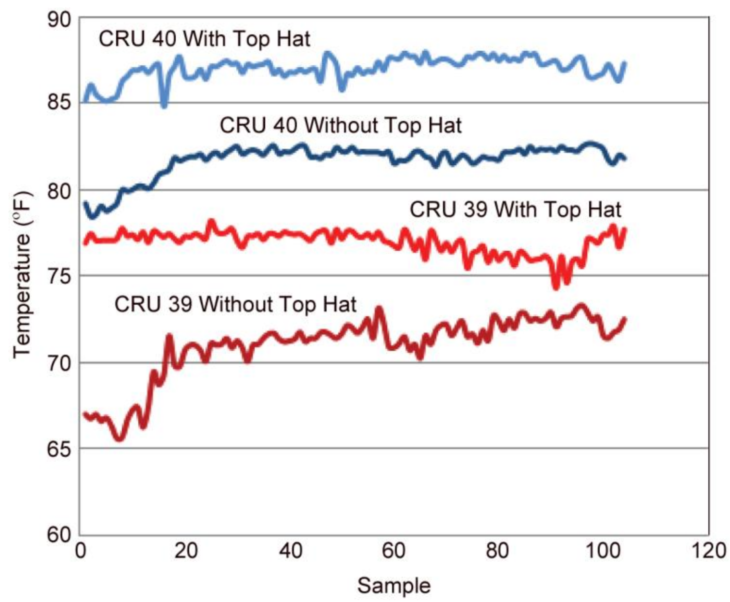

Figure 3.6 The Effect of Top Hats on CRU Efficiency

A number of activities continue, including studies on the effectiveness of hot/cold air separation techniques; use of water-side economizers; addition of VFDs to Central Energy Plant chillers and chilled-water pumps; cool-roof technologies; new controls for chilled-water delivery that optimize cooling load, environmental conditions, and available equipment; increasing the delivered chilled-water temperature; chilled-water storage; and load shedding.

Table 3.7 The Positive Impact on CRU Return-air Temperatures with Top Hats

	CRU 39			CRU 40		
Degrees Fahrenheit	71.0	76.9	6.0	81.7	87.0	5.3
Configuration	Without top hat	With top hat	Temp. increase (measured, average)	Without top hat	With top hat	Temp. increase (measured, average)

4. STRATEGIC RESULTS

CHARGE QUESTION 4: ***Is the facility enabling scientific achievements consistent with the Department of Energy strategic goals 3.1 and/or 3.2[1]?***

OLCF RESPONSE: The Center continues to enable high-impact science results through access to the leadership-class systems and support resources. The allocation mechanisms are robust and effective.

2011 Operational Assessment Guidance – Strategic Results

In this section the Facility reports:

- Science Output;
- Scientific Accomplishments; and
- Allocation of Facility Director's Reserve Computer Time (HPC only).

2011 Approved OLCF Metrics – Strategic Results

Strategic Metric 1: ***The OLCF will report numbers of publications resulting from work done in whole or part on the OLCF systems.***

636 publications in 2010 and 181 publications in 2011 YTD have been the result of work carried out by users of OLCF resources

Strategic Metric 2: ***The OLCF will provide a written description of major accomplishments from the users over the previous year.***

Several representative highlights are provided below. Additional significant accomplishments are also available in *INCITE in Review*[2]

Strategic Metric 3: ***The OLCF will report on how the Facility Director's Discretionary time was allocated.***

Section 4.4.3 provides details about the OLCF strategy for allocation of Director's Discretionary (DD) time (Reference Appendix A for a list of 2010-2011 DD projects). The DD projects cover a broad range of science domains and organizational affiliation types (university, government, private). The Industrial Partnerships Program projects, a subdomain within DD projects, are also listed.

The 2006 DOE Strategic Plan focused on themes of "Scientific Breakthroughs" and "Foundations of Science" aimed at strengthening U.S. scientific discovery and economic competitiveness and improving

---

[1] These goals are from the 2006 DOE Strategic Plan. Strategic Goal 3.1, Scientific Breakthroughs: Achieve the major scientific discoveries that will drive U.S. competitiveness, inspire America, and revolutionize approaches to the Nation's energy, national security, and environmental quality challenges Strategic Goal 3.2, Foundations of Science: Deliver the scientific facilities, train the next generation of scientists and engineers, and provide the laboratory capabilities and infrastructure required for U.S. scientific primacy. DOE's 2006 Strategic Plan, including both Strategic Goal 3.1 and Strategic Goal 3.2, is available at http://energy.gov/sites/prod/files/edg/media/2006StrategicPlanSection7.pdf. The DOE 2011 Stragic Plan is available at http://science.energy.gov/~/media/bes/pdf/DOE_Strategic_Plan_2011.pdf.

[2] http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/INCITE_IR.pdf

quality of life through innovations in science and technology. In the 2011 DOE Strategic Plan, the Department target is to continue to feed technology development through scientific discovery and "the Department will strive to maintain leadership in fields where this feedback is particularly strong, including…high-performance computing." The critical nature of simulations are highlighted in the thematic science areas in the Strategic Plan, and the targeted outcome for leading computational sciences and high-performance computing is to "continue to develop and deploy high-performance computing hardware and software systems through exascale platforms." The OLCF continues to lead the way in identifying and pursuing the requirements for next-generation computing.

In the 2010 OA report, 2009 was labeled the dawn of the petascale era. Now, just one short year later, the catch phrase is "general purpose GPU" (GPGPU) or the equally ubiquitous "CPU-GPU," thus proving once again that change is the only constant—even more so in the world of HPC than elsewhere. But there is a tendency to get caught up in the hype over the technology. As Axel Kohlmeyer, Associate Director of the Institute for Computational and Molecular Science (ICMS) at Temple University in Philadelphia has said, "it is the people who make the difference, the ingenuity with which we use technology that moves us forward, not just . . . more technology. After all it doesn't help to get an answer 100 times faster if we don't ask the right questions." This is something, indeed, that we have found to be true again and again. It is our people who are the most valuable resource. To meet the promise of GPGPU computing and reach exascale will require the combined talents and expertise of software developers, computer scientists, mathematicians, and others at all of the DOE HPCCs. Over the following pages we will describe and, in some measure, quantify how the OLCF and its staff are meeting that challenge and the challenges posed by the DOE strategic goals.

4.1 SCIENCE OUTPUT

2011 Operational Assessment Guidance – Science Output

The Facility tracks and reports the number of refereed publications written annually based on using (at least in part) the Facility's resources. This number may include publications in press or accepted, but not submitted or in preparation. This is a reported number, not a metric. In addition, the Facility may report other publications where appropriate.

The OLCF currently follows the recommendation in the 2007 report[1] of the ASCAC Petascale Metrics Panel to report and track user products including, for example, publications, project milestones (requested quarterly; also examined in the INCITE renewal process), and code improvement (Joule metric). Publications are listed in Table 4.1. The 2011 YTD publications are those collected from two quarters of reports from users. At the end of the year, a library search will be carried out to identify additional publications based on work using OLCF resources. The facility also collects quarterly reports from users, in which they are asked to provide updates on accomplishments and other activities, such as presentations given describing results of work under the allocation. In CY 2011 YTD, authors reported 49 presentations.

---

[1] Panel recommendations can be found in the full report of the committee, *Advanced Scientific Computing Advisory Committee Petascale Metrics Report*, 28 February 2007, available at http://science.energy.gov/~/media/ascr/ascac/pdf/reports/Petascale_metrics_report.pdf.

Table 4.1 Publications by Calendar Year

	2010	2011 YTD
Number of refereed publications based on the use (at least in part) of OLCF resources	636	181

4.2 SCIENTIFIC ACCOMPLISHMENTS

2011 Operational Assessment Guidance – Scientific Accomplishments

The Facility highlights a modest number of significant scientific accomplishments of its users, including descriptions for each project's objective, the implications of the results achieved, the accomplishment itself, and the facility's actions or contributions that led to the accomplishment.

In the last 20 years, we've seen shifts in architectures away from single core to multicore, and we now seem poised on the verge of another shift, to GPU computing. Because nothing, especially in HPC, is as simple or straightforward as it seems, as with past shifts, this one will require the collaboration of disciplinary scientists, applied mathematicians, and computer scientists. The OLCF has always approached the delivery of science on its computational resources as a collaborative enterprise. Computational scientists and other experts at the OLCF have engaged researchers worldwide to address the leading challenges facing the nation, and this year, as in the past, the scientific results stemming from this collaborative effort show that the OLCF strategy is continuing to pay dividends. We are confronting and answering big science questions and grand challenges—in energy, climate, materials science, physics, chemistry, and environmental science—as indicated in the abstracts and stories on the following pages and in Section 4.3.

Discovery Made Using ORNL Computers Boosts Supercapacitor Energy Storage

PI: Robert Harrison, ORNL
Time Awarded: 75,000,000 hours, 2010 INCITE; 75,000,000 hours 2011 INCITE

Drexel University's Yury Gogotsi and colleagues recently needed an atom's-eye view of a promising supercapacitor material to sort out experimental results that were exciting but appeared illogical. That view was provided by a research team led by Oak Ridge National Laboratory (ORNL) computational chemists Bobby Sumpter and Jingsong Huang and computational physicist Vincent Meunier.

Gogotsi's team discovered you can increase the energy stored in a carbon supercapacitor dramatically by shrinking pores in the material to a seemingly impossible size—seemingly impossible because the pores were smaller than the solvent-covered electric charge-carriers that were supposed to fit within them (Figure 4.1). The team published its findings in the journal *Science*. "We thought this was a perfect case for computational modeling because we could certainly simulate nanometer-sized pores," Sumpter said. "We had electronic-structure capabilities that could treat it well, so it was a very good problem for us to explore."

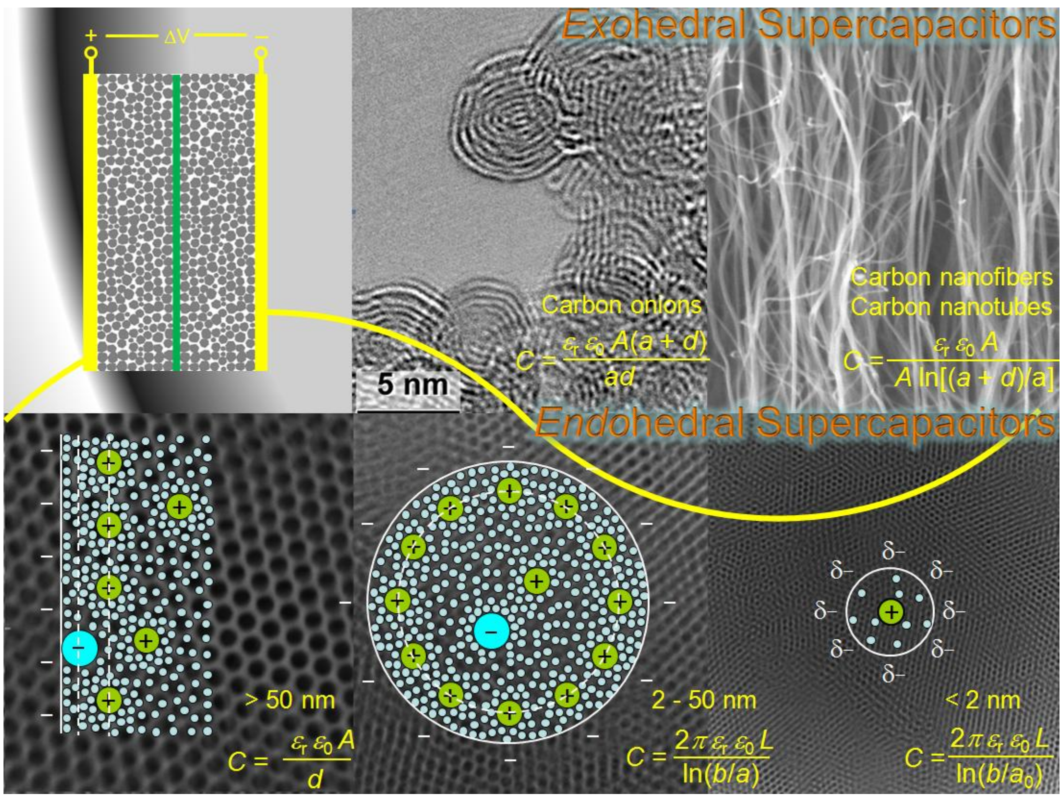

Figure 4.1. Computational modeling of carbon supercapacitors with surface curvature effects entertained leading to post-Helmholtz models for exohedral (top row) and endohedral (bottom) supercapacitors based on various high surface area carbon materials. (Image courtesy of Jingsong Huang, ORNL.)

Using ORNL supercomputers, Sumpter and his team were able to take a nanoscale look at the interaction between ion and carbon surface. A computational technique known as density functional theory allowed them to show that the phenomenon observed by Gogotsi was far from impossible. In fact, they found that the ion fairly easily pops out of its solvation shell and fits into the nanoscale pore.

Using these and other insights gained through supercomputer simulation, the ORNL team partnered with colleagues at Rice University to develop a working supercapacitor that uses atom-thick sheets of carbon materials.

"It uses graphene on a substrate and a polymer-gel electrolyte," Sumpter explained, "so that you produce a device that is fully transparent and flexible. You can wrap it around your finger, but it's still an energy storage device. So we've gone all the way from modeling electrons to making a functional device that you can hold in your hand."

BMI Uses Jaguar to Overhaul Long-Haul Trucks

PI: Mike Henderson, BMI
Time Awarded: 2,000,000 hours, Director's Discretionary

Those big rigs barreling down America's highways day and night are essential to the country's economy. They carry 75 percent of all US freight and supply 80 percent of its communities with 100 percent of their consumables. But there is a price to pay. These long-haul trucks average 6 miles per gallon or less and annually dump some 423 million pounds of CO_2 into the environment. BMI launched its SmartTruck program on a modest high-performance computing (HPC) cluster to tackle the design of new, add-on parts for long-haul 18 wheelers. "We initially ran our simulations on an HPC cluster with 96 processors," recalls BMI founder and CEO Mike Henderson. "We were unable to handle really complex models on the smaller cluster. The solutions lacked accuracy. We could explore possibilities but not run the detailed simulations needed to verify that the designs were meeting our fuel efficiency goals."

Figure 4.2. Trailers equipped with BMI Corp. SmartTruck UnderTray components can achieve a 7-12% percent improvement in fuel mileage. Representatives were on hand at ORNL on March 1, 2011 to display the components.

To beef up its computing power, BMI applied for and received a grant through the ORNL Industrial HPC Partnerships Program for time on Jaguar. Its engineers are now creating the most complex truck and trailer model ever simulated using NASA's Fully Unstructured Navier Stokes (FUN3D) application for computational fluid dynamics analysis. The team models half the tractor and trailer for simulation and analysis purposes, using 107 million grid cells in the process. To study yaw—what happens when the vehicle swerves—they mirror the grid and double it, using 215 million grid cells to accurately model the entire vehicle. BMI's ultimate goal is to design a sleek, aerodynamic truck with a lower drag coefficient than that of a low-drag car and anticipated fuel efficiencies running as high as 50 percent better than today. If all the country's 1.3 million long-haul trucks operated with the same low drag as a well-designed passenger car, the United States could annually save $5 billion in fuel costs and reduce CO_2 by 16.4 million tons (Figure 4.2).

Blood Simulation on Jaguar Takes 2010 Gordon Bell Prize

A team from Georgia Tech, New York University, and ORNL took this year's Gordon Bell Prize at SC10 by pushing ORNL's Jaguar supercomputer to 700 trillion calculations per second (700 teraflops) with a groundbreaking simulation of blood flow. The team wins a $10,000 prize provided by HPC pioneer Bell as well as the distinction of having the world's leading scientific computing application. Another team using Jaguar took an honorable mention in the competition for developing an innovative framework that calculates critical nanoscale properties of materials. The winning team used 196,000 of Jaguar's 224,000 processor cores to simulate 260 million red blood cells and their interaction with plasma in the circulatory system.

Lawrence Berkeley National Laboratory's Horst Simon, in announcing the winners, noted that the team achieved a 10,000-fold improvement over previous simulations of its type. "This team from Georgia Tech, NYU, and Oak Ridge National Lab received the award for obtaining four orders of magnitude improvement over previous work and achieved an impressive more than 700 teraflops on 200,00 cores of the Jaguar system," Simon said. "It's a very significant accomplishment." Simon noted also that the team simulated realistic, "deformable" blood cells that change shape rather than simpler, but less realistic, spherical red blood cells, calling the approach a "very challenging multiscale, multiphysics problem." The winning team included Abtin Rahimian, Ilya Lashuk, Aparna Chandramowlishwaran, Dhairya Malhotra, Logan Moon, Aashay Shringarpure, Richard Vuduc, and George Biros of Georgia Tech, Shravan Veerapaneni and Denis Zorin of NYU, and Rahul Sampath and Jeffrey Vetter of ORNL.

2010 Gordon Bell award winning team at SC10 in New Orleans, Louisiana.

An honorable mention in the Gordon Bell competition went to Anton Kozhevnikov and Thomas Schulthess of ETH Zurich, and Adolfo G. Eguiluz of the University of Tennessee, Knoxville, for reaching 1.3 thousand trillion calculations a second, or 1.3 petaflops, and scaling to the full Jaguar system in a method that solves the Schrödinger equation from first.

Building Gasifiers via Simulation

PI: Madhava Syamlal, National Energy Technology Laboratory
Time Awarded: 6,000,000 hours, 2010 INCITE

A team of scientists from the National Energy Technology Laboratory (NETL) used OLCF's Jaguar to conduct high-reliability simulations of a coal gasifier in an attempt to make the potential energy alternative more efficient and reliable. The team concluded that for engineering design of coal gasifiers, the overall resolution required in a simulation was 10 million to 20 million grid points. In 2010 the researchers presented their results at the NETL Multiphase Flow Science Workshop and published the findings in the journal Industrial & Engineering Chemistry Research. Determining the resolution requirements for simulations of coal gasifiers and their components (e.g., inlet jets) can reduce the cost and time required to develop near-zero-emissions power plants. These future plants may not only emit less carbon per unit of energy produced but also sequester carbon dioxide using water-gas shift reactions, which makes them amenable to a combined cycle where the waste heat generated during energy production is used to enhance the efficiency of the process (Figure 4.3).

Gasification is the process through which carbonaceous material such as coal, petroleum, or biomass is converted into carbon monoxide and hydrogen by reaction of the raw material with controlled amount of oxygen or steam at high temperatures. The resulting gas mixture is called syngas, which is a fuel itself.

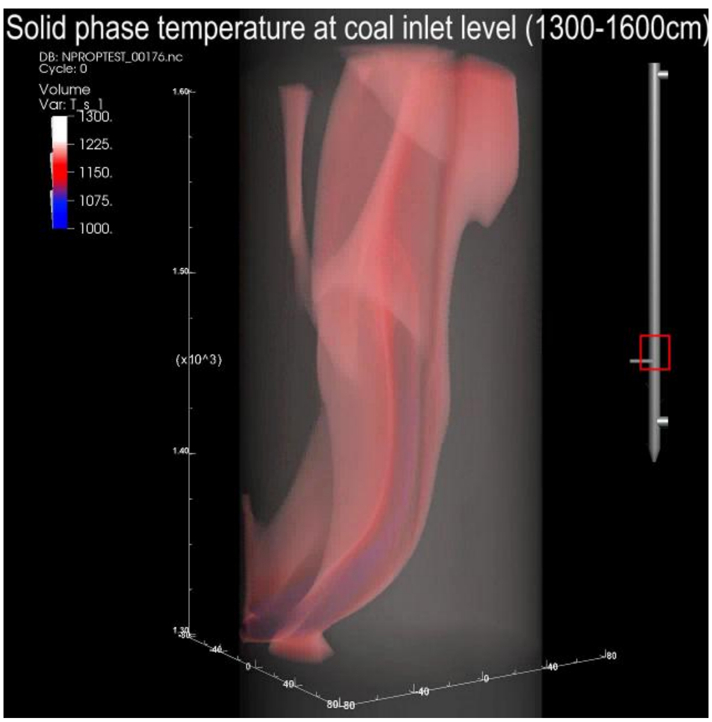

Figure 4.3. Simulation of a coal jet region (solid phase temperature, K). Image courtesy of Chris Guenther, National Energy Technology Laboratory.

The NETL team is specifically working with coal gasification. The simulations, the first of their kind at this scale and resolution, were possible only with the INCITE award, according to the researchers. The researchers pushed their simulation to a 199 million-cell resolution before their allocation ran out.

"Our work has provided an in-depth look at the interactions between the hydrodynamics and chemistry inside a commercial-scale gasifier," said Chris Guenther, research scientist in NETL's Computational Science Division and project leader. "This ability to finely resolve relevant structures inside a dense, reactive gas-solid system is not only unique, but also necessary to accelerate the commercial deployment of advanced gasification technology."

Jaguar's enormous computing power made possible the detailed simulations needed for engineering design of commercial coal gasifiers. No commercial-scale coal gasifiers exist today. Knowing the resolution required in engineering simulations allows engineers to design and place key components, such as inlet ports for coal and oxygen, which then burn incompletely to create hydrogen and carbon monoxide. Compared to the product of complete combustion (carbon dioxide and water) carbon monoxide and hydrogen have significant fuel value.

Guenther's team employs the Multiphase Flow with Interphase eXchanges (MFiX) code, used for simulating the multiphase flows within gasifiers. Multiphase refers to the process of changing a solid (in this case, coal) to a gas (syngas). MFiX was developed at NETL for describing the hydrodynamics, heat transfer, and chemical reactions in fluid–solids systems such as current gas-fired stations, which use very large boilers to produce steam for turbines.

Now, the scientists are able to run detailed simulations on the coal inlet region into the gasifier, allowing them to observe the dynamics. They are also able to do grid independence studies, which means refining the simulations until the results no longer change. This lets them know where they need to be with the simulation resolution and what information might be lost if the simulations are conducted at lower resolutions.

The project is also working on creating several high-resolution gasifier simulations to provide feedback on the design of a commercial-scale gasifier system intended for NETL's Clean Coal Power Initiative. The initiative is a cost-shared venture by the government and industry to develop advanced technologies to supply clean, reliable, and affordable electricity to the United States. Its goal is to sequester 90 percent of the carbon from coal with minimal impact to the cost of electricity.

Madhava Syamlal, principal investigator of the project, summed it up as follows: "High-performance computing is allowing us to reveal and study features of the gas–solids flow in a gasifier to a degree never before possible, experimentally or computationally. The knowledge created from the study will help improve commercial gasifier design."

Simulation Provides a Close-Up Look at the Molecule that Complicates Next-Generation Biofuels

PI: Jeremy Smith, University of Tennessee and ORNL
Time Awarded: 25,000,000 hours, 2010 INCITE; 30,000,000 hours, 2011 INCITE

A team led by Oak Ridge National Laboratory's (ORNL's) Jeremy Smith has taken a substantial step in the quest for cheaper biofuels by revealing the surface structure of lignin clumps down to 1 angstrom (equal to a 10 billionth of a meter, or smaller than the width of a carbon atom). The team's conclusion, that the surface of these clumps is rough and folded, even magnified to the scale of individual molecules, is presented in *Physical Review E* 83, 061911 (2011) (Figure 4.4).

Smith's team employed two of ORNL's signature strengths—simulation on ORNL's Jaguar supercomputer and neutron scattering—to resolve lignin's structure at scales ranging from 1 to 1,000 angstroms. Its results are important because lignin is a major impediment to the production of cellulosic ethanol, preventing enzymes from breaking down cellulose molecules into the sugars that will eventually be fermented.

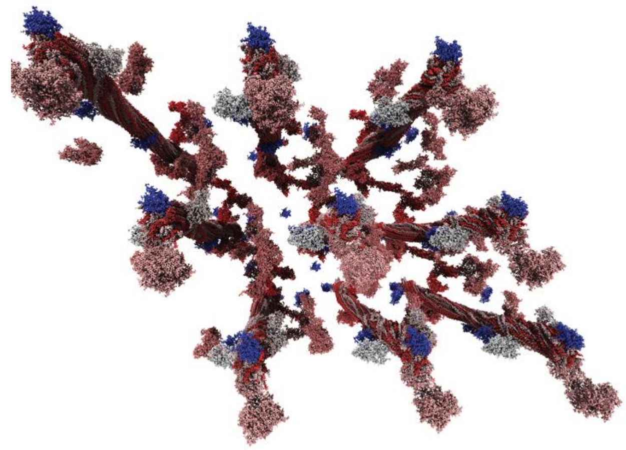

Figure 4.4. Atomic-detailed model of plant components lignin and cellulose. The leadership-class molecular dynamics simulation investigated lignin precipitation on the cellulose fibrils, a process that poses a significant obstacle to economicallv-viable bioethanol

Lignin itself is a very large, very complex polymer made up of hydrogen, oxygen, and carbon. In the wild its ability to protect cellulose from attack helps hardy plants such as switchgrass live in a wide range of environments. When these plants are used in biofuels, however, lignin is so effective that even expensive pretreatments fail to neutralize it.

Switchgrass contains four major components: cellulose, lignin, hemicellulose, and pectin. The most important of these is cellulose, another large molecule, which is made up of hundreds to thousands of glucose sugar molecules strung together. In order for these sugars to be fermented, they must first be broken down in a process known as hydrolysis, in which enzymes move along and snip off the glucose molecules one by one.

According Petridis, the team used neutron scattering with ORNL's High Flux Isotope Reactor to resolve the lignin structure from 1,000 down to 10 angstroms. A molecular dynamics (MD) application called NAMD (for Not just Another Molecular Dynamics program) used Jaguar to resolve the structure from 100 angstroms down to 1. The overlap from 10 to 100 angstroms allowed the team to validate results between methods.

Smith's project is the first project to apply both MD supercomputer simulations and neutron scattering to the structure of biomass. While this research is an important step toward developing efficient, economically viable cellulosic ethanol production, much work remains. For example, this project focused only on lignin and included neither the cellulose nor the enzymes; in other words, it can tells us where the enzymes might fit on the lignin, but it has not yet told us whether the enzymes and lignin are likely to attract each other and attach.

Moving forward, the team is pursuing even larger simulations that include both lignin and cellulose. The latest simulations, on a 3.3 million-atom system, are being done with another MD application called GROMACS (for Groningen Machine for Chemical Simulation).

This research and similar projects have the potential to make bioethanol production more efficient and less expensive in a variety of ways, Petridis noted. For example, earlier experiments showed that some enzymes are more likely to bind to lignin than others. The understanding of lignins provided by this latest research opens the door to further investigation into why that’s the case and how these differences can be exploited.

Nanoscale Simulation of Electron Flow to Elucidate Transistor Power Consumption

PI: Gerhard Klimeck, Purdue University
Time Awarded: 18,000,000 hours, 2010 INCITE; 15,000,000 hours, 2011 INCITE

A team led by Gerhard Klimeck of Purdue University has broken the petascale barrier while addressing a relatively old problem in the very young field of computer chip design.

Using Oak Ridge National Laboratory's Jaguar supercomputer, Klimeck and Purdue colleague Mathieu Luisier reached more than a thousand trillion calculations a second (1 petaflop) modeling the journey of electrons as they travel through electronic devices at the smallest possible scale. Klimeck, leader of Purdue's Nanoelectronic Modeling Group, and Luisier, a member of the university's research faculty, used more than 220,000 of Jaguar's 224,000 processing cores to reach 1.03 petaflops.

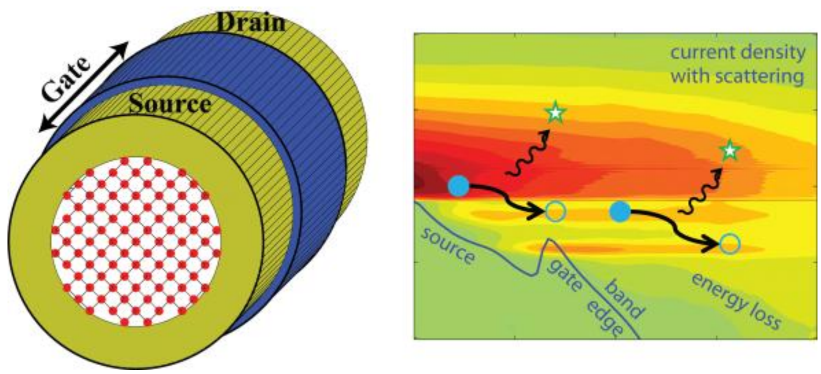

Figure 4.5. Nanowire transistor. At left, schematic view of a nanowire transistor with an atomistic resolution of the semiconductor channel. At right, illustration of electron-phonon scattering in nanowire transistor. The current as function of position (horizontal) and energy (vertical) is plotted. Electrons (filled blue circle) lose energy by emitting phonons or crystal vibrations (green stars) as they move from the source to the drain of the transistor.

"What we do is build models that try to represent how electrons move through transistor structures," Klimeck explained. "Can we come up with geometries on materials or on combinations of materials—or physical effects at the nanometer scale—that might be different than on a traditional device, and can we use them to make a transistor that is less power hungry or doesn't generate as much heat or runs faster? "

The team is pursuing this work on Jaguar with two applications, known as Nanoelectric Modeling (NEMO) 3D and OMEN (a more recent effort whose name is an anagram of NEMO). The team calculates the most important particles in the system—valence electrons located on atoms' outermost shells—from their fundamental properties. These are the electrons that flow in and out of the system. On the other hand, the applications approximate the behavior of less critical particles—the atomic nuclei and electrons on the inner shells (Figure 4.5).

The team is working with two experimental groups.. One is led by Jesus Del Alamo at the Massachusetts Institute of Technology, the other by Alan Seabaugh at Notre Dame. With Del Alamo's group the team is looking at making the electrons move through a semiconductor faster by building it from a material called indium arsenide rather than silicon. With Seabaugh's group the modeling team is working on band-to-band-tunneling transistors. These transistors bear some promise in lower-voltage operation, which could dramatically reduce the energy consumption in traditional field-effect transistors.

Computational End Station Provides Climate Data for IPCC Assessment Reports

PI: Warren Washington, National Center for Atmospheric Research
Time Awarded: 70,000,000 hours, 2010 INCITE; 70,000,000 hours, 2011 INCITE

Supercomputers serve as virtual time machines by allowing scientists to construct and execute mathematical models of the climate system that can be used to explore climate's past and present, and to simulate its future. The results of these complex simulations inform policy and guide climate change strategies, including approaches to mitigation adaptation. Led by Warren Washington of NSF's National Center for Atmospheric Research (NCAR), INCITE projects at the ALCF and OLCF continue to contribute to formulation improvements that lead to improved simulation fidelity, and contribute to experimental archives designed to quantify our knowledge about and uncertainties in the climate system.

The involved researchers have also developed a Climate-Science Computational End Station (CCES) to solve grand computational challenges in climate science. End-station users have continued to improve many aspects of the climate model and then use the newer model versions for studies of climate change with different emission scenarios that would result from adopting different energy policies. Climate community studies based on the project's simulations will improve the scientific basis, accuracy, and fidelity of climate models. Validating that models correctly depict Earth's past climate improves confidence that simulations can more accurately simulate future climate change. Some of the model versions have interactive biochemical cycles such as those of carbon or methane. A new DOE initiative for its laboratories and NCAR is Climate Science for Sustainable Energy Future (CSSEF), which will accelerate development of a sixth-generation CESM. The CCES will directly support the CSSEF effort as one of its main priorities.

The CCES will advance climate science through both aggressive development of the model, such as the CSSEF, and creation of an extensive suite of climate simulations. Advanced computational simulation of the Earth system is built on a successful long-term interagency collaboration that includes NSF and most of the major DOE national laboratories in developing the CESM, NASA through its carbon data assimilation models, and university partners with expertise in computational climate research. Of particular importance is the improved simulation of the global carbon cycle and its direct and indirect feedbacks to the climate system, including its variability and modulation by ocean and land ecosystems.

Washington and collaborators are now developing stage two of the CCES with a 2011 INCITE allocation of 70 million processor hours at the OLCF and 40 million at the ALCF. The work continues development and extensive testing of the CESM, a newer version of the CCSM that came into being in 2011.

The CCES INCITE project will provide large amounts of climate model simulation data for the next IPCC report, AR5, expected in 2014. The CESM, which will probably generate the largest set of publicly available climate data to date, will enable comprehensive and detailed studies that will improve the level of certainty for IPCC conclusions.

Getting much more realism requires running simulations at the highest possible resolution. Increasing resolution by a factor of two raises the calculating time by nearly an order of magnitude, he added. More grid points in the horizontal plane mean the supercomputer has to take smaller steps—and more computational time—to get to the same place.

The quest for greater realism in models requires ever more powerful supercomputers. Having learned a great deal about Earth's climate, past and present, from terascale and petascale systems, scientists look longingly to future exascale systems. A thousand times faster than today's quickest computers, exascale supercomputers may enable predictive computing and will certainly bring deeper understanding of the complex biogeochemical cycles that underpin global ecosystems and make life on Earth sustainable.

Medal of Science Winner

Warren Washington, who was named Oct. 19 by President Obama as a National Medal of Science winner, is a familiar name around the OLCF. The National Center for Atmospheric Research senior scientist and former chair of the National Science Board has collaborated with ORNL on climate modeling since the earliest days of the laboratory's supercomputing renaissance, going back to the Intel Paragon.

According to James Hack, director, of the OLCF and Climate Change Science Institute, Washington has been seminally involved in adapting global climate models to distributed-memory parallel computing environments, which has been a major thrust of ORNL supercomputing. He has served as a principal investigator and advisor on OLCF allocations, including Jaguar's role in simulations cited in the fourth Intergovernmental Panel on Climate Change assessment report.

Read the full press release here: http://www.whitehouse.gov/the-press-office/2010/10/15/president-obama-honors-nations-top-scientists-and-innovators.

Whole-Genome Sequencing Simulated on Jaguar

PI: Aleksei Aksimentiev, University of Illinois at Urbana-Champaign
Time Awarded: 10,000,000 hours, 2010 INCITE

The Human Genome Project paved the way for genomics, the study of an organism's genome. Personalized genomics can establish the relationship between DNA sequence variations among individuals and their health conditions and responses to drugs and treatments. To make genome sequencing a routine procedure, however, the time must be reduced to less than a day and the cost to less than $1,000—a feat not possible with current knowledge and technologies. Using ORNL's Jaguar, Aleksei Aksimentiev, assistant professor in the physics department at the University of Illinois–Urbana-Champaign, and his team are developing a nanopore approach, which promises a drastic reduction in time and costs for DNA sequencing (Figure 4.6). Their research reveals the shape of DNA moving through a single nanopore—a protein pore a billionth of a meter wide that traverses a membrane. As the DNA passes through the pore, the sequence of nucleotides (DNA building blocks) is read by a detector.

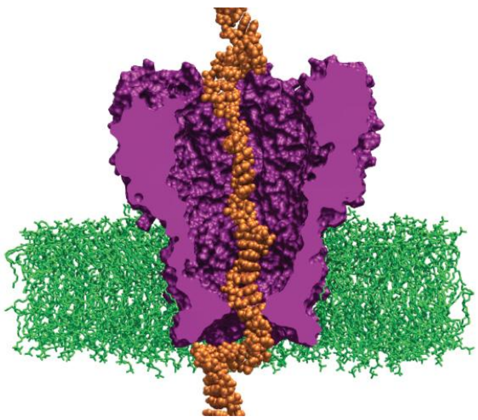

Figure 4.6. Scientists simulate DNA interacting with an engineered protein. The system may slow DNA strands travelling through pores enough to read a patient's individual genome. (Image courtesy of Aleksei Aksimentiev.)

Aksimentiev's group uses the nanopore MspA, an engineered protein. Its sequence must be altered to bind more strongly to the moving DNA strand. MspA is an ideal platform for sequencing DNA because scientists can now measure dams in the pore, which could slow DNA's journey through the protein. Altering the MspA protein to optimize dams is both time-consuming and costly in a laboratory but simple on a computer. The team received 10 million processor hours on Jaguar through the INCITE program. With the INCITE allocation, the scientists were able to reproduce the dams in the MspA nanopore for the type of DNA nucleotides confined to it, slowing down the sequence movement through the nanopore. "We have carried out a pilot study on several variants of the MspA nanopore and observed considerable reduction of the DNA strand speed," said Aksimentiev. "These very preliminary results suggest that achieving a 100-fold reduction of DNA velocity, which should be sufficient to read out the DNA sequence with single-nucleotide resolution, is within reach. Future studies will be directed toward this goal."

Simulations Explore Interactions of Quarks and Gluons and Reveal a New Bound State of Baryons

PI: Paul Mackenzie, Fermilab
Time Awarded: 40,000,000 hours, 2010 INCITE; 30,000,000 hours, 2011 INCITE

Protons and neutrons in an atom contain smaller particles called quarks and gluons. Nearly all the visible matter in the universe is made up of these subatomic particles. Quarks and gluons interact in fascinating ways. For example, the force between a quark and an antiquark remains constant as they move apart. Quarks are classified into six "flavors"—up, down, charm, strange, bottom, and top—depending on their properties. Gluons, for their part, can capture energy from quarks and function as glue to bind quarks. Groups of gluons can also bind, forming glueballs. Scientists have identified another unique property of gluons, which they describe as color. Quarks can absorb and give off gluons, and when they do so, they are said to change color. Scientists believe quarks seek to maintain a state of color balance, and to do so are bound into the protons and neutrons that make up our world.

The scientific community recognizes four fundamental forces of nature—electromagnetism, gravity, the strong force (which holds an atom's nucleus together), and the weak force (responsible for the ability of a quark to change its "flavor"). With the exception of gravity, all these forces are believed to be described in terms of "gauge theories". The gauge theory describing the strong interaction in terms of quarks and gluons is called quantum chromodynamics, or QCD. A team of scientists collaborating under the leadership of Paul Mackenzie of Fermi National Accelerator Laboratory has been awarded a total of 70 million processor hours at the Oak Ridge Leadership Computing Facility (OLCF) and the Argonne Leadership Computing Facility (ALCF) to understand the consequences of QCD.

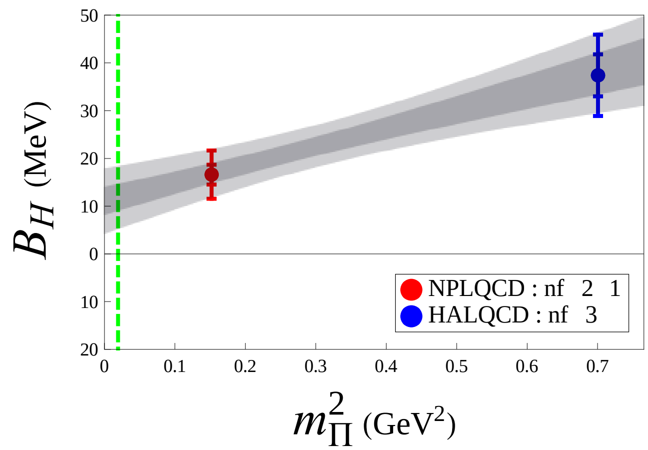

Figure 4.7. Lattice QCD calculations of strongly interacting particles. The binding energy of two Λ baryons by the NPLQCD team and by HaLQCD. The results suggest the existence of a bound H dibaryon or near-threshold scattering state at the physical up and down quark masses. (Image courtesy NPLQCD Collaboration, S. Beane et al.)

"Leadership class computing makes it possible for researchers to generate such precise calculations that someday theoretical uncertainty may no longer limit scientists' understanding of high-energy and nuclear physics," said Mackenzie.

Using Monte Carlo techniques to predict the random motions of particles, the simulations generate a map of the locations of up, down, and strange quarks on a fine-grained lattice. The up and down quarks have masses sufficient to enable researchers to extrapolate physical properties. The team is studying three distinct quark actions – clover, domain wall and improved staggered – to explore different facets of QCD. For the clover quarks, the team has used OLCF to generate a set of lattices with spacing 0.12 femtometers, and extents up to 4 femtometers. These lattices are subsequently used to compute properties of baryons, such as protons and neutrons, and mesons, such as the pion, and their interactions.

The Nuclear Physics with Lattice QCD (NPLQCD) Collaboration investigated a two-baryon system with two strange quarks, and compared its mass with that of two free **Λ** baryons, each comprising one up, one down and one strange quark. By performing the calculations at several volumes, the team found evidence for a new bound state, the "H dibaryon." These calculations will further a description of the nucleus in terms of the fundamental quarks and gluons of QCD, and by exploring the interactions of baryons, such as the **Λ**, for which there is little experimental data, address key astrophysical questions such as core collapse in supernovae.

Scientific Support

4.2.1 Scientific Liaisons

The OLCF pioneered a total user support model widely recognized as a best practice for HPCCs: the SciComp liaison program, comprising experts in their scientific discipline, including PhD-level researchers, who are also specialists in developing code and optimizing HPC systems. Support ranges from basic support—access to computing resources—to complex, multifaceted support for algorithm development and performance improvement. The liaison program is one of the reasons for the success of the OLCF.

Today, OLCF liaison support encompasses a range of activities, including the following:

- Improving performance and scalability of project application software
- Assisting in redesign, development, and implementation of strategies that increase effective use of HPC resources
- Implementing scalable algorithm choices and library-based solutions
- Providing an advocacy interface to OLCF resource decisions, including the RUC
- Performance modeling, including anticipating the impact of upgrades and fine-tuning applications for maximum efficiency
- Scaling applications to make effective use of the OLCF's petascale resources
- Assisting with code development and algorithms
- Preparing for the next generation of supercomputing
- Being members of the computational science teams

This approach provides a nurturing, exhilarating environment not only for scientists and engineers using OLCF resources but also for OLCF staff members. And the need has never been greater. We are poised on the precipice of a great leap forward in computing. To paraphrase Rob Farber, a senior research scientist at Pacific Northwest National Laboratory, in the future we may look back on these next few years as the era of the GPU,[1] for certainly the concept of the general purpose GPU (GPGPU) has become a reality. And as Farber has indicated, woe to those who don't adapt to the future (i.e., adapt legacy code to GPGPU and hybrid CPU-GPU technology). Which means that in addition to the support services SciComp liaisons typically provide, they are now reviewing software and rewriting code in preparation for the next generation of machines and this new era, which is reflected in many of the success stories detailed on the following pages.

One Eye Always on the Future

With one eye on the future and one on customer support, SciComp's Mike Brown, a molecular dynamics (MD) specialist with a background in both the biomedical and the computer sciences, is working on adapting LAMMPS (Large-Scale Atomic/Molecular Massively Parallel Simulator) and other codes to run on hybrid CPU-GPU machines like the OLCF's next generation Titan. LAMMPS is a classical MD code that can be used to model atoms or, more generically, as a parallel particle simulator at the atomic, meso, or continuum scale (Figure 4.8). LAMMPS is open source; highly portable; and easy to download, install, and run. Because of this it is much in demand.

[1]Farber, Rob, "Redefining what is possible," in *Scientific Computing* [http://www.scientificcomputing.com/articles-HPC-GPGPU-Redefining-What-is-Possible-010711.aspx (last accessed 7-12-11)].

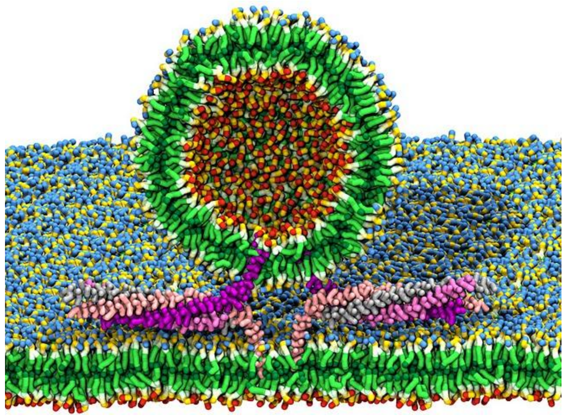

Figure 4.8. Coarse grain representation of a SNARE. [SNAP (soluble NSF attachment protein) REceptor] complex tethers a vesicle to a lipid bilayer. Used for MD simulations to study how SNARE proteins mediate the fusion of vesicles to lipid bilayers, an important process in the fast release of neurotransmitters in the nervous

This past year, working with Axel Kohlmeyer, ICMS associate director and an expert on MD codes like LAMMPS; NVIDIA's Peng Wang; SNL's Steve Plimpton, lead developer of LAMMPS; and Arnold Tharrington, lead on the OLCF LAMMPS CAAR effort, Brown researched algorithms that would allow the LAMMPS MD simulator to run with GPU acceleration on the OLCF's CPU-GPU test cluster. The main focus was twofold: (1) minimizing the amount of code that needed to be ported for efficient acceleration (to avoid rewriting the legacy code in its entirety) and (2) efficiently using the processing power from both the CPU and the GPU resources (the team intuited that some tasks might be better suited to one or the other of the platforms and using the CPU cores could reduce the amount of code that had to be ported to the accelerators). The LAMMPS Accelerator Library (http://users.nccs.gov/~wb8/gpu/lammps.htm), now distributed as part of the main LAMMPS software package and thus freely available to all MD researchers, is one of the main outcomes of this research to date. (A detailed description of the algorithms used for acceleration of short-range models has been published,[1] and publications on algorithms and simulation results for long-range models are in preparation.) The library, which allows simulations to be run between 2 and 14 times faster on InfiniBand GPU clusters, is already being applied by LAMMPS users for science applications and will facilitate an early capability for INCITE users to utilize the impressive floating-point capabilities on the Titan machine with full compatibility with all of LAMMPS traditional CPU features.

Improving Performance and Scalability

Tools for performance measurement and analysis in the HPC environment are not well understood outside university computer science departments and HPCCs like the OLCF. Consequently, users of HPC resources tend to make guesses about the performance of their codes or, worse, ignore performance entirely—highly problematic in terms of efficient, effective use of compute resources. SciComp staff members like Rebecca Hartman-Baker are addressing this head-on through aggressive use of advanced profiling tools like the Vampir (Visualization and Analysis of MPI Resources) suite of tools added last year. VampirTrace instruments codes and produces trace files when run. The trace files are then loaded into Vampir, which is used to visualize the trace; the output is a timeline trace of the workings of an application with the timeline of the code along the x-axis and processor numbers along the y-axis. Events are represented by colored blocks, dots, and lines, and subroutines or functions of particular interest can be color-coded to stand out.

"I liken profiling to getting an energy audit of your home," says Hartman-Baker. "An energy audit can tell you where you are consuming and possibly wasting energy… and you can analyze the results and figure out what changes to make. Likewise, profiling tells you where your code is spending its time so you can analyze the results and fix the code."

[1]Brown, W. M.; Wang, P.; Plimpton, S. J., and Tharrington, A. N., "Implementing molecular dynamics on hybrid high performance computers—short range forces," *Computer Physics Communications*, 182, pp. 898–911 (2011).

When the Vampir suite of tools was added last year, SciComp staff immediately commenced putting it through its paces, with some surprising—and exciting—results. A good example is the BIGSTICK configuration-interaction shell-model code, which is used to solve the general many-fermion problem (important in nuclear physics). While the code is supposed to work well on both serial and parallel machines, when Hartman-Baker profiled it using VampirTrace and Vampir, she found that the code had a number of inefficiencies in its implementation of the Lanczos method for eigenvalues and eigenvectors. This is a case of an algorithm that looked good on paper not performing well in practice. She compiled the results and supporting visualizations into a report in which she outlined suggestions for improving the algorithm, based on both the Vampir analysis and her own expertise in numerical algorithms. Hartman-Baker's analysis and suggestions were discussed at the 2011 UNEDF (Universal Nuclear Energy Density Functional) meeting,[1] and the project team is now planning to submit a request for a DD allocation to test the reformulated code in preparation for an INCITE application in 2013. Because of Hartman-Baker's initiative, a potential future INCITE awardee has been helped to "get up to speed," which Hartman-Baker finds particularly gratifying as the OLCF is always looking for new projects. It's also a great example of how the OLCF and its staff members provide continuous support to the larger HPC community.

In a similar case, Hartman-Baker was asked by the code developers to profile the j-coupled version of NUCCOR. This is a nuclear physics code that takes advantage of symmetries in certain nuclear configurations to perform energy calculations on larger nuclei than can currently be studied with this code in the nonsymmetric case. Profiling showed that on the small test problem, the code was spending more than half its time in a subroutine called sort. This sort subroutine was an implementation of an algorithm reminiscent of bubble sort, a particularly inefficient sort algorithm with a complexity proportional to the square of the number of items to be sorted. Using Hartman-Baker's previous analogy of an energy audit, this was "equivalent to running air conditioning with all the windows open and then not even realizing that the power bill is too high." Hartman-Baker's suggested solution was to implement a heap sort, which would reduce sorting to about 3% of the total run time; however, in consultation with the authors of the code, it was determined that sorting was unnecessary, so the sorting subroutine was removed altogether, resulting in a 30% speedup on the full problem. This is not inconsequential. Anytime you can get 30% more, it's a good thing, but in this case, the more is 30% more science for the same cost in CPU hours—a real deal for tax payers and the nation.

Supporting Software

VASP. One of the most important services the SciComp staff provides, and one that often goes unnoticed, is support for the software running on OLCF platforms that makes user codes run faster. VASP, the Vienna Ab-initio Simulation Package, is a workhorse in the materials science world, used at more than 1,400 sites worldwide, and one of the premier electronic structure codes used by a number of INCITE and DD projects. However, according to Markus Eisenbach, who is primarily responsible for maintaining VASP and assisting OLCF users with it, it doesn't scale particularly well. What makes this particularly challenging is that it isn't open source software, so he can't really develop it, yet he must find a way to optimize it on OLCF platforms. What Eisenbach does is provide precompiled versions of both of the commonly used VASP releases (4.6 and 5.2, released in 2010), optimized for OLCF, to licensed users on OLCF systems. The most recent version, 5.2, provides significant new physics capabilities such as exact exchange and hybrid functionals, and while it ports reasonably well, Eisenbach's background in condensed matter physics, combined with his HPC expertise, enables him to better understand the needs of users and help them get the most from the VASP code on OLCF machines.

[1]Johnson, Calvin; Ormand, Erich; and Krastev, Plamen, "Progress report on the BIGSTICK configuration-interaction code," presented at the UNEDF 2011 Annual/Final Meeting June 20–24, 2011, East Lansing, Michigan (available at http://unedf.org/content/MSU2011_talks/Wednesday/Johnson_UNEDF2011.pdf).

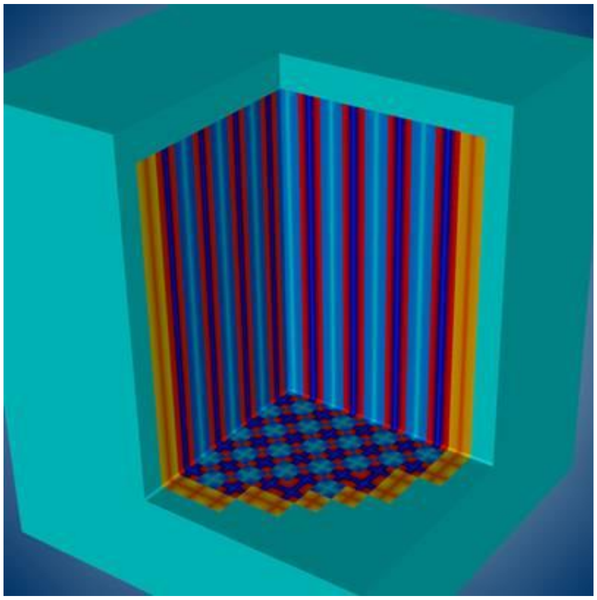

Figure 4.9. Simulation of PWR900 core model, 3-D view showing axial (z-axis) geometry. The assembly enrichments are low-enriched uranium (light blue), medium-enriched uranium (red/blue), and highly enriched uranium (yellow/orange).

Denovo. Denovo is the ORNL radiation transport code developed specifically to take advantage of the computational power of high-performance computers such as Jaguar. See Figure 4.9 for a sample simulation of a PWR900 core. Because of Denovo's broad applicability to radiation transport modeling, new applications continue to be found, including assistance with the Fukushima reactor (see separate visualization story below). Last year's OA report discussed some of the changes Denovo developer Tom Evans was making in concert with SciComp liaison Wayne Joubert, including optimizations for GPU processors. Thanks to Joubert and the Denovo team's work, the latest version of Denovo runs 2 × faster than the previous code on conventional processors, runs an astounding 40 × faster on the NVIDIA Fermi GPU compared to a Jaguar processor core, and is significantly more scalable than its predecessor (scaling up to 200,000 cores). However, as Joubert says, "it's the nature of the business that we're always looking at the slowest part of a code for ways to speed it up or otherwise improve it."

Such was the case with Denovo this past year. Changes had previously been made in the Denovo algorithms to make the code run efficiently on the new OLCF GPU-based system, Titan. This involved introduction of a whole new dimension of parallelism into the code—parallelism across energy groups to improve scalability and GPU performance. Continuing to look for ways to improve the code, the Denovo team found that the energy-set reduction operation was the least scalable part of the enhanced code. After studying it briefly, team members asked Joubert to help them with a solution to the problem. The code originally used MPI_Allreduce, a generic function, for the energy-set reduction operation. Using his knowledge of MPI, Joubert was able to recommend a fairly obscure offshoot, MPI_Reduce_scatter, that could be used for this case as an alternative method to perform the reduction operation. By simplifying the information that the various processors get, MPI_Reduce_scatter improves communication performance and memory usage, making the reduction step run 3 × faster. This is a classic example of the type of work that liaisons do regularly for their projects. Though this magnitude of improvement is not as high as is sometimes possible from incorporating an entirely new algorithm, it is still an important improvement going forward because the time spent by Denovo in the energy-set reduction operation will become increasingly significant for larger problems and future parallel systems. And with that same eye on the future common to all OLCF staff members, Joubert is currently implementing new algorithms that will allow Denovo to exploit the power of GPUs on a much broader range of problems of interest to Denovo users—for the machines of the future . . . for Titan.

4.2.2 Visualization Liaisons

Most projects are assigned a visualization liaison in addition to a primary scientific liaison to maximize opportunities for success on the leadership computing resources. This approach stems from recognition that scientific discovery relies on more than just volumes of data. The ultimate goal is to make sense of the data, and visualization schemes are key to this. In fact, OLCF visualization scientists do more than strengthen a project's data analysis and help illuminate project results; in many cases they also help in detecting and fixing problems. In addition to customary visualization support services, OLCF visualization experts frequently find themselves developing custom software and algorithms to address unique user challenges—and in some cases responding to emergencies.

Responding to Emergencies

What we do is critically important, not only to national, but also to world security. This was never more evident than in the OLCF's rapid response to the Fukushima nuclear accident. In the days following the March 11, 2011, massive earthquake and subsequent tsunami, DOE staff and experts from ORNL and other national laboratories sprang into action to help collect, analyze, and interpret data to provide the Japanese government and others with critical information. One of these groups consisted of OLCF visualization experts Jamison Daniel, Mike Matheson, and Dave Pugmire. According to Pugmire, one of the major issues was the state of rods in the spent fuel pool. Following the earthquake and tsunami, there was concern that the spent fuel pool had been compromised and that water had leaked out as a result. A loss of water could result in fuel rod heating and damage. Further, because the spent fuel pool consisted of rods that had been removed from the reactor at different times, the response to the level of the water would be different for each set of rods.

Working with ORNL Reactor & Nuclear Systems Division staff members, the visualization liaisons took blue prints and CAD models of the reactor building, spent fuel pool, and fuel bundle layouts to create a three-dimensional (3-D) model of the Fukushima plant. This 3-D model was then read into Maya and Blender (high end rendering packages) where camera animation could be added to explore the condition of the reactor (Figure 4.10). Two simulations were incorporated into the visualizations, which showed the temperature of fuel rods, the temperature of the water, and the dose levels as a function of the level of the water.

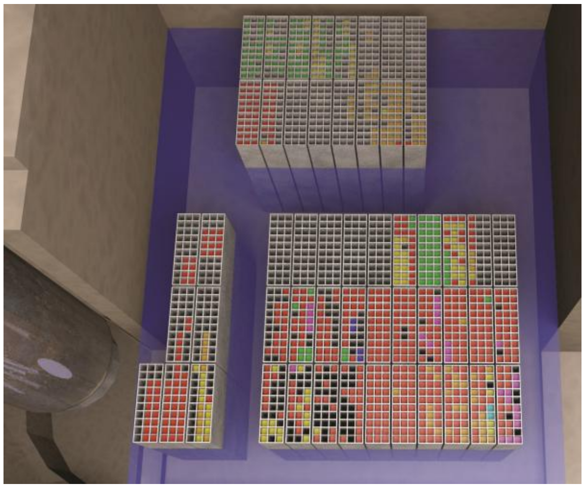

Figure 4.10. Rendering of the Fukushima reactor building spent fuel rod pool.

This illustrates how the reactor simulation capability at ORNL can be used to model a very complex, time critical event. All of this was accomplished within an incredibly short time frame. In the weeks and months since then, the visualization team has continued to refine their analyses and visualizations. Even though the accident has been contained, shutdown and cleanup of the facility will likely take years, and the ORNL team will continue to play an important role in these efforts.

Pulling Information from Raw Data

The production of ethanol from cellulose is a clean, nearly carbon neutral technology. Thus, efficient production of ethanol through hydrolysis of cellulose into sugars is a major energy-policy goal. With an INCITE grant of 25,000,000 hours, Jeremy Smith is performing highly parallelized multi-length-scale computer simulations to help understand the physical causes of the resistance of plant cell walls to hydrolysis—the major technological challenge to developing cellulosic bioethanol. Using the power of HPC, Smith and his team hope to derive information about lignocellulosic degradation unprecedented in its detail. As might be suspected, the atomistic MD simulations of lignin molecules involved create large amounts of data. This was problematic in two respects: (1) the time dependent nature of the simulations was difficult to understand with simple graphics and (2) some of the large amount of data to be processed obscures other data key to gaining insights. Because advanced visualization techniques, including animation, can aid in the analysis of such data, Mike Matheson, a visualization liaison with a background in engineering, was assigned to the team. Mike's experience with HPC, and especially visualization, enabled him to select the software most suitable to this application. Using Tachyon and the Blender 3D software, Matheson developed a method to deal with the obscuring data in an intelligent manner so Smith and his team could "see" what was important. The high quality renderings combined with this technique enhanced the team's ability to explain the simulations, especially to others, and enabled them to gather detailed knowledge of the fundamental molecular organization, interactions, mechanisms, and associations of bulk lignocellulosic biomass (Figure 4.11), as well as other insights, from the data.

Initial results were presented on the EVEREST powerwall, but later versions using the technique have been delivered as portable animations that can be played on desktops or laptops at conferences and during presentations. As with other SciComp success stories, the success of this work was based on the close collaboration between Matheson and the project team. They discussed the problems, talked about potential solutions, and tried various solutions to converge on the successful strategy together.

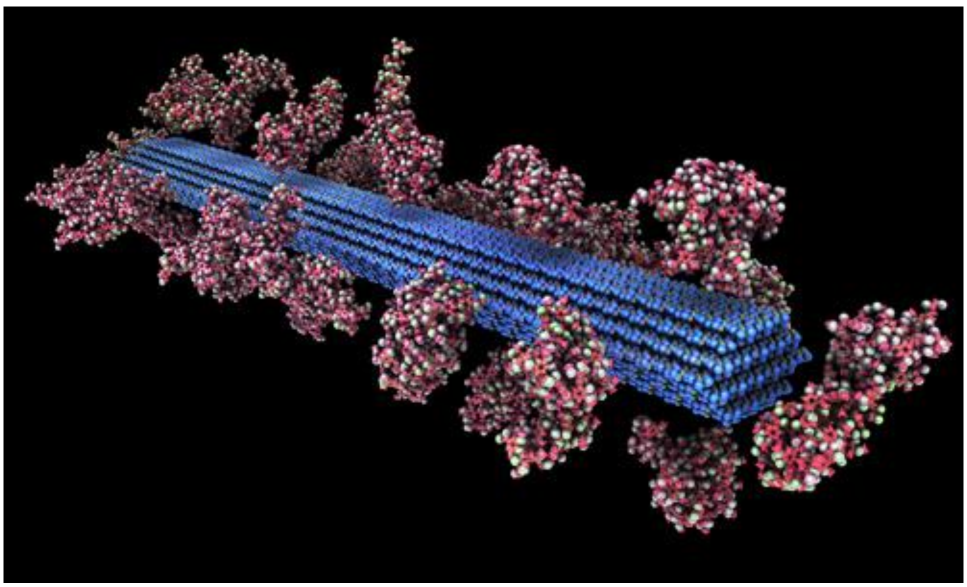

Figure 4.11 Lignin molecules aggregating on a cellulose fibril.

4.3 ALLOCATION OF FACILITY DIRECTOR'S RESERVE

2011 Operational Assessment Guidance – Allocation of Facility Director's Reserve Computer Time

The Facility describes how the Director's Reserve is allocated and lists the awarded projects, showing the PI name, organization, hours awarded, and project title.

The OLCF allocates time on leadership resources primarily through the INCITE program and through the facility's Director's Discretionary (DD) program. The OLCF seeks to maximize scientific productivity via capability computing through both programs. Accordingly, a set of criteria are considered when making allocations, including the strategic impact of the expected scientific results and the degree to which awardees can make effective use of leadership resources. Further, up to 30% of the facility's resources are allocated through the Advanced Scientific Computing Research Leadership Computing Challenge (ALCC) program.

4.3.1 Innovative and Novel Computational Impact on Theory and Experiment

In early 2011, DOE initiated a review of the INCITE program to assess the processes the Argonne and Oak Ridge Leadership Computing Facilities (ALCF, OLCF) use to solicit, review, recommend and document proposal actions and monitor active project[s] and evaluate their INCITE portfolio. The six-member panel of national and international experts met in June with the INCITE manager and OLCF and ALCF senior management. There were no negative findings. The panel judged that the program has addressed the 2008 Committee of Visitors recommendations from the previous review of INCITE and had few additional suggestions. The INCITE manager and center directors were complimented for their effective management of the program.

A total of approximately 1.7 billion processor hours were allocated to 57 INCITE projects in CY 2011. (930 billion hours on OLCF's Cray XT Jaguar were awarded to 32 projects and 732 billion hours on ALCF's IBM Blue Gene/P were awarded to 30 projects; several projects received awards of time at both centers). The scientific peer-review was carried out with nine panels of experts, nearly seventy reviewers in total. INCITE is open to researchers from around the world and the panels reflect this: 15% of the reviewers were from outside of the U.S.

The 2012 INCITE Call for Proposals (CFP) yielded a total of 119 submittals. These submittals are currently undergoing computational readiness and scientific review. The demand for time on the leadership systems continues to be high. In the 2012 CFP INCITE received requests for 5 billion hours of time, nearly 3× greater than the combined OLCF and ALCF hours available for allocation.

Peer review represents a best practice for the assessment of programmatic efficacy as well as for the identification of high-impact research activities. For INCITE, not only are the proposals peer reviewed, we also ask the scientific panels to provide INCITE management with feedback about the quality of the submittals and the operation of the program. To gauge the quality of the proposals received, the panel reviewers are asked to rate their response to the statement "INCITE proposals discussed in the panel represent some of the most cutting-edge computational work in the field." On a scale from 1 (strongly disagree) to 5 (strongly agree), the reviewers in the 2010 and 2011 CFP strongly agreed, with average ratings of 4.51 and 4.52, respectively. 94% of the attending panel reviewers last year responded. See Table 4.2 for the survey questions and average responses. Average scores are based on ratings between 1 ("strongly disagree") to 5 ("strongly agree").

Table 4.2 Results of survey of INCITE scientific peer-reviewers at the annual panel review meeting

	2010 INCITE CFP Avg. Scores	2011 INCITE CFP Avg. Scores
INCITE proposals discussed in the panel represent some of the most cutting-edge computational work in the field	4.51	4.52
The proposals were comprehensive and of appropriate length given the award amount requested	3.89	4.15
Please rate your overall satisfaction with the 2010 [2011] INCITE Science Panel review process (where 1 is "very dissatisfied" and 5 is "very satisfied)	4.67	4.79

Refinements to the program policies and procedures were introduced in April 2010 for the 2011 CFP: see the 2010 Operational Assessment Report for details. These changes resulted in an improvement in the panel rating for the second survey statement "The proposals were comprehensive…" with an increase in average rating from 3.89 to 4.15. Additional changes were introduced in April 2011 for the 2012 CFP. For example, the program revised the renewal proposal form (the new-submittal form was previously re-done) and emphasized the authors' achievements to date. After the 2012 CFP ended, the authors were invited to respond to a short survey asking for input about the proposal form and templates. Nearly 20% of the authors responded and expressed satisfaction with the INCITE proposal form. Several suggested modifications which will be incorporated into the 2013 INCITE CFP. Some comments are provided here.

- "Templates were great, wish other programs such as Teragrid, GENCI or PRACE provided these."
- "I really think the increased emphasis on results for renewals is a good change. Previous years it seemed like the important thing was how many jobs were run and at what size for each objective, and not so much what you get out of the simulations. Since obtaining science results is the ultimate objective, this change is appropriate, and prevents users spending time collecting statistics that are not particularly enlightening themselves when it comes to science results."

Authors also provided suggestions for future consideration.

- "I would like to see in the proposal the section devoted to a position of the proposed project as compared with the existing 'state of the art' in the field of the proposal."
- "I had trouble figuring out how the best way to report some of our Computing Resource Allocations. They did not follow a fiscal year pattern and the webpage only allowed one to enter fiscal years. Maybe having the option to give start and end date would help."

4.3.2 ASCR Leadership Computing Challenge Program

Open to scientists from the research community in academia and industry, the ALCC program allocates up to 30% of the computational resources at NERSC and the leadership computing facilities at Argonne and Oak Ridge for special situations of interest to DOE, with an emphasis on high-risk, high-payoff simulations in areas directly related to the department's energy mission in areas such as advancing the clean energy agenda and understanding the Earth's climate, for national emergencies, or for broadening the community of researchers capable of using leadership computing resources. The call for proposals will be issued annually for single year proposals; however, proposals for single year allocations may be submitted at any time during the calendar year. Proposals submitted to the ALCC program will also be subject to peer review of scientific merit based on guidelines established in 10 CFR Part 605.

4.3.3 Director's Discretionary Program

The DD program provides a valuable mechanism for the investigation of rapidly changing technology or unanticipated scientific opportunities that frequently arise outside the standard (INCITE) annual proposal cycle. The goals of the DD program are threefold: development of strategic partnerships, leadership computing preparation, and application performance and data analytics.

Strategic partnerships are partnerships aligned with strategic and programmatic ORNL directions. These are entirely new areas or areas in need of nurturing. Example candidate projects are those associated with the ORNL Laboratory Directed Research and Development Ultrascale Computing Program, programmatic science areas (bioenergy, nanoscience, climate, energy storage, engineering science), and key academic partnerships (e.g., that with the ORNL Joint Institute for Computational Sciences).

The DD program must help to identify and develop new computational science areas expected to have significant leadership class computing needs in the near future as well as exploit existing computational science areas where a leadership computing result can lead to new insight, an important scientific breakthrough, or program development. Candidates for such leadership preparation projects include those from industry, the SciDAC program, end station development, and exploratory pilot projects.

The DD program must also enable porting and development exercises for infrastructure software such as frameworks, libraries, and application tools; and support research areas for next-generation OSs, performance tools, and debugging environments. Candidates for such application performance and data analytics projects include application performance benchmarking, analysis, modeling, and scaling studies; end-to-end workflow, visualization, and data analytics, basic computer science research; and system software and tool development.

The Industrial Partnerships Program is part of the DD program and reflects the laboratory's strategy to provide opportunities for researchers in industry to access the leadership-class systems to carry out work that would not otherwise be possible.

The duration of DD projects is typically shorter than INCITE projects for two reasons: DD projects are intended to solve a problem within a finite period of time (e.g., scalability development) or be a prelude to a formal INCITE submittal, which is the appropriate vehicle for long-term research projects. The actual DD project lifetime is specified upon award, where most allocations are for less than 1 year.

The Resource Utilization Council (RUC, Reference Section 3) makes the final decision on DD applications, using written input from subject matter experts. Once allocations are approved, DD users are held to basically the same standards and requirements as INCITE users.

Since its inception in 2006, the DD program has granted allocations in virtually all areas of science identified by DOE as strategic for the nation (Table 4.3). Additional allocations have been made to promote science education and outreach. Requests and awards have grown steadily each year (Table 4.4). The complete list of current Director's Discretionary projects is provided in Appendix A.

Table 4.3 Director's Discretionary Program: Domain Allocation Distribution

Time Period	Biology	Chemistry	Computer Science	Earth Science	Engineering	Fusion	Materials Science	Nuclear Energy	Physics
2008	19%	8%	28%	4%	8%	15%	3%	1%	14%
2009	5%	3%	19%	6%	8%	6%	33%	1%	19%
2010	9%	6%	10%	8%	19%	6%	16%	3%	23%
2011 YTD	8%	5%	11%	18%	17%	3%	14%	6%	18%

Table 4.4 Director's Discretionary Program: Awards and User Demographics

Year	Project Awards	Project Requests	Hours Available (M)	Hours Allocated (M)	User Demographics (%)
2008	36	38	18.33	8.5	42.7 DOE 3.8 Government 6.4 Industry 47.1 Academic
2009	47	51	125	38	55.9 DOE 0.7 Government 9.9 Industry 33.5 Academic
2010	77	85	160	85	46.0 DOE 2.3 Government 12.2 Industry 39.5 Academic
2011 YTD	88	95	160	110	41.4 DOE 1.7 Government 9.1 Industry 47.1 Academic 0.7 Other

Annual DD allocations are typically less than the available hours. We review and allocate DD proposal requests on a weekly basis through the RUC. With this approach, the OLCF can remain flexible and responsive to new project requests and research opportunities that arise during the year. The leadership computing resources are effectively utilized because INCITE and ALCC users are not "cut off" when they overrun their allocation. Rather, they are allowed to continue running at lower priority to make use of potentially available time.

The DD allocation is an important resource and necessary for ORNL to advance computational science priorities, and the OLCF will continue to actively manage this allocation. Jack Wells, OLCF director of science, is currently leading a review of DD policies to evaluate their effectiveness and consider possible modifications.

4.3.4 Industrial Partnership Program

The Industrial HPC Partnership Program is gaining traction and attracting both large and small firms (Table 4.5 lists projects active in in CY 2010 and/or CY 2011 YTD). Excluding the INCITE preparatory projects, one-fourth of the industry projects were from small businesses, affirming that large complex problems are not the exclusive purview of large companies. Small companies, the backbone of a growth economy and the source of many advances in innovation, also are tackling tough scientific challenges and relying on modeling and simulation with high performance computing to achieve their results.

Table 4.5 Industry Projects at the OLCF

Corporate Partner	Program	Description
Boeing	INCITE	Development and correlation of computational tools for transport airplanes
General Motors	INCITE	Electronic, Lattice, and Mechanical Properties of Novel Nano-Structured Bulk Materials
Ramgen	ALCC	High resolution design-cycle computational fluid dynamics analysis supporting CO_2 compression technology development
BMI Corporation	DD	Class 8 long-haul truck optimization for greater fuel efficiency
GE Global Research	DD	Unsteady Performance Predictions for Low Pressure Turbines
Caitin	DD	Parallel computing performance optimization for complex multiphase flows in cooling technologies
United Technologies Research Center	DD	Nanostructured catalyst for water-gas shift and biomass reforming hydrogen production
United Technologies Research Center	DD	Multiphase injection for jet engine combustors
GE Global Research	DD	Investigation of Newtonian and non-Newtonian Air-Blast Atomization Using OpenFoam
GE Global Research	ALCC	High fidelity simulations of gas turbine combustors for low emissions engines
United Technologies Research Center	DD	Surface Tension Predictions for fire-fighting foams
GE Global Research	DD	Engineered icephobic surfaces (INCITE Preparatory)
GE Global Research	DD	Engineered surfaces for water treatment (INCITE Preparatory)
Northrop Grumman	DD	Proof of Concept project to develop regional climate models, projections, and decision tools for local planners

Many of the industry projects complement DOE's strategic focus on addressing the nation's energy challenges. The cost and availability of energy, coupled with heightened environmental concerns, are causing companies to reexamine the design of products from large jet engines and industrial turbines to fire fighting foams. Their customers and the country are demanding products that have lower energy requirements and reduced environmental impact. However, the complexity of these design and analysis problems, coupled with the need for nearer term results, often requires access to computing capabilities that are far more advanced than those available in corporate computing centers. The OLCF is helping to address this gap by providing access to leadership systems and experts not available within the private sector.

For example, GE and United Technologies Research Center (UTRC) are both using Jaguar to tackle different problems related to jet engine efficiency. The impact of even a small change is enormous. A 1% reduction in specific fuel consumption can save $20B over the life of a fleet of airplanes (20,000 engines × 20-year life).

Access to Jaguar is allowing GE for the first time to study unsteady flows in the blade rows of turbomachines, such as the large diameter fans used in modern jet engines. Unsteady simulations are orders of magnitude more complex than simulations of steady flows, and GE was not able to attempt this on its in-house systems. By comparing its results to current steady flow solutions, GE will be able to determine whether unsteady flow analysis can lead to more energy efficient designs.

UTRC is using Jaguar to better understand the air-fuel interaction in combustors, a critical component of aircraft engines. They are validating first principles methods against experimental measurements, a first in this field given the complexity of the problem. Better understanding of the air-fuel interaction will enable UTRC to develop more efficient combustors that will reduce the emissions, lower the noise, and enhance the fuel efficiency of aircraft engines.

Caitin, a small engineering design firm in California, is developing a unique technology solution that could substantially reduce the energy required for cooling in applications ranging from general purpose refrigeration to data centers to chip level cooling. This firm just launched a project to use Jaguar to perform a full system analysis of the Caitin cooling system, simulating nonequilibrium multiphase critical flow. Evaluation of full system performance is simply not possible on Caitin's in-house system.

Access to Jaguar and OLCF experts is helping industry accelerate time-to-insight and time-to-solution for important energy-related problems with national impact. As industry delivers more energy efficient products, ORNL and DOE are delivering an additional return on the nation's investment in the OLCF.

5. FINANCIAL PERFORMANCE

CHARGE QUESTION 5: ***Are the costs for the upcoming year reasonable to achieve the needed performance?***

OLCF RESPONSE: The OLCF carefully managed costs in FY10 to execute the FY10 OLCF operational requirements and meet the targeted system availability and number of hours delivered. During the July 2011 Budget Deep Dive, the DOE program manager reviewed the proposed budget and concurred with the priorities reflected therein. In the August Lehman review, the OLCF presented the same DME project budget and enumerated how this fit into the overall operational budget.

2011 Operational Assessment Guidance – Financial Performance

The Facility presents financial performance information as follows:

- Presents a cost breakdown based on the budget taxonomy DOE created, which includes efforts, lease payments, operations (including DME, power cost, etc.,) and security;
- Compares current performance with a pre-established cost baseline;
- Explains variances between the baseline and actual and projected differences between current year and future year (FY11 to FY12);
- Identifies any entries that deviate from an established pattern with explanations for the deviations; and
- Explains any rebaselining that occurred during the year and reasons.

2011 Approved OLCF Metrics – Financial Performance

Financial Performance: **The OLCF will report on budget performance against the previous year's budge deep dive projections.**

The projected total OLCF cost for FY11 is $85,180K. Of this 28% is spent on effort, 36% on lease payments, 11.6% on space and utilities, 8.3% on computer system maintenance, and 16.1% on other costs. The OLCF carefully managed costs in fiscal year (FY) 2011 to accommodate a lengthy continuing resolution (CR) and to execute the FY11 operational requirements and meet the targeted system availability and number of hours delivered. The final FY11 budget and funding was not settled until June. As a result of these delays, the OLCF presented revised budgets to ASCR in December 2010 and June 2011. The December revised budget cut the FY11 budget from $96M to $87M with a full year continuing resolution. The June budget revised the spending plan based on the appropriated $96M budget and the revised spending plan for the OLCF-3 project.

The OLCF budget includes both steady state operations and the OLCF-3 upgrade project. The Development, Modernization, and Enhancement (DME) portion of the budget includes project costs related to upgrading the existing Jaguar system. This upgrade will be executed in two phases. The first phase, early in FY12, will be a processor, memory, and interconnect upgrade. The second phase, early in FY13, will add 10 to 20 petaflops of accelerators to the system. The DME work in FY11 includes project planning, system acquisition, application and tools readiness and site preparation activities. After the system acceptance in each project phase, the cost related to the system is included in the operational portion of the OLCF budget. The OLCF tracks all costs against the yearly budget in functional categories (leases, utilities, etc.) and cost types (labor, subcontracts, etc.) and by DME and operations. This allows the OLCF to monitor costs against planned budgets in numerous important ways. The OLCF is aided in this ability by a powerful SAP financial system that can provide information from the time-reporting system and the procurement system. The financial status of the OLCF is monitored daily by the OLCF finance officer and at least monthly by OLCF management. The OLCF management is experienced in mitigating potential budget impact from delays in Congressional passing of funding bills (Reference Section 7, Risk Management). The budget presented here is based on the assumption of a continuing resolution of up to 6 months and includes a carryover of $18M from FY11 to FY12 to help manage cost and cash flow.

The planned OLCF budget for FY11 (President's Budget) was $96M and full funding at this level was received in late June. See Table 5.1 for the FY11 funding and cost. Because of the extended CR and overall budget uncertainty during the majority of the year, the OLCF spent very conservatively before the funding level was resolved and therefore experienced variances in several cost categories. The current performance is compared to the pre-established cost baseline in Table 5.2.

The DME budget was a placeholder for the OLCF-3 project in the pre-established budget and was replaced by the proposed OLCF-3 project baseline budget that will be reviewed as part of the Office of Science CD-2 review in August 2011. The actual cost aligns with this new proposed cost baseline.

Actual effort costs were less than budgeted because the OLCF experienced the loss of several staff members (Kothe, Carpenter, Rosinski, Barrett, Henley, Frederick, Buchanan and Zhang) during FY11.

Table 5.1 OLCF FY11 funding and cost table

Category		Subcategory			$K
	Budget				$96.000
	Carry-in				$7.795
	Total Budget				$103.795
1	Effort				
		1.1	DME		$3.310
		1.2	Steady State		$20.572
2	Leases				
		2.1	Advance Payments		
		2.2	Leases	(Lease payments, financial charges, TN tax and OH)	$31.000
3	Security				
4	Operations				
		4.1	DME (excluding effort)		$0.399
		4.2	Subcontractors/Students		$2.707
		4.3	Maintenance		$7.074
		4.4	Center Balance	(Local storage, Networking, Infrastructure, Visualization, Testbeds, Software development, Software licenses)	$4.729
		4.5	Other Major HW	(HPSS, End to end)	$4.195
		4.6	Other	(Travel, Training, User meeting, Workshops, Department materials, Outreach materials)	$1.341
		4.7	Center Charges	Computer Center Operators	$0.400
				Computer Center Improvements	$0.710
				Space Cost	$0.590
				Utilities (power, cooling)	$8.151
	Total				$85.180
	Carry-out				$18.615

During the CR, hiring for these open positions and other planned staffing was slowed until June when the full year funding became known and available.

Subcontracts/Student costs were less than budgeted because the support for Lustre was achieved in a new, less costly way and a management advisory subcontract was not yet required.

Maintenance and Center charges (utilities) costs were less than budgeted because the XT4 system was decommissioned in February. The decommissioning was part of the conservative spending strategy enacted, in part, because of the funding uncertainty during the fiscal year. Additionally, Adaptive Computing/MOAB maintenance, originally budgeted for FY11, was prepaid with FY10 funds made available late in FY10.

Center Balance (Cybersecurity, local storage, networking, infrastructure, visualization, testbeds, software development, software licenses) costs were less than budgeted because network operations/infrastructure budgeted for a new computing facility were not required. Additionally, some budgeted testbed expenses were reduced.

Other major hardware costs were greater than planned because OLCF invested in additional HPSS tapes and in tape cleaning to support the growth requirements of the archival storage system.

The FY12 target budget includes $95M of new budget authority (BA). The FY12 baseline budget includes $88M of new BA. Under either budget scenario for FY12 the budget will be identical with the exception of the investment in the file system/storage. Depending of the actual funds received the OLCF will adjust the strategy for acquiring this equipment. With the target budget, the file system/storage will be purchased during FY12 and 13. With the baseline budget, the file system may need to be leased or acquired through a combination of purchase and lease. The option to lease the file system is not preferred and would cost more because of the fees associated with a loan agreement. The target and baseline FY12 budget scenarios are shown in Table 5.2.

There are several areas where the FY12 budget deviates from the previous year budget. These are identified below.

Because a portion of the OLCF budget is allocated to the DME project, the budget for operations must be adjusted for DME expenses which fluctuate from year to year depending of the schedule of project activities and their anticipated costs derived from the OLCF-3 project controls system. In FY12 the operations budget must accommodate a DME budget of $11.3M which is significantly more that in FY11.

The operations effort budget has been adjusted for current FY12 planning salary rates and FTE levels.

The Maintenance budget no longer includes maintenance for the XT4 system and is adjusted for the upgrade of the XT5.

The Center balance budget for FY12 does not include expenses for upgrading the visualization equipment which was done in FY11. Additionally, the networking budget will be lower because network investments made in prior years are not needed again in FY12.

The budget for other major hardware will increase to accommodate the new file system and disk storage purchase or lease as well as the continued growth in HPSS.

The FY12 budget will include the final payment on the XT5 lease and the beginning of the lease stream for phase one of the system upgrade. The new lease will require the upfront payment of a loan origination fee as well as the appropriate Tennessee use tax.

The FY12 Center charges budget has been adjusted to reflect the utilities associated with the XT5 system as it is currently configured as well as the upgraded system. The XT4 system utility costs have been removed from the FY12 budget.

The OLCF budgets for FY11 through FY16 have been reestimated to reflect the new plan for the OLCF-3 project. The original plan included the purchase of a new computer, the build out of a new facility, and the overlap of providing two systems for a year while transitioning users to the new system architecture. The new plan for OLCF-3 is significantly different as it only includes a two-phase upgrade to the existing XT5 system in the existing computer facility. The new plan reduces the planned costs for site preparation and the utilities associated with operating two systems for a year, but it does cause some system downtime while the upgrades are taking place

Table 5.2 OLCF FY11 Budget vs Actual Cost

	Carry-in from FY10	DME	OPS Effort	Subcontracts/ Students	Maintenance	Center Balance	Other HW	Leases	Other	Center Charges (Utilities)	Mgmt Reserve	Carry-out	Total
Budget	7.8	6.8	24.7	3.1	9	6.1	2.8	31	1.3	10.7	0.5		103.7
Actual		3.7	20.6	2.7	7	4.7	4.2	31	1.3	9.9	0	18.6	103.7

Table 5.3 OLCF FY12 Target and Baseline Budgets

		DME	OPS Effort	Subcontracts/ Students	Maintenance	Center Balance	Other HW	Leases	Other	Center Charges (Utilities)	Mgmt Reserve	Carry-out	Total
Target	$95M Budget	11.3	21.8	2	6.9	2.9	14.3	35.8	1.7	6.5	1	9.4	113.6
Baseline	$88M Budget	11.3	21.8	2	6.9	2.9	9.6	35.8	1.7	6.5	1	7.1	106.6

6. INNOVATION

CHARGE QUESTION 6: ***What innovations have been implemented that have improved the facility's operations?***

OLCF RESPONSE: The OLCF actively engages in innovation activities that enhance facility operations. Through collaborations with users, other facilities, and vendors, many of these innovations are disseminated and adopted across the country.

2011 Operational Assessment Guidance

The Facility highlights innovations that have improved its operations, focusing especially on best practices:

- that have been adopted from other Facilities,
- those the Facility has recommended to other Facilities, and those other Facilities have adopted.

2011 Approved OLCF Metrics – Innovation

Innovation Metric 1: ***The OLCF will report on new technologies that we have developed and best practices we have implemented and shared.***

The OLCF has carried out numerous activities ranging from working with users to update their applications to maximize their effective use of anticipated systems, to technology innovations that streamline workflow, to tools development. See additional comments for Innovation Metric 2.

Innovation Metric 2: ***The OLCF will report on technologies we have developed that have been adopted by other centers or industry.***

The OLCF has developed a number of technical innovations that have been adopted by other centers and industry. Our work on exploiting hierarchical parallelism within applications to better map to next-generation architectures is being adopted by the communities who developed these applications. To this end, the OLCF established the Center for Accelerated Application Readiness (CAAR). A guiding principle of this effort has been to directly integrate these capabilities into the canonical source tree of each application thereby easing longer-term maintenance of the application and portability of these enhancements. The OLCF's work in topology aware I/O, specifically our topology aware Lustre network routing capabilities have been incorporated into the canonical Lustre source tree and the knowledge required to make use of these capabilities have been disseminated through a number of publications and presentations by OLCF staff. Our work on the Common Communication Interface (CCI) is a collaborative development effort conducted in concert with other laboratories (SNL, INRIA) and industry (Cisco, Myricom). The OLCF has funded and managed contract development of scalable and heterogeneous debugging features that have been incorporated into the Allinea DDT debugging tool. To improve code portability and ease porting to advanced architectures the OLCF has funded and managed contract development of accelerator enhancements in the CAPS HMPP compiler, a commercially available product. Finally, the OLCF has funded and managed contract development of scalable performance analysis for heterogeneous systems in the widely used Vampir tool set allowing these capabilities to be utilized by HPC centers around the world. Through direct engagement with other HPC centers, vendor partners, and application development teams, the OLCF is ensuring that ASCR investments that culminate in technical innovations have broad impact to the entire HPC ecosystem.

Innovation is the heart of HPC. Innovation not just in the science enabled by the computing power inherent in high-performance computers, but in HPC itself. The increasing complexity of the world we live in is making innovation increasingly a matter of careful, long-range planning.[1] OLCF activities this past year reflect this, with staff members across the organization contributing to planning for the next generation of HPC. Judging by the results, the OLCF will be more than ready to take advantage of the technological breakthroughs looming with the advent of such leading edge technologies as multithreaded parallelism, general purpose GPUs, and multicore-aware software. The following pages describe some of these exciting new developments, pioneered and led by OLCF staff.

6.1 THE ACCELERATOR CHALLENGE

In 2012 the OLCF will deploy a large-scale, hybrid- multicore node-based system known as Titan for use as a major compute resource for DOE SC. The nodes on this system will have an industry standard x86-64 architecture processor paired with a GPU-based application accelerator. The resulting node will provide a peak performance of more than 1 teraflop.

The new hybrid node architecture will require application teams to modify their codes to take advantage of the accelerator. Given the marked difference in node architecture, substantial effort will be needed to bring scientific applications to the point of effective use of the new platform. The primary challenges involved in marshaling the GPUs are threefold:

- recognition and exploitation of hierarchical parallelism by scientific applications, including distributed memory parallelism via message passing interface (MPI), symmetric multiprocessing (SMP)-like parallelism via threads (OpenMP or pthreads), and vector parallelism via the GPU programming;
- development of effective programming tools to facilitate this (often) substantial rewrite of the application codes; and
- deployment of useful performance and debugging tools to speed this refactoring.

To lead the way, in 2010 the OLCF established the Center for Accelerated Application Readiness (CAAR), whose members include application teams, vendor partners, and tool developers. CAAR is charged with preparing six representative applications for Titan. The six applications, selected from among 50 of the most productive applications running on Jaguar, were chosen because they represent much of the range of demands that will be placed on Titan from a variety of scientific domains. application and Software development leadership

---

[1]Dosi, G., “Technological paradigms and technological trajectories,” *Research Policy,* 11 (1982), pp. 157–162.

Each of the CAAR teams is led by an OLCF staff member from the Scientific Computing Group. The teams also include representatives from the individual code development groups, engineers from OLCF vendor partners Cray and NVIDIA, and, in some cases, other OLCF and ORNL staff members. The SciComp CAAR team leaders are responsible for coordinating the work of their teams and have shared responsibility with the code owners in formulating the science targets for OLCF-3. One of the most important responsibilities of the CAAR team leads is to ensure that changes made to facilitate the port to OLCF-3 are retained in the production trunk of each code. This vital step helps assure portable performance, as changes made that increase data locality and expose hierarchical parallelism prove useful even on non-hybrid architectures.

The totality of each CAAR code port experience, like much of the work the SciComp liaisons produce in support of production work on Jaguar, will be transmitted to the wider community through several means, including dissemination of best practices and the availability of production software packages and libraries (e.g., the Multi-level Summation Method kernel from the CAAR code LAMMPS will be made available as a library to other MD practitioners). The CAAR experiences and lessons-learned will lead to the most complete and sustainable set of practices available for hybrid multicore computing for the near future.

Researchers Gather at ORNL to Explore Petascale While Looking to Exascale Future

About 70 researchers working on some of the nation's most pressing scientific missions gathered at ORNL for the Scientific Applications (SciApps) Conference and Workshop August 3–6, 2010. An interdisciplinary team of computational scientists shared experience, best practices, and knowledge about how to sustain large-scale applications on leading HPC systems while looking toward building a foundation for exascale research.

SciApps 2010 was funded by the American Recovery and Reinvestment Act. The OLCF Scientific Computing Group leader, Ricky Kendall, and then OLCF director of science, Doug Kothe, cohosted the conference. "While many of the scientific disciplines have little in common, there is a tremendous algorithmic commonality among some of them, and they all share a need for ever expanding computational resources to help them meet their scientific goals and missions," Kendall said. "One finding was that all disciplines represented at the meeting had a strong use case for sustained petascale computing and many had well-thought-out ideas about the next steps towards exascale computing."

LBNL and ORNL Organize First SciDAC Software Workshop for Industry

About 60 software experts gathered in Chicago on March 31, 2011, for the first Workshop for Independent Software Developers and Industry Partners, sponsored by the DOE Advanced Scientific Computing Research office. Jointly organized by Lawrence Berkeley and Oak Ridge National Laboratories, this workshop introduced independent software vendors (ISVs) and industrial software developers to software resources that can help ease the private sector's transition to multicore computer systems. These tools, libraries, and applications were developed through DOE's Scientific Discovery through Advanced Computing (SciDAC) program to enable DOE's own critical codes to run in a multicore environment.

The cost and difficulty of scalably parallelizing legacy codes (codes written for nonoperational or outdated operating systems or computer technologies) often are prohibitive to independent software vendors, particularly if they are small businesses. They also hamper many firms that, for proprietary and competitiveness reasons, maintain their own code in addition to using commercial options. The problem is becoming acute as desktop workstations and small clusters are rapidly being designed and built using multicore processors.

The 1-day workshop was an important contribution to addressing these hurdles. It gave participants an overview of the SciDAC program and more than 60 SciDAC-developed software packages and outlined the process to obtain them, often at no cost. In addition, DOE explained its role in providing research grants through the U.S. Small Business Administration's Small Business Innovation and Research (SBIR) grant program. This program ensures that the nation's small, high-tech, innovative businesses are a significant part of the federal government's research and development efforts. Workshop participants then provided feedback on private sector software development requirements that could help DOE shape future SBIR research topics and jumpstart areas for collaboration.

"SciDAC has spent a decade developing world class software to ensure DOE can operate successfully in a multicore environment," explained David Skinner, workshop cochair and director of the SciDAC Outreach Center at Lawrence Berkeley. "The private sector software developers who participated now have direct links to key developers who can provide expertise in developing software for multicore systems and help guide integration of SciDAC software into commercial applications. We hope to extend these links to those who could not attend."

The workshop's participants represented 49 organizations, including small and large ISVs, companies with internal software development capabilities, academic institutions, other national laboratories, and HPC system vendors.

"This event launched a new opportunity to leverage DOE's investment in SciDAC for an additional return on investment for the country," said fellow chair Suzy Tichenor, director for the HPC Industrial Partnerships Program at Oak Ridge. "Most of the ISVs and companies that attended had never heard of the SciDAC program. Now they are aware of SciDAC's valuable software resources and how to access them."

6.2 CENTER TECHNOLOGY INNOVATIONS

Flash Storage Technologies

Solid-state disks (SSDs) offer significant performance improvements over hard disk drives on a number of levels. However, SSDs can exhibit significant performance degradations when garbage collection (GC) conflicts with processing the request stream. The frequency of GC activity is directly correlated with the pattern, frequency, and volume of write requests, and scheduling of GC is controlled by logic internal to the SSD.

When using SSDs in a redundant array of independent disks (RAID),[1] the lack of coordination of the local GC processes amplifies these performance degradations. No RAID controller or SSD available today has technology to overcome this limitation.

OLCF has proposed a new technology, global garbage collection (GGC), to address these problems and enhance both storage and retrieval performance in existing computer systems for SSDs in RAID configurations. This new technology functions on both servers and mass consumer computers. The OLCF technology uses SSDs in a coordinated RAID configuration.

The invention includes a high-level design for an SSD-aware RAID controller and GGC-capable SSD devices and algorithms to coordinate the global GC cycles. An optimized redundant array of solid-state devices includes an array of one or more optimized solid-state devices and a controller coupled to the solid-state devices for managing the solid-state devices. The controller can be configured to globally

[1]RAID is an umbrella term for computer data storage that can divide and replicate data among multiple disk drives. Data are stored across all disks in such a way that if a single drive fails, the data can be retrieved and reconstructed by the remaining disks.

coordinate the GC activities of each of the optimized solid-state devices (e.g., to minimize the degraded performance time and increase the optimal performance time of the entire array of devices). The controller can also schedule and perform a globally coordinated memory scan over all disks in a given RAID—reclaiming space when possible. In addition, the controller can arrange the GC in an active mode so that collection cycles begin on all disks in the array at a scheduled time or it can query the disks to determine the best time to start a global collection.

Simulations have shown that this design improves response time and reduces performance variability for a wide variety of enterprise workloads. For "bursty," write dominant workloads, response time was improved by 69% while performance variability was reduced by 71%.

A patent application for this invention, titled "Coordinated Garbage Collection for RAID Array of Solid State Disks," was filed with the U.S. Patent Office on August 5, 2010. The Patent Application Number is 61,370,908. The inventors are David A. Dillow, Youngjae Kim, H. Sarp Oral, Galen M. Shipman, and Feiyi Wang.

Spider and Topology Aware I/O

While computation is the heart and soul of a scientific application, there are many I/O tasks required to make that computation feasible.

Applications must read in their input decks, write out their results, and perform defensive I/O to protect against machine faults. Time spent performing these operations represents time that could be used to improve the resolution of the science or give a reduction in time-to-answer, further improving productivity. In support of this goal in 2011, the user-achievable bandwidth on Spider was more than doubled. This was accomplished without purchasing any additional hardware by carefully considered configuration changes.

Spider is a "routed" file system, which means that it uses I/O nodes on the Jaguar system to move information between two physically incompatible interconnect topologies; in this case, the Cray SeaStar network on Jaguar and the 20 Gbps InfiniBand on Spider. Because Spider offers aggregate bandwidth far in excess of the single-link speeds of either interconnect, avoiding congestion is fundamental to achieving efficient I/O. Unfortunately, simple configurations of Lustre at large scale inherently induce congestion in the InfiniBand fabric. By default, Lustre disperses traffic to all routers in a round-robin fashion. This causes traffic to be injected into the InfiniBand fabric's fat-tree topology in nonoptimal locations, which in turn causes oversubscription and congestion on internal links of the fabric. Significant performance degradation due to this issue has been measured. Additionally, this dispersal of traffic to the routers prevents using locality information to optimize application I/O performance, as it is impossible to know which router will service each request.

The OLCF has completely eliminated congestion inside the InfiniBand fabric by pairing routers with individual Spider servers. This one-to-one mapping keeps traffic inside the crossbar switch and prevents it from traversing the internal links of the fat-tree. In addition, traffic for a given server takes a more direct route within the torus. This configuration change improved demonstrated read bandwidth by 101% and gave a 93% improvement for write bandwidth for applications without regard to their locality. For tests in which the I/O targets were chosen based upon location in the torus, the new routing configuration allows improvements of up to 115% for reads and 137% for writes.

This information was shared with the larger user community during the 2011 Cray User Group meeting and is available as an ORNL technical report via http://info.ornl.gov/sites/publications/Files/Pub30140.pdf.

I/O Management and Tools

Part of the work of any HPC facility is improving its core competencies in the operational management of large-scale file systems, including developing improved tools to manage the file systems. Day-to-day operations such as generating candidates for purging or maintaining server balance often involve querying the file system metadata. Additionally, there is an occasional need to determine the file name affiliated with an error message or a set of files impacted by an outage. As file systems age and more files are added, the amount of time such management tasks take increases in proportion to the number of files in the system. The OLCF has had more than 445 million files in Spider during times of peak usage and currently contains about 210 million files.

Operations at this scale take many hours and in some cases many days. For example, generating the candidate list for purging takes between 6 and 21 hours on Spider, depending on the I/O load of the running science applications, the number of files stored, and the past peak usage. The vendor recommended methods for determining the files associated with a given storage target take more than 5 days when run to completion, and even recent tools required more than 2 hours to associate an error message with a file name.

The OLCF has developed tools to reduce these times in order to increase management productivity and to improve responsiveness in the event of an unplanned outage. With the improved I/O patterns of these new tools, the time to generate a purge candidate list has been reduced to about an hour on Spider. Other management tasks requiring a full scan of the file system metadata now take similar times. Determining which files are potentially impacted by an outage, for example, now takes less than 1 hour, which is a substantial improvement over the 5 days required by first generation tools. The file associated with an error message can now be named in less than 15 minutes, compared to the hours it would require without the OLCF tools. These enhanced tools have led to greater responsiveness and user peace of mind when dealing with outages, planned or not. Over the next few months the OLCF will be packaging these tools for distribution to the broader HPC community.

Data Management for Climate Science

The Earth System Grid Federation (ESGF) is a large-scale, multi-institution, interdisciplinary project to provide climate scientists worldwide, as well as climate impact policy makers, a web-based platform to publish, disseminate, compare, and analyze ever increasing amounts of climate-related data. ORNL is a key contributor to the ESGF project with development and data publication efforts funded by the DOE Office of Science - Biological and Environmental Research. While BER funds the development and software maintenance of ESGF at ORNL, the OLCF has assisted in the architecture and deployment of the system infrastructure required to provide climate scientists with access to the high-value datasets resident within the OLCF. This involved a hardware and software setup consisting of the following:

- two data nodes for end users to publish their data sets,
- one production gateway running the latest Gateway portal software, and
- one 250 TB storage backend for on-disk data access.

The HPSS deployed at OLCF is capable of storing multi-petabytes of data for long-term archival purposes; the Earth System Grid's (ESG's) current online disk capacity is limited in comparison. Project participants determined, therefore, that it would further ESGF goals for climate scientists to be able to access data stored in HPSS via the ESG. The basic problem was one of security: ESG is publicly accessible while HPSS has security restrictions. Only a small amount of the data in HPSS, that pertaining to the ESG program, should be accessible from ESG, so the issue devolved to one of ensuring that the rest of the data in HPSS would not be inadvertently compromised. To do this, ORNL designed and implemented an ESG HPSS access framework (Figure 6.1), which leveraged the OLCF infrastructure.

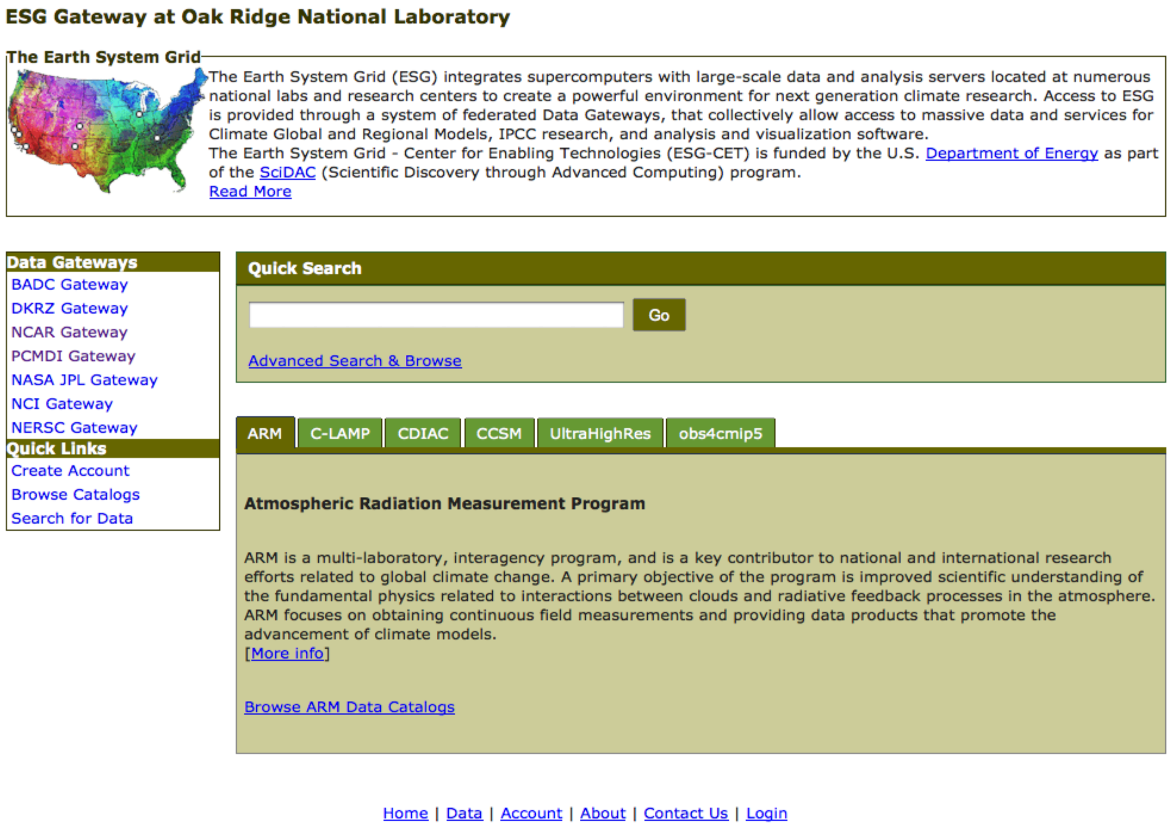

Figure 6.1 ORNL Secure ESG Gateway.

As a result of this work, the ORNL-ESG system hosted within the OLCF provides access to a number of high use, high value data sets, including the following.

- Climate Modeling Best Estimate atmospheric, cloud, and radiation quantities showcase data sets from the Atmospheric Radiation Measurement Program
- Carbon-Land Model Intercomparison Project data set
- Ameriflux (part of the FLUXNET global network of towers making continuous measurements of CO_2, water vapor, and radiation via eddy covariance in terrestrial ecosystems) and Fossil Fuel (gridded fossil-fuel CO_2 emission estimates) data from the Carbon Dioxide Information Analysis Center data set
- The Ultra High Resolution Global Climate Simulation project

The availability of these datasets on the ORNL-ESG system provides climate scientists with direct access to high-value simulation results and observations. Further integration of ESG within our operational environment will provide remote analysis and data-subsetting, much needed capabilities when working with geographically distributed, multi-terabyte datasets.

Open Scalable File Systems, Inc.

The Lustre parallel file system is the most used parallel file system technology in HPC, with use on more than 70 of the top 100 HPC systems and all of the top 5 systems in the November 2010 Top500 list. As the only open-source, vendor-neutral parallel file system capable of supporting leadership-class HPC systems, the Lustre file system is a critical technology used across DOE sites. Originally developed under the auspices of the DOE National Nuclear Security Administration path-forward effort by Cluster File Systems, Inc., the Lustre file system is now broadly supported by a variety of system integrators and storage system vendors. Because of the breadth of Lustre use in HPC and the criticality of this technology to the marketplace, in 2010 the OLCF teamed with Cray, DDN, and LLNL to form Open Scalable File Systems, Inc. (OpenSFS), a nonprofit mutual benefit corporation for development of high-end open-source file system technologies, with a focus on the Lustre parallel file system. OpenSFS is specifically geared to meet the needs of the Lustre community by providing a forum for collaboration among entities deploying file systems on leading edge HPC systems, communicating future requirements to developers, and supporting a development of advanced features designed to meet these goals. OpenSFS supports the Lustre community by holding annual scalable file systems workshops and providing a variety of services such as education and community outreach, testing, documentation, and project management.

OpenSFS is now embarking on the development of next-generation features within the Lustre file system, allowing the OLCF to meet its current and future HPE requirements. Whereas in the past this development would require direct funding solely by the OLCF or would rely upon development activities funded by other organizations but with no direct oversight by the OLCF, the OpenSFS model allows the OLCF to leverage others' investment in the Lustre file system while preserving its ability to oversee collaborative development efforts. Having released a request for proposal in April 2011, OpenSFS is now in contract negotiations to develop a variety of features in the Lustre file system aimed at meeting member operational requirements.

The OLCF's leadership role in OpenSFS has resulted in a single Lustre community represented by OpenSFS and the European Open File System consortium (EOFS). This collaboration, the first of its kind in the HPC world, was announced at the first Lustre User Group Meeting (organized by the OLCF) and ratified through a memorandum of understanding between OpenSFS and EOFS signed at this year's International Supercomputing Conference (June 19–23, 2011, Hamburg, Germany). OLCF leadership fostered this collaborative approach to continued Lustre development and thus ensured the future of the Lustre file system.

Common Communication Interface

The sheer size of the OLCF imposes scalability issues for everything from storage to debugging tools. In addition to Jaguar, the OLCF includes many different types of hardware including multiple types of network infrastructures. Each network provides at least two application programming interfaces (APIs); BSD sockets; and the network's native interface, which provides better performance through direct access to the network hardware. Jaguar, for example, provides Portals while the storage system uses Verbs. Cray's next generation of hardware replaces SeaStar with Gemini.

Applications must be ported (i.e., modified) to use each network's native API to obtain the best performance (i.e., lowest latency, highest throughput, and lowest CPU utilization), and various groups within the OLCF port applications for each new generation of hardware.

The Technology Integration Group (TechInt) is working on a new programming interface that will provide a common API for applications, allowing them to take advantage of current networking hardware and next generation hardware as it is acquired. This new API, known as the Common Communication

Interface (CCI), is being jointly developed by ORNL, SNL, University of Tennessee, Myricom, and Cisco.

CCI is designed for portability, scalability, and performance. For portability, CCI provides a simple interface that is similar to BSD Sockets yet provides remote memory access if the hardware supports it. CCI achieves scalability by bounding memory usage per communication end point (e.g., application) rather than per communication peer. CCI delivers performance via access to the underlying hardware capabilities such as OS bypass, zero copy, and remote memory access.

TechInt is working on refining the API, with support for Portals nearly complete. Initial testing on Jaguar shows that CCI adds just 200 nanoseconds of overhead (about 3%) to small messages. For large transfers, the overhead is less than 1% (nearly unmeasurable). A version for BSD sockets for general testing and to support other networks until the native versions are ready is in progress, and TechInt will soon begin work on CCI over Verbs and Gemini.

The software is expected to be ready for adoption soon. Once CCI is released, TechInt will work with application developers and maintainers to add support for it.

OLCF HPSS Development Activities

HPSS was created more than 15 years ago by a collaboration of IBM and five DOE laboratories: LANL, LLNL, LBNL, ORNL, and SNL. At that time it was recognized that no single laboratory or corporation had the expertise or resources to create the product alone. HPSS continues to depend upon and to grow from the joint contributions of all collaboration members.

Over the past year, OLCF HPSS developers have contributed to several parallel development efforts: release 7.4, RAIT, and release 8.1.

HPSS version 7.4 development was completed this year. The integration tests are now being upgraded and integration and system tests will follow, with a target release date of January 2012. The new version adds the following features.

- Dynamic drive updates. This builds upon the dynamic drive add and delete functionality which was first provided in HPSS 7.1. Device configurations can now be updated without system downtime.
- HPSS High Availability on Linux.
- Repack enhancements. The repack utility copies data from old volumes to new ones so that sparse volumes can essentially be defragmented and outdated technology can be replaced. Version 7.4 is capable of repacking old nonaggregated tapes, where files are stored individually, into tape aggregates on the new volumes.
- hpssadm enhancements. hpssadm is the command line interface to SSM. In 7.4 it was extended to provide complete HPSS configuration capability. Lengthy system configuration changes can now be automated in a batch script, reducing downtime. A complete system can now be configured from a script, enabling quick set up of new test systems or of production systems at new sites.
- Logging enhancements. Logfiles were changed from binary to text format, a tremendous boon to real time debugging. Log archiving was improved to be more flexible and to avoid potential loss of logging data during times of high activity; previous systems could lose some log data when a log file could not be archived quickly enough.

ORNL has primary responsibility for the development of a number of important subsystems of HPSS: the storage system manager (SSM), the graphical and command line interface for monitoring, configuring, and controlling the system; the bitfile server (BFS), one-third of the core server; the logging subsystem; and the accounting subsystem.

OLCF HPSS developers contributed the necessary SSM modifications to support all of these innovations, particularly dynamic drive updates, and were fully responsible for the logging and hpssadm features.

The collaboration is in the process of developing an implementation of RAIT, redundant array of individual tape. This is targeted for a release sometime after 7.4 or 7.5. OLCF HPSS developers have made contributions to RAIT in the areas of logging, SSM, and BFS.

The OLCF HPSS developers are continuing to work with other collaborators on the design and development of HPSS version 8.1.

6.3 TOOLS DEVELOPMENT

Debugging: Allinea DDT

A scalable, hybrid, platform-aware debugger is an essential component for the programming environment (PE) of OLCF-3 to work well on a massive, hybrid, GPU-based cluster system. OLCF is working with Allinea to make their debugger, Distributed Debugging Tool (DDT), scale to more than 200,000 cores and handle the debugging of GPU data.

The Allinea collaboration allows the OLCF to address the requirements of the OLCF-3 GPU-based architecture by using sophisticated tree topology and tight integration with Cray’s advanced PE features such as scalable breakpoints, stepping and program stack queries, scalable process management, scalable visualization of variable values using statistical analysis and prefetching techniques, distributed core file generation with abnormal process termination, and Cray’s process launcher. All of these DDT capabilities provide the basic building blocks for creating an efficient debugger for the OLCF-3 PE. Figure 6.2 shows the time that it takes to set a breakpoint or step over program statements during a debugging session on up to 200,000 MPI processes. The figure clearly shows that the debugger is scalable.

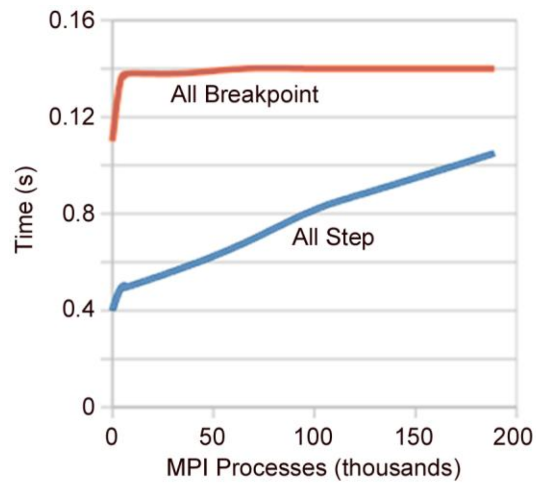

Figure 6.2. DDT scalable breakpoints. DDT scalable breakpoints and stepping for large MPI process counts in Jaguar XT5.

In addition, Allinea has enhanced its existing DDT debugger capabilities to support CUDA and the hybrid multicore parallel programming (HMPP) compiler. The current implementation supports stepping over CUDA kernels and automatic detection of HMPP fragments, step over HMPP codelets, and report error codes from the HMPP run time. Figure 6.3 shows setting a breakpoint in an HMPP region directive in one of the Community Atmosphere Model (CAM)–spectral element (SE) kernels. The DDT debugger is able to recognize the HMPP directives and step over them correctly.

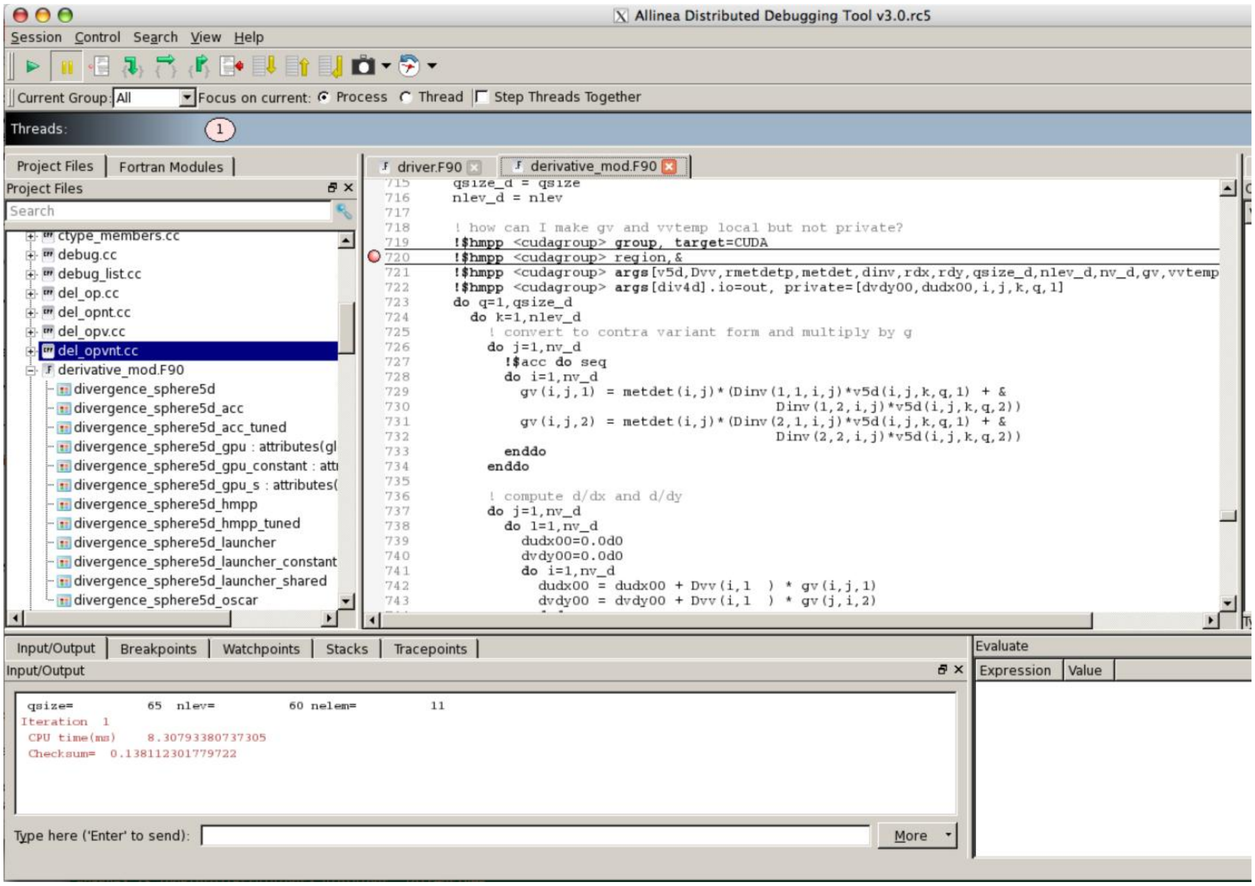

Figure 6.3 The DDT debugger applied to the HMPP codelet.

Compiling: CAPS HMPP

Applications of interest to OLCF-3 are written in C/C++ and Fortran 77/90, with MPI; OpenMP; and, in some cases, DSL. To improve user code porting and development productivity, the OLCF-3 project will support the use of high-level languages with accelerator directives. The Center is exploring the use of Cray, PGI, and HMPP accelerator directives and has initial performance assessments on kernels written in C and Fortran, which requires minor modification to the original source code and can be retargeted to different platforms. As part of this process, the Applications Performance Tools group is working with CAPS enterprise (www. caps-enterprise.com) to come up with a set of directive requirements to port OLCF-3 applications to the new system.

Copying data in and out of accelerator devices is a time-consuming process, as the data do not always have a flat layout (e.g., an array of primitive data types). As part of the OLCF-3 effort, HMPP has been extended to support user-defined data types and data structures holding pointer fields; OLCF applications such as CAM-SE rely on user-defined data types to store the cubed elements information. With the introduction of dynamic CPU/GPU coherency management, OLCF users are relieved from manually mirroring host/device images of data structures upon modification. Requesting coherency maintenance through a directive as opposed to implementing it by hand reduces code size greatly, is type agnostic, and raises programming productivity.

Users often need to contrast the performance of or incorporate hand-tuned, compiler-generated, and external (e.g., library-provided) kernels to their code using directives. The implementation of User-Kernel

Integration instructs HMPP to bypass its own code generation and utilize user-supplied code directly, and thus, it achieves the desired effect. The TechInt LSMS team is in the process of modifying the LSMS application so that it can make use of CULA, a GPU-accelerted linear algebra library. The CAPS partnership has also led to the formation of HMPP++. HMPP++ bridges HMPP and object-oriented programming by allowing application C++ classes to inherit from the HMPP run time's classes while fully utilizing the HMPP directives (extended to by C++ scope-aware, etc.); this hybrid model has been tested successfully in the context of the Multiresolution Adaptive Numerical Environment for Scientific Simulation (MADNESS) application.

Data staging is not always a single copy operation; data may need certain accelerator-specific processing such as transferring them to the device, reformatting them while on the device, and placing them in shared memory. HMPP's CUDA-specific direct shared memory operations achieve this. The staging process is also affected by the affinity of data. Certain enhancements to the data residency qualifiers have helped with data structures that are only "live" on the GPU. Host-device data transfers can be expensive and advantage needs to be taken of the nonblocking data-transfer opportunities next to the transfers' planning and strategic placement. Improvements against the HMPP asynchronous I/O mechanism combined with the mechanism's type-awareness have simplified these tasks.

Performance Analysis: Vampir

The Vampir (Visualization and Analysis of MPI Resources) tool set is used for performance analysis in OLCF-3. We are working together with Vampir's vendor, the Technical University of Dresden, to make this tool set ready for the targeted OLCF-3 system. Vampir uses program tracing to record a detailed list of events during the execution of an application. Using a set of compiler wrappers for C, C++, and FORTRAN, the application can be built with specific instrumentations.

VampirTrace provides instrumentation of the parallel paradigms MPI and OpenMP/Threads, as well as generic recording of function invocations through compiler or manual instrumentation. Vampir then provides a postmortem visualization of the program execution based on the recorded trace. This visualization features a set of different displays to help understand the behavior of the application. The analysis for visualization is provided by a parallel server and a GUI application, allowing the processing of large traces. The entire tool chain is tailored for a scalable parallel analysis. To match the scale of the target OLCF-3 system, additional improvements have been and are being incorporated in Vampir. Specific optimizations in the communication behavior of VampirServer now enable the use of more than 10,000 analysis processes. Multiple improvements target the handling of an increasing amount of trace data from hundreds of thousands of processes. Pattern matching–based compression will improve the recording, while filtering and the highlighting of irregularities will support the evaluation of large-scale traces.

The other important contribution is the integrated CUDA support in VampirTrace. CUDA-API calls are captured and recorded. GPU events such as kernel execution and memory copies are mapped to CUDA streams. Those events can be invoked asynchronously and are correctly embedded into the timeline of traditional program events. The support for GPU performance counters adds information to the trace. This integrated approach allows analyzing hybrid MPI/OpenMP/CUDA applications as a whole and provides a better picture of the application's performance characteristics than just looking at isolated CUDA kernels. Figure 6.4 displays a timeline of four MPI processes, each with an associated CUDA stream that runs the GPU-accelerated version of LAMMPS. With these improvements, Vampir provides a comprehensive performance analysis tool for the upcoming OLCF-3 system. It helps application developers to port and adapt their codes to this system and therefore increases its utilization and facilitates the solution of new scientific problems.

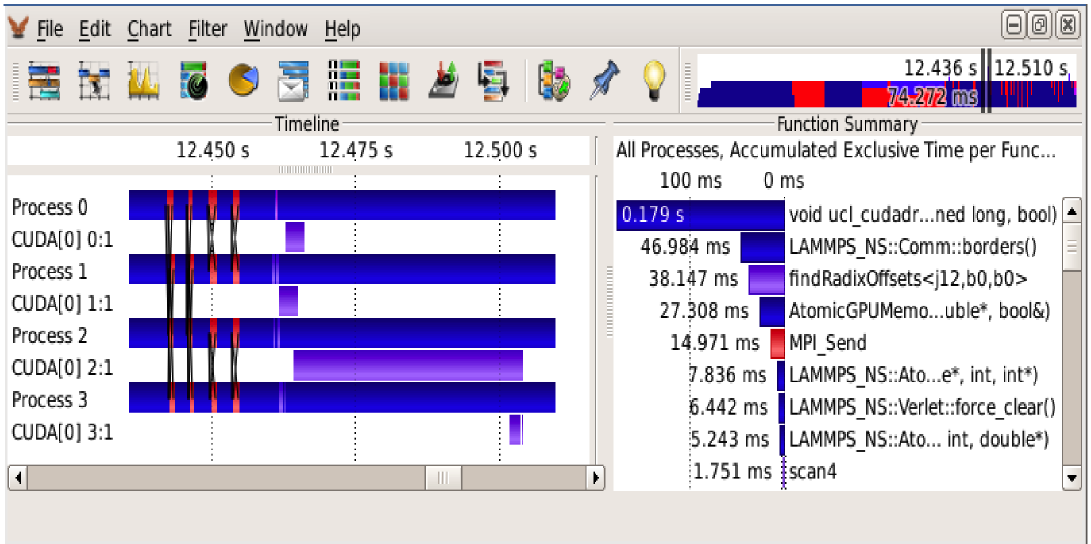

Figure 6.4 Vampir when Applied to LAMMPS Accelerated with GPU.

It is possible to analyze GPU applications that have been developed with HMPP in Vampir. The code generated by HMPP uses the CUDA run time library as a backend. The calls to the CUDA library are wrapped by VampirTrace in the same way this is done for manually developed CUDA applications. The same functionality is therefore available for HMPP applications, including memory copies, kernel (codelets) executions, and performance counters. Vampir exposes details on how HMPP maps the codelets to the GPU but might lose some information about the high level HMPP code. This preservation of high level HMPP semantic is subject to ongoing development. HMPP and VampirTrace both use compiler wrappers for their functionality. Those compiler wrappers have to be chained for the integration. This is done by using *vtcc* as a compiler for *hmpp*.

6.4 INNOVATION UPDATES

Dashboard—electronic Simulation Monitoring (eSiMon)

Computational scientists have a new weapon at their disposal. On February 1, 2011, the electronic Simulation Monitoring (eSiMon) Dashboard, version 1.0, was released to the public, allowing scientists to monitor and analyze their simulations in real time. Developed by the Scientific Computing and Imaging Institute at the University of Utah, North Carolina State University, and ORNL, this “window” into running simulations shows results almost as they occur, displaying data just a minute or two behind the simulations themselves (Figure 6.5). Ultimately, the Dashboard allows scientists to concentrate on the science being simulated rather than having to learn HPC intricacies, an increasingly complex area as leadership systems continue to break the petaflop barrier. This work was funded through a collaboration between DOE/ASCR, DOE/FES, and the OLCF.

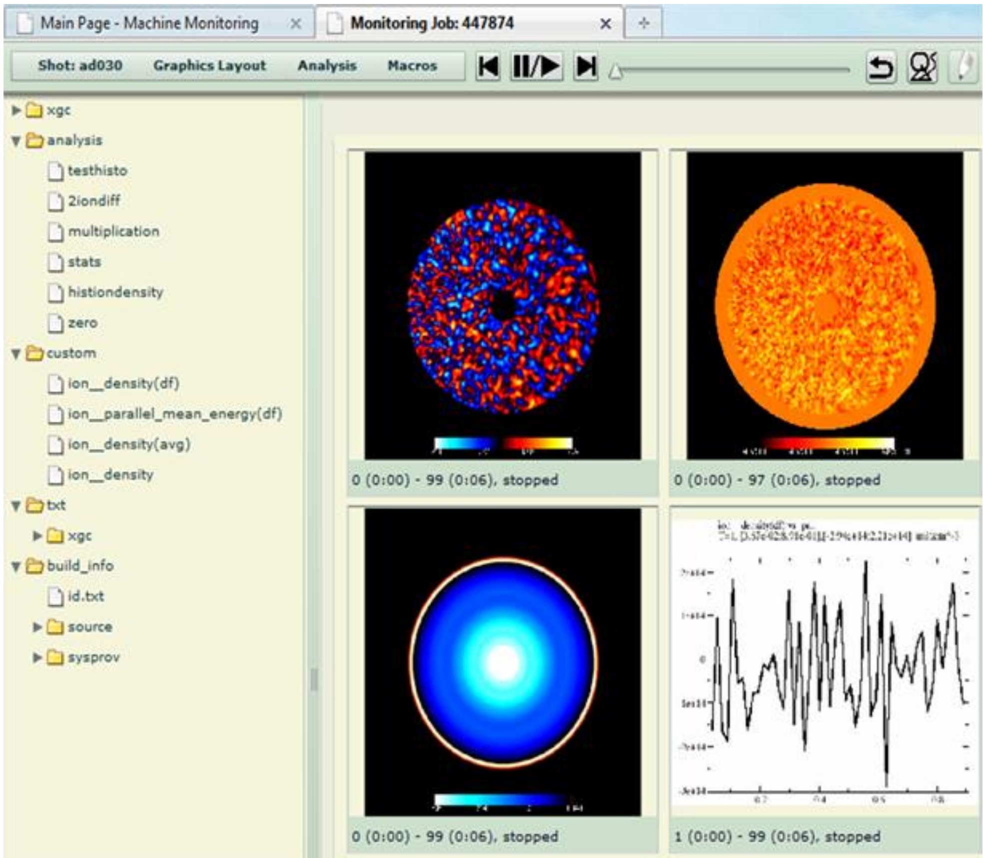

Figure 6.5. Screenshot of a XGC1 simulation monitoring. Fusion scientists are monitoring their Plasma Edge Simulation code via eSiMon. Images and/or movies are tracked as the simulation is running and researchers can check for any problems.

7. RISK MANAGEMENT

CHARGE QUESTION 7: ***Is the Facility effectively managing risk?***

OLCF RESPONSE: The OLCF has a very successful history of anticipating, analyzing and rating, and retiring risk for both project-based and operations-based risks. Our risk management approach uses the Project Management Institute's best practices as a model. Risks are tracked and, when appropriate, are retired, re-characterized, or mitigated. The major risks currently being tracked are listed and described below. Any mitigation(s) planned for or implemented are included in the descriptions. The operational risks are broadly categorized as across the board; system utilization; outages; performance; file systems–operations; and development environments. Table 7.1 references the "low", "medium", and "high" definitions used by the OLCF for operational risks. The OLCF has two "high" operational risks: that the funding is inadequate to cover the projected spend plan, and availability of an exascale facility. To address this, the OLCF maintains close contact with the federal project director and ASCR program office to understand the changing funding projections so alternative plans can be made in a timely manner.

2011 Operational Assessment Guidance – Risk Management

Each Facility utilizes a risk management plan and procedures to document operational risks. The risk management plan describes how risks are identified, rated, and monitored.

The Facility documents its risk management plan and provides information about the development, evaluation, and management of the most significant operating and technical risks encountered during the reporting period.

The Facility should highlight various risks to include:

- Major risks that were tracked for the current year:
- The risks that occurred and the effectiveness of their mitigations:
- A discussion of risks that were retired during the current year:
- Any new or re-characterized risks since the last review: and
- The major risks that will be tracked in the next year, with mitigations as appropriate.

2011 Approved OLCF Metrics – Risk Management

Risk Management: **The OLCF will provide a description of major operational risks.**

Risk Management

The OLCF's Risk Management Plan (RMP) describes a regular, rigorous, proactive, and highly successful review process first implemented in October 2006. Operations and project meetings are held weekly, and risk is continually being assessed and monitored. The Federal Project Director (residing at the DOE Oak Ridge Office (ORO)) attends each monthly project/operation risk meeting. The OLCF sends aggregated risk reports monthly to the DOE program office.

The OLCF has a highly successful history of anticipating, analyzing and rating, and retiring risk for both project- and operations-based risks. Our risk management approach uses the Project Management Institute's best practices as a model. The RMP includes:

- identifying and analyzing potential risks,
- ensuring that risk issues are discovered and understood early on,
- ensuring that mitigation plans are prepared and implemented, and
- developing budgets with consideration of risk.

Risk assessment is a major consideration for the DOE SC. OLCF staff have attended DOE sponsored risk management events including the 2008 Risk Management Techniques and Practice (RMTAP) workshop. This workshop concluded that HPC projects often require a tailoring of standard risk management practices and that the special relationship between the HPCCs and HPC vendors must be reflected in the risk-management strategy.

Several of the workshop best practices recommendations are standard OLCF practice, including

- developing a prioritized risk register with special attention to the top risks,
- establishing a practice of regular meetings and status updates with the platform partner,
- supporting regular and open reviews that engage the interests and expertise of a wide range of staff and stakeholders, and
- documenting and sharing the acquisition/build/deployment experience.

OLCF risk assessment is a six-step process. Once a risk is identified through a discussion of threats and vulnerabilities, the chance of occurrence is determined and its impact on project or operations scope, cost, and schedule are assessed. Then a (typically informal) cost/benefit analysis is performed to determine if mitigation activities are called for. If so a plan is made and executed when appropriate. Mitigation activities are reported and tracked as with any other project work breakdown structure (WBS) activity element, or if there are operational risks, they are reported and tracked as part of the periodic OLCF risk meetings.

Risk planning focuses on likelihoods and consequences. Likelihood is assigned as "very likely" (over 80%), "likely" (between 80% and 30%), and "unlikely" (below 30%). Impact category thresholds differ according to the impact area and whether the impact is to a particular project or to operations. For OLCF operations, the Table 7.1 is used.

Table 7.1 Risk Planning Focuses on Likelihoods and Consequences

Category	Impact on Project		
	Low	Medium	High
Cost	<$250,000	>$250,000 and <$500,000	>$500,000
Schedule	<1 month	>1 month and <3 months	>3 months
Scope (based on performance metrics)	<10%	>10% and <20%	>20%
Other	Depends on the area of concern and is usually a subjective evaluation.		

A risk management software application provides a risk register repository and helps the team to record, track, and report on identified project risks. The application uses the assessment to rate and rank them as they are entered and updated over time. A risk rating is a dimensionless numeric score generated from a combination of likelihood and the highest rated impact, which is used to give a sense of relative importance.

The risks to be tracked next year are in the Operational Risk Register, which is reviewed and updated on a regular basis. The highest priority risk is projected to be funding uncertainty.

At its periodic risk reviews, weekly staff meetings, and ad hoc discussions, the OLCF management team continues to focus attention on the high and moderate risks while keeping an eye on low risks, which may increase in importance over time. The managers and group leaders benefit from a thorough familiarity with previous risk profiles as they review the risk register, and they are in a strong position to anticipate future events. There were 173 risks registered for the OLCF-1 project that have been retired, and the OLCF-3 project team is collecting and assessing the risks associated with that new project.

At the time of this writing, 34 entries in the OLCF operations risk register. They fall into two general categories: risks for the entire facility and risks particular to some aspect of it.

Across-the-board risks are concerned with such things as safety, funding/expenses, and staffing. More focused risks are concerned with reliability, availability, and use of the system or its components (e.g., the computing platforms, power and cooling, storage, networks, software, and user interaction).

Costs for handling risks are integrated within the budgeting exercises for the entire facility. Risk mitigation costs are estimated as any other effort cost or expense would be. For projects, a more formal bottom-up cost analysis is performed on the WBS. However, for operations, costs of accepted risks and residual risks are estimated by expert opinion and are accommodated as much as possible in management reserves. This reserve is re-evaluated continually throughout the year. The following are the known risks in the OLCF Operations Risk Register.

7.1 ACTIVE RISKS

Across-the-board

Funding uncertainty is one of the highest risks for the OLCF. Annual budgets are set with guidance from the ASCR office, but actual allocated funds are unknown until Congress passes funding bills. Continuing resolutions are common, and we often go several months before actual funding is resolved. The risk is that we may have to delay some purchases, activities, hiring, etc., or possibly adjust lease payment schedules, resulting in substantially higher costs and perhaps

- schedule delays. We will continue to maintain close contact with the Federal Project Director and ASCR Program Office to understand the changing funding projections so that alternative plans can be made in sufficient time. Where possible, we will structure contracts to accommodate flexible payment terms. Rating: **High**
- DOE’s long term plans include pre-exascale and exascale systems before the end of this decade. ORNL has a plan to house the exascale system in building 5600 by moving other systems out of the building. However, the much preferred approach would be to build a new building that is designed for exascale from the beginning. OMB has rejected third party financing as a method of building such a facility so this will need a congressional line item. Rating: **High**
    - This is a new risk, introduced in the past year.
- Labor and/or utility costs may increase over time at rates higher than expected. We will accept the risk and conservatively budget for utilities. Where possible, we will purchase energy efficient computing and storage systems to minimize the impact. We will work closely with laboratory

leadership to control labor cost increases and budget for reasonable escalations in labor rates. Rating: **Low**

  - This risk was recharacterized in June, 2011 to cover labor and utility costs. Previously only utility costs were considered.

- Staffing is a concern. Much of the effort within the OLCF is provided by highly trained and highly experienced staff. The loss of critical skill sets or knowledge in certain technical and managerial areas may hinder ongoing progress. Good career development programs have been implemented within the division to retain high-quality personnel. Succession planning is promoted, and there are active laboratory-wide recruiting campaigns and outreach programs. Despite the best efforts in recruiting, training, etc., funding uncertainty continues to be a concern for the OLCF's ability to attract and keep the high-quality staff essential to its success. For example, several other risk register entries describe risk mitigation efforts involving Scientific Computing, HPC Operations, and Technology Integration Groups, whose contributions are critical to the mission of both the OLCF and DOE. Demands on these groups of specialists are increasing at an extraordinary rate and the danger remains that staff burnout will take its toll. Rating: **Low**

- There is always a risk that the facility experiences a safety occurrence resulting in serious personal injury. We work to reduce these risks with monitoring of worker compliance with existing safety requirements, daily tool box safety meetings, periodic surveillances using independent safety professionals, joint walk-downs by management and work supervisors, and encouragement of stop-work authority of all personnel. Observations from safety walk-downs will be evaluated for unsatisfactory trends (e.g., recurring unsafe acts). Unsatisfactory performance will receive prompt management intervention commensurate with severity of the safety deficiencies. Integrated Safety Management is a core performance metric for the entire laboratory. Safety is a top UT-Battelle priority that carries throughout the laboratory, and the OLCF understands that it is critical to its success to provide a safe working environment. Rating: **Low**

- System cyber security failures involving unauthorized access or use of systems may force a shutdown for extended periods or otherwise degrade system productivity. We have developed a cyber security plan that implements a security level of Moderate for security objective of confidentiality as defined in the Federal Information Security Act of 2002, P.L. 107-34T. This includes such things as continual monitoring for security breaches, user identity checks prior to granting accounts, two factor authentication, and periodic formal tests and reviews. A U.S. government laboratory is subject to intense external assaults on its IT systems and networks. OLCF staff, in concert with ORNL's cyber security technical and policy teams, are constantly looking for ways to balance the protection of its IT resources with its need to continue its science mission. Rating: **Low**

System utilization

- The impending OLCF-3 system upgrade has a new computer architecture, using both traditional x86 CPUs and GPUs to achieve unprecedented performance and energy efficiency. OLCF-3's architecture with both Opteron processors and GPUs gives the users the opportunity to port codes from Jaguar, Intrepid, or other traditional systems to run on just the Opteron, while continuing to work on using the GPUs. As pointed out at the July 2009 Lehman review of the project, we must develop a strategy to allow applications to be ported to OLCF-3 and still have portability to more traditional architectures. The risk is that users will be slow to adopt this programming model, resulting in application performance on the OLCF-3 system that would be lower than what it could be. As a mitigation strategy, we have decided to get an early delivery of 960 Fermi+ cards to be integrated into Jaguar to allow staff, developers, and users to have access to a GPU based system to begin early work on porting applications. It is important to work with users early to begin porting to the system so that the machine will be judged as successful by delivering breakthrough science. Rating: **Medium**

  - This risk was recharacterized from Low to Medium after gaining a better understanding of the capabilities and intentions of the user community.

- Related to the risk above is the situation where leadership-level computing is not achieved. Too many application runs may be submitted that do not achieve “leadership” status. The OLCF has established job queue policies with high preference for leadership jobs and continually evaluates their effectiveness. The OLCF is involved with the INCITE proposal selection process, which ensures that leadership projects receive allocation preference. The Scientific Computing Group has been established to help users scale their applications to leadership levels. Leadership computing is defined as utilizing a certain percentage of the available computing capability of a system. In CY 2011 YTD, Jaguar XT5 has been running at 54% capability usage.. Continued improvement is enabled by the Scientific Computing Group helping scientists scale up their applications. Rating: **Medium**

- Upgrade of system takes too long, causing users to seek other alternatives. With a new system of this size and complexity, there may be problems that delay completion of the acceptance tests, thus delaying user access. There is very low risk with the initial XK6 processor and memory upgrade. The new Interlagos processor with the Bulldozer core has been undergoing extensive testing at AMD and Cray. We will be early in the delivery cycle, but not the first customers to receive the processor. The Gemini interconnect has been in the field for a year with no major unresolved problems. There is risk that at the new scale Gemini will exhibit some problems, but we will test this in acceptance. We will also require Cray to keep the existing Seastar based boards for a period of time to make sure that the Gemini is working properly before those boards are surplused. Rating: **Low**

  - This is a new risk, introduced in the past year.

Outages

- Power outages from external causes may create delays in user job completion or otherwise hinder system performance. The OLCF constantly evaluates risk in this area. It has installed cost-effective back-up capabilities (generators, UPS, dual-power cabinet designs, etc.). Cooling equipment failures are also possible. HPC systems operate with fairly strict temperature requirements. OLCF systems have automatic shutdown features in case temperatures rise above a set threshold. In addition, there are redundant chillers (five, where the systems could run on as few as two). There are also redundant cooling towers and pumps, and buildings 5309, 5800, and 5600 are interconnected, allowing them to distribute chilled water among themselves as necessary

  Network outages could prevent effective system use. If networks are inoperable or degraded, some users could lose access to the OLCF systems. There is some redundancy in ESNet with a backup OC-48c connection, but there is some residual risk there. To mitigate this risk ORNL is implementing physically diverse network paths to connect Lab to the internet with goal of full redundancy by end of CY 2011. The ANI program will provide a 100G/sec connection by 2012. Additionally, ORNL has contracted with a commercial network provider to supply alternative network capability, although that would be at reduced performance. Rating: **Low**

Performance

- Maintaining high availability and stability of systems is critical to users and for the OLCF to meet DOE performance targets. There is a risk that the system stability and availability may not be sufficient to meet these requirements. This risk includes the disruptions of the impending XK6 upgrade. One risk in this installation is the scaling of the Gemini interconnect to a 200 cabinet system. The largest system built to date is 96 cabinets. In general, policies have been implemented that control availability to minimize maintenance downtimes, coordinate upgrades, maximize fault-tolerant HW and SW, etc. Availability and stability are continually monitored in order to detect trends in time to take remedial action. With respect to mitigations specific to the upgrade, we have built the upgrade schedule to minimize the period of disruption, at the expense of total available resources at times. If there are problems with the installation, we can retain the XT5 capability until the problem is resolved. Rating: **Low**
  - Updated for current technological scope, e.g., XK6 board upgrade
- There is a risk that INCITE hour goals may not be met because the upgrade to Jaguar may require downtimes longer than expected or longer than user have planned for. DOE has set aside 125,000,000 ALCC hours to account for the time lost during the upgrade. Moreover, some projects may be extended into 2012, since the first few months of the calendar year are typically low utilization times. Rating: **Low**Users require support (e.g., account management, help desk, training, etc) to use large-scale computing systems effectively. There is a risk that the support we provide in one or more of these areas will not be adequate. To mitigate this risk, OLCF staff communicate frequently and directly with users, measure satisfaction with formal surveys, and use liaisons to get better insight into users' problems and issues. OLCF will also develop and conduct training classes for both users and staff in effective ways to take advantage of the new architecture. This risk is somewhat different from user dissatisfaction with system use due to technological inadequacies (e.g., poor system performance, unscheduled downtimes, lost data). Those are covered in other registry entries. This risk has to do with the interactions users have with the User Assistance & Outreach Group. Rating: **Low**
  - Recharacterized from an earlier risk, which introduced the training element.
- The restructuring of applications may not be sufficient to maintain portability of a given application. The level of portability of a given code is a function of the domain specific and architectural specific implementations in that application. The goals of the OLCF-3 project are to first port six specific applications to the new hybrid architecture. To support ongoing operations we are developing generalized prescription to transform applications to a hybrid architecture and to preserve or enhance the level of portability of the current application. The programming model that we propose to use requires a restructuring to utilize the standard distributed memory technologies in use today (e.g., MPI, Global Arrays etc.) and then a thread based model (e.g., OpenMP or Pthreads) on the node that captures larger granularity work than that is typically done in applications today. In the case of OpenMP the compiler can facilitate and optimize this thread level of concurrency. This restructuring is agnostic to the particular multi-core architecture and is required to expose more concurrency in the algorithmic space. Our experience to date shows that we almost always enhance the performance with this kind of restructuring. The utilization of directives based methods will allow the lowest level of concurrency to be exposed (e.g., vector or streaming level programming) concomitantly. This means that that bottom level of concurrency can be generated by a compiler directly. We expect this kind of restructuring will work effectively with portable performance on relevant near-term architectures (e.g., IBM BG/Q, Cray Hybrid, and general GPU based commodity cluster installations). he adoption of multiple instantiations of compiler infrastructure tools to maximize the exposure of multiple levels of concurrency in the applications. This will be abetted by publishing the case studies and

experience with the six project applications coupled with the appropriate training of our user community. Rating: **Low**

- Scientists may decline to port to heterogeneous architecture. Some users may determine that it is too much effort to port their code to the new heterogeneous architecture. Outreach, training, and the availability of libraries and development tools will ameliorate some of the resistance. Current trends in publication venues imply that many development teams are exploring architectures with accelerators which is contrary to this risk. Rating: **Medium**

- Communication library (MPI) may not be able to survive system failures, causing running jobs to fail. Fault tolerance for the MPI standard is currently being defined, with ORNL leading the effort, and developing the support within the Open MPI library. Rating: **Low**

File systems—operations

- With Oracle's acquisition of Sun, and the Lustre file system IP, followed by a halt to future development of the Lustre file system by Oracle, there is a risk that future development of Lustre will stagnate. Features needed for Lustre to be viable for OLCF-3 or future systems may not be developed. We have put in place the OpenSFS consortium to begin addressing the issue. OpenSFS will address the longer term operational risk via collaborative and contract development of Lustre on Linux for HPC. In the short term, we will transfer the risk to a contractor to upgrade the metadata handling in Lustre and the resiliency to server failure of the Lustre file system. Rating: **Low**
- Metadata performance is critical to a wide variety of leadership applications. There is a risk that single metadata server performance will not be adequate and may adversely impact both applications and interactive users. This risk has already occurred and will continue impacting performance. The OLCF is working with other major Lustre stakeholders through OpenSFS to develop features to improve single metadata server performance and follow-on support of multiple metadata servers for the Lustre file system. Contract development through the OLCF with Whamcloud is accelerating the deployment of Lustre 2 on Jaguar which has demonstrated improved performance, confirmed during dedicated Lustre test shots on Jaguar. The OLCF is working with application teams to reduce their metadata workloads through code restructuring and the use of middleware I/O libraries. Tools have been developed to monitor and respond to metadata performance slowdowns in order to minimize the impact to the overall user population. Multiple file systems have been deployed reducing load on the metadata server. Rating: **Medium**

- There is a risk that the file system will become unstable at larger scales. The introduction of new features within Lustre and the transition to a new Lustre release may exhibit instability at larger scales. Our transition to Lustre 1.8.6 and later Lustre 2.x may present software bugs or scalability limitations that must be addressed prior to returning the system to operations. The OLCF will leverage contractual development of Lustre features and stabilization of these features at scale. Contractual development of improved metadata performance and improved resiliency at scale are underway via the Scalable File Systems Center (SFSC) at the OLCF. The SFSC includes an onsite Lustre engineer presence at the OLCF. Testing of these features at progressively larger scales will be conducted utilizing the storage testbed systems and dedicated test shots on Jaguar XT5 and upgraded XK6 platforms. In addition to these activities the OLCF will leverage joint development of Lustre scalability and stability features within the Open Scalable File Systems consortium and testing of these features using testbed resources at Cray, DDN, LLNL, ORNL and other OpenSFS member sites. The Technology Integration group will work closely with Cray to ensure that the required version of Lustre is supported on the Jaguar and subsequent Titan platforms. Rating: **Low**

- The scale of the data volume increases the probability that data integrity will fail somewhere. The risk is not being able to identify corrupt data and manage it appropriately. The OLCF will work closely with others in the Lustre community via OpenSFS to reduce the probability of data corruption via improved resiliency mechanisms. The OLCF will work on improved detection of data corruption once occurred and develop tools to quickly identify data within the file system that could be impacted by a component failure. Newly established procedures will minimize the likelihood and impact of failures. Rating: **Low**

Development environments

- To use HPC effectively, a fully functional software development environment is necessary. The risk is that some of these tools may be inadequate to allow practical levels of productivity. As was pointed out by the CD-1 Lehman Review panel, the OLCF-3 system will not be perceived as successful if programming the system requires that the users are required to use a very different programming method that would not be compatible with other large system such as Jaguar and the new BG/Q system at ANL. We have developed a strategy to prevent this problem by using compilers, debuggers, and performance measurement tools that are compatible with other systems for the programming environment of OLCF-3. We also created the Application Performance Tools Group within the NCCS to own the problem. We surveyed users on their requirements in this area and the adequacy of the tools available or planned. We have initiated contracts with vendors to supplement the work of the Tools Group to obtain additional functionality. Rating: **Low**
    - Compilers. Platforms are changing rapidly, with increasing system heterogeneity as well as the requirement to extract unprecedented levels of parallelism from the applications. The commodity market is operating at a much lower scale and is not funding the development of compiler technology at the levels needed for HPC systems. The OLCF will track system requirements and compiler vendors and make targeted investments to meet specific OLCF needs. Additionally, the research community is being tracked for ways to bring needed capabilities into vendor-supported compiling systems. The OLCF participates in actions to develop a large HPC community that works in concert to remedy the situation.
    - Debuggers. On today's large-scale systems, debugging support is limited, with only one debugger vendor (DDT) capable of providing debugging support at large scale (after our investment). As system scales continue to grow at a rapid pace, the scalability of debugging solutions needs to increase as well. In addition, high-performance analysis tools for inspecting data for the source of code errors are extremely inadequate. The OLCF will continue with targeted investment in improving debugging capabilities. Additionally, the research community is being tracked for ways to bring needed capabilities into vendor supported debugging systems. The OLCF participates in actions to develop a large HPC community that works in concert to remediate the situation.
    - Application performance tools. Detailed trace-based performance analysis is limited to runs of, at most, a few tens of thousands of cores. Our ability to understand application performance at the scales leadership applications are expected to run is extremely limited. The commodity market is operating at a much lower scale and is not funding the development of performance tool technology needed for HPC systems. The OLCF will continue with targeted investment in improving performance analysis capabilities. Additionally, the research community is being tracked for ways to bring needed capabilities into vendor-supported debugging systems as the volume of data generated at large scale is large and new analysis techniques need to be developed. The OLCF participates in actions to develop a large HPC community that works in concert to remediate the situation.

7.2 RETIRED RISKS

Three risks were retired or recharacterized this past year.

- Contention between systems for Spider adversely impacts applications. We will work with Sun to establish requirements for quality of service mechanisms. Develop patches to Lustre to add critical features to support QoS.
  - RETIRED: 4/1/2011 – Following full deployment of the Spider file system infrastructure and substantial experience in operations this has proven not to be a risk to operations. Adequate bandwidth has been provisioned to each system ensuring a balanced configuration for OLCF computational assets.
- Differences between Lustre versions on Spider and the Cray systems impedes integration. Lustre currently provides backward compatibility between major releases. Our operational environment includes both Lustre 1.6.x and Lustre 1.8.x systems and will soon include Lustre 2.x. We conduct testing of mixed Lustre versions prior to deployment on our production systems. When Lustre versions exhibit incompatibility we work with the vendor to address these issues.
  - RETIRED: 8/6/2011- We have developed operational processes to test and integrate new Lustre releases and stage upgrades to maintain compatibility of systems across the OLCF complex.
- Future disk technology may be different from expected. In order to remain within budget and achieve the performance needed the OLCF staff will have to ensure that it sets the performance requirements at a level that stretches the manufacturers capabilities and are yet still achievable. Once a manufacturer is chosen, ORNL will actively work with the manufacturer by providing feedback on the product to ensure that the performance requirements are achieved.
  - RETIRED: 8/6/2011 - We have a very good understanding of what disk technologies will be available for our next procurement through careful market analysis.
- Applications are not ready for new technologies. As new or upgraded computing platforms are acquired, the applications may not be sufficiently prepared to take advantage of the increase computing capabilities. Continue efforts by the OLCF Scientific Computing Group which works closely with the HPC user community to improve their codes to take maximum advantage of any new technology that OLCF introduces. Continue to acquire testbeds to provide early access to new technologies. The User Services and Scientific Computing Groups also conduct education, outreach, and training to continually expand and extend and skill levels of the HPC user base and ORNL staff.
  - RETIRED: 8/6/2011 - Restated as Risk #912, 361, 906
- Sun may eliminate or reduce availability of support for Lustre. Sun has recently indicated that their support model for continued Lustre development may change significantly. Lustre is open-source software. Should Sun reduce their support below acceptable levels, we will increase our engagement with, and financial support to, the Lustre open source developer community.
  - RETIRED: 8/6/2011 – Restated as risk #913 to recognize Oracle acquisition of Sun.
- Lack of availability of on-site support for Vampir. On site support is used in the collaboration with TU-Dresden, to work with users to help in identifying missing functionality/capabilities. The on-site support has been very helpful in identifying issues, and rapidly obtaining fixes for these. A reduction in such support will slow down progress. We will accept this risk. Work early on with the vendor on identifying potential candidates.
  - RETIRED: 8/6/2011 – We now have adequate on-site support for Vampir.

8. CYBER SECURITY

CHARGE QUESTION 8: ***Does the facility have a valid cyber security plan and authority to operate?***

OLCF RESPONSE: Yes, the most recent OLCF Authority to Operate (ATO) was granted on June 21, 2011. The current ATO expires on June 20, 2012.

2011 Operational Assessment Guidance – Cyber Security

The Facility provides information on its approved Cyber Security Program Plan and approved Cyber Security Certification and Accreditation, in accordance with DOE Orders and Federal Regulations.

2011 Approved OLCF Metrics – Cyber Security

The OLCF will provide the date of approval and expiration of our Authority to Operate.

All information technology (IT) systems operating for the federal government must have certification and accreditation (C&A) to operate. This involves the development of policy, the approval of policy, and the assessment of how well the organization is managing those IT resources—an assessment to determine that the policy is being put into practice.

The OLCF has the authority to operate for 1 year under the ORNL C&A package approved by DOE on June 21, 2011. The ORNL C&A package uses the National Institute of Standards and Technology Special Publication 800-53 revision 3 as a guideline for security controls. The OLCF is accredited at the moderate level of controls, which authorizes the facility to process sensitive, proprietary, and export-controlled data.

In the future, it is inevitable that cyber security planning will become more complex as the Center continues in its mission to produce great science. As the facility moves forward, the OLCF is very proactive, viewing its cyber security plans as dynamic documentation and responding to and making modifications as the needs of the facility change to provide an appropriately secure environment.

9. SUMMARY OF THE PROPOSED METRIC VALUES FOR 2012 OAR

The OLCF provides (below) a summary table of the metrics proposed for the 2012 Operational Assessment Review and the values for 2011.

Are the processes for supporting the customers, resolving problems, and communicating with key stakeholders and Outreach effective?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Customer Metric 1: Customer Satisfaction		
Overall OLCF score on the user survey will be satisfactory (3.5/5.0) based on a statistically meaningful sample.	Overall user satisfaction rating for the 2010 user survey was 4.3, "very satisfied."	Overall score on the OLCF user survey will be satisfactory (3.5/5.0) based on a statistically meaningful sample.
Annual user survey results will show improvement in at least ½ of questions that scored below satisfactory (3.5) in previous period.	None of the user responses in the previous period (2009 user survey) were below the 3.5 satisfaction level.	Annual OLCF user survey results will show improvement in at least ½ of questions that scored below satisfactory (3.5) in the previous period.
Customer Metric 2: Problem Resolution		
N/A	N/A	OLCF survey results related to problem resolution, if any, will be satisfactory (3.5/5.0) based on a statistically meaningful sample.
80% of OLCF user problems will be addressed within three working days, either resolving the problem or informing the user how the problem will be resolved.	In CY 2011 YTD, 89.5% of queries were addressed within 3 working days.	Target: 80% of OLCF user problems will be addressed within three business days, by either resolving the problem or informing the user how the problem will be resolved.
Customer Metric 3: User Support		
OLCF will report on survey results related to user support	The OLCF does not have a survey question specifically targeted at the full range of user support from OLCF staff members, and instead solicits an overall user satisfaction rating and comments about support, services, and resources.	OLCF survey results related to User Assistance and Outreach, if any will be satisfactory (3.5/5.0) based on a statistically meaningful sample.
N/A	N/A	OLCF will provide a summary of training events including number of attendees. Target: At least 4 training events.

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Customer Metric 4: Communications with Key Stakeholders		
N/A	N/A	OLCF survey results related to communication, if any will be satisfactory (3.5/5.0) based on a statistically meaningful sample.
N/A	N/A	OLCF will provide representative communications with key stakeholders. Target: An example of at least one representative communication with users and one representative communication with DOE ASCR.

Is the facility maximizing the use of its HPC systems and other resources consistent with its mission?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Business Metric 1: System Availability (for a period of one year following a major system upgrade, the targeted scheduled availability is 85% and overall availability is 80%)		
Scheduled Availability: 95%	XT5 (93.9%); XT4 (97.6%); HPSS (99.9%); Spider (98.5%); Spider2 (99.9%); Spider3 (99.9%).	Scheduled availability Target: 85% (lower in FY12 due to the compute system upgrade).
Overall Availability: 90%	XT5 (88.7%); XT4 (97.1%); HPSS (98.9%); Spider (96.5%); Spider2 (99.1%); Spider3 (99.2%).	Overall availability Target: Jaguar: 80%; HPSS 90%; Spider 80%
Business Metric 2: Resource Utilization		
OLCF will report on INCITE allocations and usage.	CY 2011 INCITE allocations of 930 million hours. INCITE usage in CY 2011 to date (6/30/2011) is 375 million core-hours, or 40.3% of the total allocation.	Target: OLCF INCITE usage will be at least 60% of total system usage of the Opteron processors in CY2012
Business Metric 3: Capability Usage		
At least 40% of the consumed core hours will be from jobs requesting 20% or more of the available cores.	The OLCF is on track to exceed the capability usage metric in CY 2011 (achieved 54% YTD).	At least 30% of the consumed processor hours will be from jobs requesting 20% or more of the available Opteron cores.

Is the facility enabling scientific achievements consistent with the Department of Energy strategic goals 3.1 and/or 3.2?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Strategic Metric 1: Scientific Output		
The OLCF will report numbers of publications resulting from work done in whole or part on the OLCF systems.	181 publications in 2011 YTD have been the result of work carried out by users of OLCF resources.	The OLCF will report numbers of publications resulting from work done in whole or part on the OLCF systems. Target: On average, two publications per INCITE project.
Strategic Metric 2: Scientific Accomplishments		
The OLCF will provide a written description of major accomplishments from the users over the previous year.	Reference Section 4.	The OLCF will provide a written description of major accomplishments from the users over the previous year. Target: Descriptions of at least 5 major accomplishments.
Strategic Metric 3:		
The OLCF will report on how the Facility Director's Discretionary time was allocated.	Reference Section 4 and Appendix A.	The OLCF will report on how the Facility Director's Discretionary time was allocated, including project title, PI, PI's home organization, processor hours allocated and usage to date. Target: None

Are the costs for the upcoming year reasonable to achieve the needed performance?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Financial Performance		
The OLCF will report on budget performance against the previous year's Budget Deep Dive projections.	Reference Section 5	The OLCF will report on monthly budget performance against the current baseline agreed. Reporting categories will include effort, lease payments, operations and cyber security. The baseline will be revised as needed with the ASCR PM to reflect updated budget actions. Target: Within 10% variance between then-current baseline spend plan and actual spending for the year.

What innovations have been implemented that have improved the facility's operations?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Innovation Metric 1: Infusing Best Practices		
The OLCF will report on new technologies that we have developed and best practices we have implemented and shared.		The OLCF will report on new technologies that we have developed and best practices we have implemented and shared. Target: at least 1
Innovation Metric 2: Technology Transfer		
The OLCF will report on technologies we have developed that have been adopted by other centers or industry.		The OLCF will report on technologies we have developed that have been adopted by other centers or industry. Target: None

Is the Facility effectively managing risk?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Risk Management		
The OLCF will provide a description of major operational risks.	Reference Section 7.	The OLCF will provide, a description of major operational risks, including realized or retired risks: Target: at least 5 risks discussed.

Does the facility have a valid cyber security plan and authority to operate?

CY 2011 Target	CY 2011 YTD Achieved	CY 2012 Target
Cyber Security Plan		
The OLCF will provide the date of approval and expiration of our authority to operate.	The OLCF authority to operate was granted on June 21, 2011.	Target: Maintain valid authority to operate.

APPENDIX A. OLCF DIRECTOR'S DISCRETIONARY AWARDS: CY 2010 AND 2011 YTD

Table A.1 OLCF Director's Discretionary awards: CY 2010 and 2011 YTD

PI	Affiliation	2010 Allocation	Carryover to 2011	New 2011 Allocation	Project Name
Shaikh Ahmed	Southern Illinois University Carbondale	1	0		Multimillion-Atom Modeling of Harsh-Environment Nanodevices
Leslie Hart	NOAA-ESRL	50,000	50,000		NOAA Benchmark Portability
John Cobb	ORNL	50,000	50,000		Neutron Scattering Science Exploratory Projects
Amra Peles	United Technologies Research Center	100,000	7,979		Nanostructured Catalyst for WGS and Biomass Reforming Hydrogen Production
John Dutton	Prescient Weather	100,000	100,000		CFS Reanalysis Extension
Christopher Lynberg	Centers for Disease Control and Prevention	100,000	100,000		CSC Scientific Computing Architecture
Kenneth Smith	United Technologies Research Center	100,000	94,333		Surface Tension Predictions for Fire-Fighting Foams
Srdjan Simunovic	ORNL	100,000	14,493		Development of a Global Advanced Nuclear Fuel Rod Model
Stephen Nesbitt	UIUC	165,000	115,797		Dynamically Downscaling the North American Monsoon Using the Weather Research and Forecasting Model with the Climate Extension (CWRF)
Patrick Joseph Burns	Colorado State University	200,000	200,000		Parallel Lagged Fibonacci Random Number Generation
Christopher Taylor	LANL	200,000	89,501		Fundamental Properties of the Stability of Exposed and Oxygen-covered Tc-Zr Alloy Surfaces from Density Functional Theory
Emilian Popov	ORNL	200,000	188,718		Testing STARCCM+ on Jaguar for Computing Large Scale CFD Problems
Stephen Poole	ORNL	300,000	0		FASTOS Community Allocation

PI	Affiliation	2010 Allocation	Carryover to 2011	New 2011 Allocation	Project Name
Oleg Zikanov	University of Michigan	400,000	396,401		Effect of Liquid-Phase Turbulence on Microstructure Growth During Solidification
Ilian Todorov	STFC Daresbury Lab	500,000	440,888		An Investigation of the Channel-Opening Movements of the Nicotinic Acetylcgikube Receptor
David Erickson	ORNL	500,000	172,260		WRF Downscaling
Dale I Pullin	California Institute of Technology	500,000	194,776		Direct Numerical Simulation of the Mach Reflection Phenomenon and Diffusive Mixing in Gaseous Detonations
Marco Arienti	United Technologies Research Center	500,000	467,095		Multiphase Injection
James Chelikowsky	University of Texas Austin	500,000	406,510		Simulating the Emergence of Crystallinity: Quantm Modeling of Liquids
James Nutaro	ORNL	500,000	500,000		Qualitative System Identification for Massive Data Sets: Knowledge Discovery from Observations of Biological Systems
Michael Matheson	ORNL	500,000	1,084,560		Exploration of High Resolution Design-Cycle CFD Analysis
Alexei Khokhlov	University of Chicago	600,000	600,000		First-principles Petascale Simulations for Predicting DDT in H2-O2 Mixtures
Pablo Carrica	University of Iowa	750,000	20,167		Large-scale Computations of Wind Turbines using CFDShip-Iowa Including Fluid-Structure Interaction
Tommaso Roscilde	Ecole Normale Superieure de Lyon	800,000	0		Emulating the Physics of Disordered Bosons with Quantum Magnets
Jason Hill	University of Minnesota	900,000	900,000		Air Pollution Impacts of Conventional and Alternative Fuels
Salman Habib	LANL	1,000,000	999,735		Dark Universe
Patrick Fragile	ORAU	1,000,000	1,000,000		Radiation Transport in Numerical Simulations of Black-Hole Accretion Disks
Lei Shi	Cornell University	1,000,000	999,980		Transport Mechanism of Neurotransmitter: Sodium Symporter
Jean-Luc Bredas	Georgia Institute of Technology	1,000,000	1,000,000		Electronic and Geometric Structure of Inorganic/Organic and Organic/Organic interfaces Relevant in Organic Electronics
Erik Deumens	University of Florida	1,000,000	777,712		EOM-CC calculations on diamond nano crystals
Moetasim Ashfaq	UT-Knoxville	1,000,000	993,364		Quantification of Uncertainties in Projections of Future Climate Change and Impact Assessments
Gregory Laskowski	GE Global Research	1,000,000	890,854		Investigation of Newtonian and non-Newtonian Air-Blast Atomization Using OpenFoam
George I-Pan Fann	ORNL	1,000,000	0		Prototype Advanced Algorithms on Petascale Computes for IAA II
Zizhong Chen	Colorado School of Mines	1,000,000	0		Fault Tolerant Linear Algebra Algorithms and Software for Extreme Scale Computing
Robert Patton	ORNL	1,000,000	934,680		High Performance Text Mining

PI	Affiliation	2010 Allocation	Carryover to 2011	New 2011 Allocation	Project Name
Kalyan Kumaran	ANL	1,000,000	1,000,000		Performance Measurements Using ALCF Benchmarks
Omar Ghattas	University of Texas Austin	1,000,000	150,618		Forward and Inverse Modeling of Solid Earth Dynamics Problems on Petascale Computers
Stephen Poole	ORNL	1,000,000	1,000,000		Gov-IP
Bhagawan Sahu	University of Texas Austin	1,000,000	990,876		Gap Engineering in Trilayer Graphene Nanoflakes
Gary Grest	SNL	1,000,000	1,000,000		Assembly of Nanoparticles at Liquid/Vapor Interface
Brian J Albright	LANL	1,000,000	2,000,000		Kinetic Simulations of Laser Driven Particle Acceleration
Nikolai Pogorelov	University of Alabama Huntsville	1,000,000	480,051		Modeling Heliospheric Phenomena with an Adaptive, MHD-Boltzmann Code and Observational Boundary Conditions
George Karniadakis	Brown University	1,500,000	1,276,488		NektarG-INCITE
Branden Moore	GE Global Research	2,000,000	172,836		Unsteady Performance Predictions for Low Pressure Turbines
Thomas Miller	California Institute of Technology	2,000,000	10,104		Proton Coupled Electron Transfer Dynamics in Complex Systems
Kalyan Perumalla	ORNL	2,000,000	1,999,980		An Evolutionary Approach to Porting Applications to Petascale Platforms
Barry Schneider	National Science Foundation	2,000,000	18,574		Time-Dependent Interactions of Short Intense Laser Pulses and Charged Particles with Atoms and Molecules
Dinesh Kaushik	ANL	2,000,000	2,000,000		Scalable Simulation of Neutron Transport in Fast Reactor Cores
Phil Colella	LBNL	2,500,000	228,877		Applied Partial Differential Equations Center. APDEC.
George Vahala	College of William and Mary	2,500,000	461,737		Lattice Algorithms for Quantum and Classical Turbulence
David Bowler	University College London	2,650,000	2,321,114		Modeling of Large-Scale Nanostructures using Linear-Scaling DFT
Gil Compo	University of Colorado	3,000,000	2,769,235		Developing a High Resolution Reanalysis Data set for Climate Applications (1850 to present)
Lee Berry	ORNL	3,000,000	80,635		Wave-Particle Intercations in Fusion Plasmas
Homayoun Karimabadi	University of California San Diego	3,000,000	3,000,000		Enabling Breakthrough Kinetic Simulations of the Earth's Magnetosphere through Petascale Computing
Paul Ricker	UIUC	3,150,000	2,000,000		Testing Active Galaxies as a Magnetic Field Source in Clusters of Galaxies
Mike Henderson	BMI Corporation	4,000,000	2,695,917		Smart Truck Optimization
Pratul Agarwal	ORNL	4,000,000	0		High Throughput Computational Screening Approach for Systems Medicine
Sean Ahern	ORNL	8,000,000	1,516,488		SciDAC 2 Visualization Center and Institute
Kate Evans	ORNL	5,000,000	0		Decadal Prediction of the Earth System after Major Volcanic Eruptions

PI	Affiliation	2010 Allocation	Carryover to 2011	New 2011 Allocation	Project Name
James Joseph Hack	ORNL			15,000,000	Ultra High Resolution Global Climate Simulation to Explore and Quantify Predictive Skill for Climate Means, Variability and Extremes
John Turner	ORNL			15,000,000	Fundamental studies of multiphase flows and corrosion mechanisms in nuclear engineering applications
Thomas Maier	ORNL			10,000,000	Predictive simulations of cuprate superconductors
Jerome Baudry	ORNL			6,000,000	High Performance Computing for Rational Drug Discovery and Design
Pui-kuen Yeung	Georgia Institute of Technology			3,000,000	Frontiers of Computational Turbulence
Zhengyu Liu	University of Wisconsin Madison			2,000,000	Assessing Transient Global Climate Response using the NCAR-CCSM3: Climate Sensitivity and Abrupt Climate Change
Thomas Jordan	University of Southern California			2,000,000	Deterministic Simulations of Large Regional Earthquakes at Frequencies up to 4Hz
Bobby Sumpter	ORNL			2,000,000	Computational Resources for the Nanomaterials Theory Institute at the Center for Nanophase Materials Sciences and the Computational Chemical and Materials Sciences group in the Computer Science and Mathematics Division
Terry Jones	ORNL			1,000,000	HPC Colony II
Sean Ahern	ORNL			1,000,000	Large-Scale Data Analysis and Visualization
William Martin	University of Michigan			1,000,000	Development of a Full-Core HTR Benchmark using MCNP5 and RELAP5-ATHENA
Xiao Cheng	University of Nebraska Lincoln			1,000,000	Exploration of Structural and Catalytic Properties of Gold Clusters
Rong Tian	Institute of Computing Technology, Chinese Academia of Sciences			900,000	Petascale simulation of fracture process
Praveen Ramaprabhu	University of North Carolina			862,160	Simulations of turbulent mixing driven by strong shockwaves
Aytekin Gel	ALPEMI Consulting			600,000	Mitigation of CO2 Environmental Impact Using a Multiscale Modeling Approach
Thomas Gielda	Caitin Inc.			500,000	Parallel Computing performance Optimization for Complex Multiphase Flows in Strong Thermodynamic Non-equilibrium
Xiaolin Cheng	ORNL			500,000	Scalable bio-electrostatic calculation on emerging computer architectures
Cristiana Stan	Center for Ocean-Land-Atmosphere Studies			500,000	Simulations of Antropigenic Climate Change Effect Using a Multi-Modeling Framework
David Rector	PNNL			400,000	Solid-liquid tank mixing using the implicit lattice kinetics method
Kai Germaschewski	ORNL			200,000	Load balancing particle-in-cell simulations
Don Lucas	LLNL			100,000	Uncertainty Quantification of Climate Sensitivity

PI	Affiliation	2010 Allocation	Carryover to 2011	New 2011 Allocation	Project Name
Masako Yamada	GE Global Research			100,000	Engineered icephobic surfaces
Atul Jain	University of Illinois			30,000	Land Cover and Land Use Change and its Effects on Carbon Dynamics in Monsoon Asian Region
Paul Sutter	University of Illinois			5,000,000	Exploring the origins of galaxy cluster magnetic fields