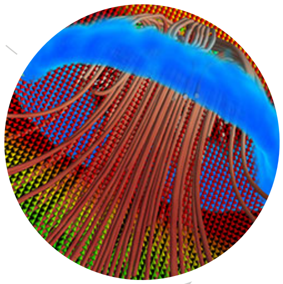
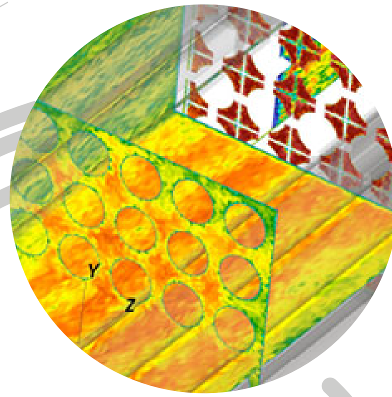
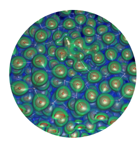






Argonne **Leadership Computing** Facility

Operational Assessment Report

March 2013

Argonne NATIONAL LABORATORY

On the cover

Clockwise from top left:

Billion-atom reactive molecular dynamics simulation of nanobubble collapse in water near a ceramic surface under shock compression. Computed on Intrepid.

2012 INCITE Project: *Petascale Simulations of Stress Corrosion Cracking*
PI: Priya Vashishta, University of Southern California

Velocity magnitude distribution in a flow through the 25-pin swirl-vane spacer grid of Matis benchmark. Computed on Intrepid with Nek5000 and visualized on Eureka with VisIt.

2012 INCITE Project: *Advanced Reactor Thermal Hydraulic Modeling*
PI: Paul Fischer, Argonne National Laboratory

Electron density obtained from a density functional theory (DFT) calculation of lithium oxide (Li_2O) performed with the GPAW code. Computed on Intrepid.

2012 Early Science Project: *Predictive Materials Modeling for Li-Air Battery Systems*
PI: Larry Curtiss, Argonne National Laboratory

Mira, the ALCF's 10-petaflops IBM Blue Gene/Q.

This page intentionally left blank.

Contents

Figures

Tables

This page intentionally left blank.

Executive Summary

As one of two high performance computing facilities in the nation supported by the Department of Energy's Office of Advanced Scientific Computing Research (ASCR) for open science, the Argonne Leadership Computing Facility (ALCF) provides the computational science community with world-class computing resources to pursue breakthroughs in science and engineering. In operation since 2006, the ALCF provides its users with expertise and assistance to ensure that every project achieves top performance on its resources.

This year's facility operations success story is all the more impressive considering that the support team not only ran a production system, but also installed Mira, a 48-rack IBM Blue Gene/Q system that is expected to surpass 10 petaflops. The major resource at the ALCF is an IBM Blue Gene/P system, Intrepid, and two development systems, Challenger and Surveyor. Intrepid is a 40-rack, 557-teraflops system used for production science and Surveyor and Challenger are each approximately 14-teraflops, single rack systems used for tool and application porting, software testing and optimization, and systems software development.

ALCF has had an outstanding year, meeting or exceeding all metrics set for the facility. ALCF delivered 1.2 billion core hours of compute time between January 1, 2012 and December 31, 2012, with almost 796 million of those core hours being used by capability jobs (single job using 20% or more of the machine). During that same time period, the science done on the machine produced at least 180 publications, in all major areas of interest to DOE.

In 2012 ALCF supported 919 active users and 235 active projects from universities, national laboratories, and industry. The resources are allocated via four different merit-based award programs: Innovative and Novel Computational Impact on Theory and Experiment (INCITE), Advanced Scientific Computing Research Leadership Computing Challenge (ALCC), Early Science Program (ESP), and Director's Discretionary Allocation (DD). The INCITE category represents 60% of the time allocated on the machine, and its users are selected via a peer review process led by ALCF and OLCF, Oak Ridge National Laboratory's Leadership Computing Facility. The ALCC accounts for 30% of the resources and is allocated by ASCR, with the remainder of the resource allocated to the DD and ESP programs.

The annual Operational Assessment review of ALCF by ASCR provides the facility with an opportunity to receive external feedback on ways to improve the operation of the facility. The review takes into consideration agreed-upon metrics and reports describing the operation of the facility. The report is organized into six sections. The remaining sections address the 2012 OAR metrics and present User Support Results, Business Results, Strategic Results, Innovation, Risk Management, and a Summary of the Proposed Metric Values for future OAR's.

Area	Metric	2012 Target	2012 Actual
User Results	User Survey – **Overall Satisfaction**	3.5/5.0	4.4/5.0
	User Survey – **User Support**	3.5/5.0	4.5/5.0
	User Survey – **Problem Resolution**	3.5/5.0	4.5/5.0
	User Survey – **Response Rate**	25%	33.2%
	% User Problems Addressed Within Three Working Days	80%	94.4%
Business Results	Intrepid Overall Availability	90%	95.9%
	Intrepid Scheduled Availability	95%	98.5%
	Intrepid INCITE Capability Usage (20%-100% of system)	40%	63.1%

Table ES.1: Summary of the Target and Actual Data for the Previous Year (2012) Metrics.

Section 1. User Support Results

Are the processes for supporting the customers, resolving problems, and outreach effective?

ALCF Response

ALCF has established processes in place for effectively supporting customers, resolving problems, and performing outreach. The sections below document these processes, the effectiveness of the processes, and improvements to the processes that were implemented during CY2012.

User Services Overview

ALCF User Services continues to execute the following responsibilities:

- Providing a help desk,
- Stewarding key demographic information and surveys,
- Managing user communications,
- Coordinating outreach and education efforts.

Help Desk

The core business of the technical side of Users Services is to provide tier-1 and limited tier-2 help desk support for facility users. This help desk investigates, categorizes, and routes email, web, and phone support requests. In addition, the help desk administers account, project, and allocation systems. Within ALCF, User Services acts as the user representative, bringing user-centric concerns (e.g., quotas, upgrades, accounting policy, etc.) to the attention of ALCF staff members outside User Services.

User Services also provides the essential function of establishing ALCF user accounts. This includes helping the users meet the various requirements that Argonne and DOE have in terms of Foreign National Access paperwork, Master Institution Agreements, User Acknowledgements of such agreements, etc.

Demographic Information and Surveys

User Services provides extensive reporting on how the facility resources are being used. Detailed weekly usage reports are sent to INCITE and ALCC PIs and proxies with textual as well as graphical information about their project. These reports include jobs that were run on Intrepid, job sizes (bucketed by size of run on Intrepid), users who ran these jobs, actual allocation burn rate (relative to steady burn rate), etc.

ALCF staff employ aggregated information from databases stewarded by User Services. User Services also participates in division-wide efforts to collect information from various other data sources. These data are often provided for ALCF, Argonne, and DOE leadership.

In addition to demographic information, User Services manages the yearly ALCF User Survey and ad hoc, internal surveys.

Communications

The core business of the media side of Users Services is the stewardship and dissemination of science stories. These stories provide an account of the scientific or engineering accomplishments achieved using ALCF resources and staff expertise. These accounts are concise and written for targeted audiences that differ from the user-authored, peer-reviewed journals and conference proceedings that report in technical depth to domain experts.

ALCF disseminates these accounts through various channels via both traditional and new media using in-house editorial and communication experts. In addition, User Services leverages the expertise of Argonne's Communications, Education and Public Affairs (CEPA) Division. CEPA provides additional talent (such as graphic design and print management) and provides regional, national, and international press relationship management.

User Services is also responsible for the organization and administration of ALCF's online documentation. These responsibilities include leading a documentation committee, coordinating documentation content from technical experts across the organization, managing the information architecture for the website, and providing content and web analytics and reporting. The ALCF website is built upon Drupal software, and User Services is primarily responsible for the software administration and maintenance of this website.

Finally, User Services is tasked with coordinating resources for, and providing assistance with, general-purpose text editing, layout design, and video production. This encompasses coordination and development of internal and external technical publications, posters, and scientific or promotional videos.

Outreach and Education

User outreach activities are structured around the INCITE and ALCC award cycles. During these cycles, ALCF targets two different sets of communities: 1) "traditional" academic and laboratory researchers, and 2) industry researchers and engineers. Promotional materials, both print and web, are designed to attract scientific researchers with high-performance computing experience to apply for an award at ALCF. There are workshops that are designed to help researchers understand the award processes and to formulate their applications to meet the technical requirements of an award.

Once a researcher wins an award, messaging and outreach turn to support of the project goals through conference calls, quarterly reports, and appropriate e-mail messaging. Additionally, ALCF hosts workshops for awardees to ensure the success of their projects. To keep current with user perspectives, the facility makes use of the ALCF User Advisory Council. The Council proportionally represents ALCF users and provides a sounding board for issues with and enhancements to the facility.

User Services Results

Help Desk Results
To address issues raised by users who call the ALCF help desk, all members of the User Services team, not just specific help desk staff, have been trained and are available to answer phone calls. In addition, User Services worked with Argonne's Computing and Information Systems (CIS) Division and added a new phone service, Advanced Call Distribution, to manage calls. This provides call reporting as well as provides a better mechanism for call distribution.

To improve the "% of problems addressed in 3 working days" metric in 2012, ALCF User Services put into effect reminder e-mail scripts and help desk procedures. This improved the metric from 90% to 94%. ALCF continues to monitor the long-term success of these additional procedures.

In preparation for the release of Mira and the end of the ALCF-2 project, User Services has engaged in continuous improvement efforts on four fronts: 1) internal account creation processes, 2) ticket response and routing, 3) startup of INCITE and ALCC projects, and 4) end-user account management tools. Below are results from two of these four efforts:

- Through measurement, User Services found two large bottlenecks in internal account creation: foreign national status processing (Argonne's 593 process) and cryptocard shipments. ALCF leveraged shipping services with tracking information and increased communication between User Services help desk staff and Argonne administrative staff responsible for processing the 593 forms.

- User Services has created a seven-point checklist for all help requests. Staff are currently measuring their performance at completing this checklist via tagging and reporting via reports in RT, the ALCF ticketing system. During the ALCF-2 transition to operations, User Services plans to use this system to provide meaningful measurement for ALCF staff and leadership.

Demographic Information Results
Part of the continuous improvement efforts in User Services has been a focus on improving the accuracy and quality of user demographic information. This has improved the performance and accuracy of analytics and reporting.

User Communication Results
Navigational and content improvements were made to the online documentation for the Blue Gene/P from February through June. Analysis of user support tickets using Google Analytics identified what content the users were having difficulty finding.

User Services led the ALCF-2 documentation efforts. Two user focus groups were assembled to obtain feedback on current documentation and advice on moving forward with building the ALCF-2 documentation. The focus groups found that the current ALCF documentation serves advanced users well but needs enhancements for new users, such as online "getting started" documents, how-to information sheets, and online tutorials for new users. Additionally, the focus group study found advanced users and novice users use HOWTO and tutorials differently and both groups need these types of documentation.

Several months were spent improving website maintenance practices by establishing a development/stage/production server environment using svn. This new server environment has reduced site maintenance outages from several hours down to five minutes. Another benefit is the ability to stage changes before activating them, to make sure that any new modules or features do not break the site.

Outreach and Education Results
ALCF changed the 2012 Welcome Call format. In addition to the usual topics, PIs were notified about ALCF's transition plan from Intrepid to Mira and advised about the best ALCF events for their project members to attend. The results have been especially positive among new PIs. They have indicated that the call has been helpful and has cleared up questions they've had about the INCITE and ALCC processes.

User Services staff worked collaboratively during the INCITE and ALCC startup times. Mini-checklists were employed and messages simplified in PI and proxy e-mail to increase understanding and response rate.

Argonne's Computing, Environment, and Life Sciences (CELS) associate laboratory directorship recently hired a media specialist who works with ALCF to place ALCF science success stories into periodicals and media channels significant to the DOE-funded programs.

Operations Support Results

Notable User Support efforts in ALCF's Operations team primarily focused on storage and storage management. Moving to a SAN storage infrastructure for ALCF critical management nodes decreased user visible downtime. This allows the facility to recover from failed hardware more easily and quickly. Adding two new GPFS cluster token manager nodes significantly reduced user job boot failures due to timeouts - the single most common source of downtime and job interrupts on Intrepid in 2011 and 2012. (To be clear, this is a measure of the number of events, not the actual impact of the events.)

Operations enabled quota enforcement on intrepid_home, allowing ALCF to maintain a consistent level of disk space and file system performance. At the same time, Operations increased the total storage space of the intrepid-fs0 filesystem by adding another 1PB of disk. Finally, Operations implemented a preemptive replacement strategy for disk drives on the file systems that serve Intrepid, Challenger, Surveyor, and Eureka. As part of this strategy, drives that exhibited increased latency or medium disk errors were replaced.

Application Support Results

HACC
HACC (Hardware/Hybrid Accelerated Cosmology Code) is a framework designed initially on Roadrunner. After ALCF ported this to Mira and Sequoia, this application achieved a

sustained performance of 13.94 PF (69.2% of peak) on Sequoia. HACC is a software framework that simulates the behavior of galaxies on a cosmological scale. Its purpose is to help scientists to reveal the nature of dark matter and dark energy.

The code is written in modular form splitting the long-range and short-range force calculations. The long-range force computation is independent of architecture (MPI-based) while the swappable short-range forces can be optimized for different architectures and programming models. The application runs on ALCF and OLCF and has time allocated via the ESP, INCITE, and ALCC programs.

Five ALCF application and visualization staff, as well as a dedicated postdoc, provided help in porting the application to BG/Q, optimizing the short-range forces for Gordon Bell submission, incorporating parallel I/O, and analyzing the results. A few highlights are provided below:

- Designed, implemented, and tuned the RCB-tree used to organize and balance the short-range force calculations while ensuring data locality.
- Merged the initializer (the code that generates initial conditions with specified statistical properties) with the main code, and made the initializer able to use the pencil-decomposed FFT. This is necessary for running at scale on the BG/Q.
- Designed and implemented a custom pool allocator (called bigchunk) necessary for mitigating the effects of memory fragmentation. This is necessary for running on the BG/Q for long periods of time (because otherwise the code would exhaust the available memory due to fragmentation effects).
- The short force evaluation kernel was designed such that memory operations latencies could be entirely hidden by computations. The kernel was implemented with QPX (Quad FPU) intrinsics to aid the compiler to schedule instructions and registers with no penalties. The standalone kernel runs at 82% of peak performance on practical sizes of interaction lists.
- Incorporated GLEAN's topology-aware I/O methodology, which exploits the 5-D torus architecture of BG/Q, into HACC's I/O infrastructure.
- Achieved up to 170 GB/s I/O performance for checkpoints on 32 racks of Mira BG/Q. This is ~90% of the achievable peak bandwidth. This performance is ~18x improvement over the previous HACC I/O implementation based on MPI-IO single shared file.
- The new I/O mechanism has been integrated into the HACC simulation and has been used for production runs for the past two months. Over 6 PB of data has been written and read using this I/O mechanism.
- The I/O mechanism finds a middle ground between the number of processes (M) and files (N) based on the characteristics of the underlying filesystem for improved performance. This will enable the I/O mechanism to easily adapt to a parallel filesystem, including GPFS (Mira- ANL, Hopper-NERSC) and Lustre (Titan- ORNL, Hopper-NERSC). On Mira, a file is created for all processes in an I/O node (one file for every 128 nodes).

Dynamics of Conformational Transition in Thermo-Sensitive Polymers and Hydrogels

Members of this ALCC project encountered difficulties when trying to use the latest release of the LAMMPS classical molecular dynamics (MD) software package for production runs on Intrepid. ALCF staff helped them identify two opportunities for reducing the time-to-solution with the LAMMPS code:

1. Selecting the PPPM algorithm instead of the standard Ewald sum algorithm that was being used. This led to a 5X reduction in the MD time step.
2. Using the latest version of LAMMPS, which supports OpenMP. This led to a cumulative 6X improvement (20% improvement over pure MPI with the same number of cores in the best case scenario).

These improvements would allow the group to run their calculations, but only in the complete absence of I/O.

ALCF staff then suggested that the project consider moving to NAMD, a more scalable classical MD code. The group found that this move would be feasible for their simulations. ALCF staff then built the latest Pthread and memory-optimized version of NAMD (v. 2.9) for them and fixed a source code bug for the input of the fixed atom file. The staff also worked with the group to optimize the NAMD simulation configure file, including hybrid load balance frequency and number of I/O nodes for parallel I/O. Finally, ALCF staff fixed a severe memory problem at large scale runs by increasing the DCMF_RECFIFO (DCMF is the low level communication API on BG/P) value.

Currently the NAMD simulation can scale well to 16 racks for a polymer system of 8 million atoms. The team members are using OPLS and CHARMM force fields to simulate large-scale conformation changes of polymers. NAMD is over 10x faster than LAMMPS for their calculations in the multi-million atom range and will not only make their scientific campaign feasible but also enable adding scope they thought was previously impossible.

NAMD

NAMD represents a large fraction of allocations for INCITE and ALCF. In 2012, the ALCF staff made some significant contributions to NAMD. First Charm++ and NAMD were integrated with the low level Blue Gene messaging layer. This gave a 40% speedup. Additionally, and substantially, a set of Multiple-Copy Algorithms (MCAs) for quantitative biological molecular dynamics simulations were implemented in NAMD2.9 on top of Charm++6.4.0, in collaboration with Parallel Programming Laboratory at UIUC. The progress in this project will be used by two 2013 INCITE projects “Studies of Large Conformational Changes in Biomolecular Machines”(Benoit Roux) and “Thermodynamics of Binding Biomass to Cellulases for Renewable Fuel” (Micheal Crowley). Here this work is described in more detail.

The standard Charm++ was enhanced to support parallel-parallel simulations with multiple copies by splitting communicators. The replica exchange communication is implemented through new APIs in both Charm++ and the NAMD Tcl scripting interface to allow the MCAs to be extended to non-MPI communication layers.

The current implementation of MCAs supports most of the popular equilibrium biological free energy calculations with free energy perturbation and umbrella sampling, temperature exchange MD simulations, and nonequilibrium MD simulations. Large-scale benchmark results were created on BGQ Mira. For example, the nonequilibrium string method has been run on Mira to support extremely scalable multiple copy algorithm simulations. The test has been done using both non-SMP mode and SMP mode and both modes scale up to 32 racks (32,768 nodes) on Mira.

Simulating Regional Climate at Convection Permitting Resolution
This 2011 ALCC project was challenged transitioning from NCAR resources to Intrepid. Their workflow and scaling were very different. ALCF Catalysts reached out to them when they thought they would not be able to use their allocation.

ALCF Catalysts started by working out some minor running and queuing problems. They then led the project team through a systematic evaluation of scaling and checkpoint overhead to determine the most efficient job size and job length for each of the runs in the series. After working closely with ALCF staff at the 2012 *Leap To Petascale* hands-on workshop, the project team were able to forge new ground with a significant simulation at 4km resolution -- much more detailed than previous 36km and 12km runs at NCAR. The project was not only able to use their entire allocation, but also their success prompted them to request more time.

ALCF Support Metrics for 2012

Tables 1.1 and 1.2 show a summary of all user support metrics. Further details are provided in the respective metric sections below.

2011 and 2012 Metric Comparison		Value for 2011	Target for 2012	Value for 2012
Number Surveyed		959	N/A	920
Number of Respondents (Response Rate)		278 (29.0%)	(25%)	305 (33.2%)
Overall Satisfaction	Mean	4.4	3.5	4.4
	Variance	0.5	N/A	0.5
	Std Dev	0.7	N/A	0.7

Problem Resolution	Mean	4.5	3.5	4.5
	Variance	0.4	N/A	0.4
	Std Dev	0.6	N/A	0.6
User Support	Mean	4.5	3.5	4.5
	Variance	0.5	N/A	0.5
	Std Dev	0.7	N/A	0.7
Tickets Addressed in 3 days (business)		90.2%	80.0%	94.4%

Table 1.1: Comparison of Key Metrics Between 2011 and 2012 Results

2012 Metrics By Program		INCITE	ALCC	INCITE + ALCC	DD	All
Number Surveyed		368	37	405	515	920
Number of Respondents		130	23	153	152	305
Response Rate		35.3%	62.2%	37.8%	29.5%	33.2%
Overall Satisfaction	Mean	4.5	4.4	4.5	4.3	4.4
	Variance	0.4	0.6	0.4	0.5	0.5
	Std Dev	0.6	0.8	0.7	0.7	0.7
User Support	Mean	4.6	4.4	4.5	4.4	4.5
	Variance	0.4	0.4	0.4	0.6	0.5
	Std Dev	0.6	0.6	0.6	0.8	0.7
Problem Resolution	Mean	4.6	4.4	4.6	4.5	4.5
	Variance	0.3	0.3	0.3	0.5	0.4
	Std Dev	0.6	0.6	0.6	0.7	0.6
All Questions	Mean	4.6	4.4	4.5	4.4	4.5
	Variance	0.4	0.4	0.4	0.6	0.5
	Std Dev	0.6	0.7	0.6	0.8	0.7

Table 1.2: Comparison of Key Metrics between Different Allocation Programs

User Survey

In order to continue to improve the 2012 and 2013 ALCF user surveys, ALCF hired a consultant to develop and instrument surveys. The survey evaluation took on four different phases:

1. Literature review and summary of findings

2. Analysis and evaluation of 2010 and 2011 surveys and related data and reports
3. Planning of 2012 survey
4. Development, testing, and implementation the 2012 survey, and analysis and write-up of the survey data

The consultant collected a series of materials for decision makers to use as a reference, and then analyzed the 2011 questions and made recommendations for improvement. The consultant then developed a plan of action on how to change the CY2012 survey. ALCF leadership was polled for any new questions they wished to be added to the survey. The survey was tested with the User Services team; then by the members of the User Advisory Council (UAC); and then vetted with the DOE program manager.

The ALCF user survey was conducted from December 7, 2012 to January 21, 2013. Users with an active status account from January 1, 2012 to December 3, 2012 were included in the survey.

Likert Scale and Numeric Mapping

All but one Likert Scale question in the ALCF user survey used a six choice scale. This has been a standard for surveys because: 1) it is easy for the users to quickly place the response to a question within a range of options, 2) it can be mapped to a numeric scale and given a certain sample size, be used with a normal distribution. This allowed use of off-the-shelf functions in Excel to determine variance and standard deviation. ALCF maps the Likert Scale in this way:

Statement	Numeric
Strongly Agree	5
Agree	4
Neutral	3
Disagree	2
Strongly Disagree	1
N/A	(No Value)

The overall satisfaction question used a five choice scale:

Statement	Numeric
Excellent	5
Above Average	4
Average	3
Below Average	2
Poor	1

Changes to the User Survey

As part of the process to improve the user survey, ALCF management raised the following questions:

- Is the response rate low and how can it be increased?
- Are there any questions that are poorly formulated?
- How can ALCF take advantage of the comments sections of the survey?

The survey consultant studied responses from previous years and found no correlation between the time period chosen to administer the survey and a more positive or more negative result. According to the literature, this suggests that extending the response period would not change the overall distribution of positive/negative responses.

The 2012 survey was instrumented differently than the survey for 2011 in four ways:

1. A series of yes/no questions was created to eliminate certain questions for subsets of users, thus improving average time to complete the survey.
2. Possible issues with double-barreled questions were tested by re-asking the two possible choices separately.
3. The survey was tested with the ALCF UAC.
4. The introductory message to PI's and Discretionary users was tailored to encourage survey participation by respondents.

In addition to the above changes, a question was added that was recommended by the CY2011 OAR reviewers: *"As you may know, ALCF is required at times to shut down its computer for routine maintenance. This must occur on a weekday. Use the drop downs below to rank days."* Below this question, users were presented with drop down lists allowing them to rank a particular weekday on this scale: "best, good, neutral, bad, worst".

Other minor changes: ALCF replaced the open-ended questions at the beginning of the survey with multiple choice questions with an optional comment section, based upon feedback from the consultant. For all comment sections, ALCF provided checkbox options for the user to classify the feedback as praise, suggestion, problem, or complaint, in order to expedite calculations of summary results. Finally, ALCF consolidated all questions related to Catalysts into the same section.

Survey Response Results

Survey sent to 920 Users, PIs, and Co-PIs (2012)	2011 Response Rate (%)	2012 Response Rate (%)	Target Response Rate (%)	Total # of Responses
Response Rate	29.0%	33.2%	25%	305

Overall, the survey response rate increased from 29.0% to 33.2%. ALCF experimented with a few ways to increase the response rate. It was discovered during the UAC focus group that some of the users were concerned with who had access to their specific responses. ALCF created messaging in the first page of the survey that assured them that only staff handling the survey data would see information that identified them as the respondent. Sending personalized e-mails from the Division Director improved the PI response rate in 2011 and ALCF implemented this practice again. ALCF sent a tailored e-mail to increase responses to the most underrepresented group – discretionary projects. The response rate for this group increased from 20% in 2011 to 29% in 2012.

1.1 User Support Metrics

Survey Area	2011 Target	2011 Actual	2012 Target	2012 Actual
Overall Satisfaction Rating	3.5/5.0	4.4/5.0	3.5/5.0	4.4/5.0
Average of User Support Ratings	3.5/5.0	4.5/5.0	3.5/5.0	4.5/5.0

Overall Satisfaction and User Support ratings did not change. For three years running, the specific survey questions that relate to the metrics have remained unchanged to ensure proper trend analysis.

Day of the Week for Scheduled Maintenance

Users were asked to rank the best day of the week for scheduled system maintenance. The following code was used to generate numerical values for the user responses: Best was coded as +2; Good was coded as +1; Neutral = 0; Bad = -1; and Worst = -2. Summing the resulting re-coded responses for the 305 respondents who answered this question produced the following results:

Day of the Week	Monday	Tuesday	Wednesday	Thursday	Friday
Sum of results	106	-10	-42	-51	-31

Given this response, ALCF will continue to perform preventive maintenance on Mondays.

1.2 Problem Resolution Metrics

Survey Area	2011 Target	2011 Actual	2012 Target	2012 Actual
% of problems addressed in 3 working days	80%	90.2%	80%	94.4%
Average of problem resolution ratings	3.5/5.0	4.5/5.0	3.5/5.0	4.5/5.0

The survey results for problem resolution from 2012 did not vary significantly from the 2011 results. However, there was a significant increase in the percentage of addressed tickets due to processes put in place in 2011 and revised in 2012.

ALCF defines a ticket as "addressed" once the following is true: The ticket has been accepted by a staff member; the problem has been identified; the user has received notification of this; and the staff member is working on a solution or has found one. The ALCF has a specific process in the problem tracking system for addressed tickets. An ALCF staff member modifies a special field, called addressed, once the definition of addressed has been met.

Table 1.3 shows the difference in categorization of tickets between 2011 and 2012. Ticket categorization within RT, the ticketing system used by ALCF, was implemented in June 2011. Therefore, the YTD numbers in the CY2011 OAR report were an assumption of how tickets were categorized based on the internal ALCF group member who handled those tickets. After the implementation of ticket categorization within the ticketing system, ALCF staff members were told to retroactively categorize their tickets for CY2011 based on the new system. The current CY2011 categorization of tickets provides more accurate data for ticket categorization.

In 2012 ALCF added categories in order to track key issues. ALCF implemented quotas on user file systems on Intrepid in mid-2012. ALCF has also implemented quotas on all file systems for Vesta, Mira, and Cetus. Visualization was also made a separate category in response to 2010 and 2011 survey results. User Services wanted to draw out questions about quotas and visualization to see the impact on users. Finally, Allocations were added, as ALCF is moving to a different workflow for addressing Discretionary allocation requests.

Categories	CY2011	CY2012
Access	516	986
Accounts	1,394	1,679
Applications Software	217	177
I/O & Storage	104	167
Misc.	222	202
System	309	208
Visualization (previously grouped with Data Transfer)	200	41
Allocations	N/A	39
Bounces	N/A	239
Data Transfer	N/A	24
Quota Management	N/A	40
TOTAL TICKETS	**2,962**	**3,802**

Table 1.3: Difference in Categorization of Tickets Between 2011 and 2012

1.3 User Support and Outreach

1.3.1 User Advisory Council

The UAC provided direct, actionable feedback on the 2012 User Survey; advised on messaging about the progress of ALCF2; and engaged in a dialogue about new features to ALCF2, such as project quotas. The UAC advised ALCF on the following issues:

- Advice: Clearly define the role of the quarterly reports. Results: 2013 INCITE calls included a discussion on how ALCF uses the quarterly reports to understand and to help projects.
- Advice: Provide more frequent messaging about early access systems (Vesta and Mira). Result: Increased communication with the early access users.
- Advice: Make sure the user understands the policies of quotas. Result: Greater collaboration between Operations and User Services on quota management.
- Advice: Provide language in ALCF's 2012 survey that ensured identities were kept private to survey staff. Result: Increased response rate on the 2012 survey.

1.3.2 Academic, Non-Profit, and Government Laboratory Collaboration

ALCF continues to develop relationships with these more traditional communities. ALCF users play a significant role in this outreach area – referrals are an important source of new projects. In addition, ALCF Catalysts reach out into their scientific communities and use materials (text-based e-mails, fliers, and websites) created by User Services. ALCF also sends various staff and leaders throughout the world to present to organizations, often taking advantage of geographic proximity (e.g., Bay Area, Greater Boston, etc.) of educational institutions.

1.3.3 Industry Collaboration

ALCF made significant progress in improving industrial relations during 2012 by working with other Argonne divisions to present the full breadth of Argonne's capabilities. In working with Toyota Research Institute, for example, ALCF leveraged the initial contact that was focused on computing into a broader discussion involving the Center for Nanoscale Materials, the Materials Science Division, and the Chemical Sciences and Engineering Division. Such collaborations provide a stronger case for industry to work with ALCF. They also enable ALCF to leverage computational science expertise in other divisions and ease some of the workload on ALCF staff, especially the Catalysts.

ALCF has also been aided in industrial relations by a new partnership with the Technology Development and Commercialization Division (TDC) at Argonne. ALCF works closely with TDC on all existing and potential industrial engagements. TDC has cited ALCF's tight coordination as a model for other Argonne divisions.

1.3.4 2012 Workshops

ALCF hosted five major events in 2012. The *Winter Workshop* was a merged event that both assisted new and potential INCITE users as well as provided expert users an opportunity to directly engage with ALCF. The *Proposal Writing Webinar* used Adobe Connect to stream and record a live presentation and Q/A on INCITE proposal writing. The two *Early Science Program Workshops* were created to prepare ESP projects for early access to Vesta and eventually Mira. *Leap to Petascale* was the yearly workshop tailored to INCITE applicants and potential users who need to scale up their applications to run efficiently on the production Blue Gene systems. Finally, the *Blue Gene Consortium Summit* was a gathering of representatives from other institutions, national and international, which manage Blue Gene systems.

- *Winter Workshop 2012* – January 23-26
- *INCITE Proposal Writing Webinar 2012* – January 24
- *Early Science Program (ESP) Workshops* – March 19 and April 30
- *Leap to Petascale 2012* – May 22-25
- *Blue Gene Consortium Summit* – October 2

Per recommendations from the prior year, ALCF staff strived to build all event schedules with more hands-on time. In addition, ALCF User Services reached out to users prior to events, such as the *Winter Workshop*, to make sure they had the accounts and access needed to successfully use the systems once onsite.

Streaming Media, Webinars, and Video Conference Workshops

In 2012, ALCF experimented with streaming and storing the *Proposal Writing Webinar*. In addition, Argonne video staff assisted ALCF and attempted to shoot video of a roundtable meant for streaming storage.

The first experiment was acceptable. Its webinar format lent itself well to long-term storage: 1) The content was relatively scripted, 2) any non-scripted parts were significantly edited in post-production, and 3) the stream was a simple, high-quality audio overlay on PowerPoint slides. This session was recorded to an Adobe Connect format and hosted on a site contracted by Argonne National Laboratory through Adobe.

The second experiment was not acceptable and never posted. Problems included poor lighting, a single camera was not enough for ad-hoc discussions, the staff was not properly trained on presenting in this format, and extremely poor audio. Witnessing this event live, the panel session was interesting and highly interactive. Watching the recording greatly diminished the effect and the various technical problems made it unwatchable.

In 2013, ALCF has plans underway to take advantage of a more collaborative, real-time technology called the Cisco Jabber client on ESnet to accommodate a "getting started" videoconference workshop. This effort will be aimed at creating a roundtable-like, highly collaborative setting by limiting the number of video streams and sharing video from all participants. It is hoped that this technology can create a virtual space that mimics the in-person part of early-January workshops.

1.3.5 2012 Publications

ALCF publications are tailored to address targeted audiences ranging from those who are very knowledgeable about computational science (e.g., researchers, scientists) to those who need just the basics (e.g., Congress, outside media, students who visit the ALCF.) User Services has also committed to writing articles tailored to web-based distribution instead of traditional print media. This is in line with resources invested in ALCF's Drupal content management system.

This year, ALCF changed how the Industry Brochure and Science Highlights were handled. Instead of one User Services staff solely assigned to each key publication, the entire team of media-related staff (including CEPA graphic artists) worked collaboratively on its development. Instead of asking for feedback ad hoc, a schedule was developed and formal reviews started from the beginning of the process. The team had input on audience, writing guides, and design. The media team members then wrote materials based upon a variety of sources and vetted the text with catalysts, users, and ALCF scientific managers. This change ensured that the full complement of User Services skills and experience was brought to bear throughout the development of the most important publications.

In 2012, ALCF User Services found that a significant amount of work was being done on the ALCF Newsbytes Newsletter with very little indication that the piece was being read. Additionally, the piece was developed for print format, so moving stories onto the web was an afterthought. ALCF has stopped production of the print-centric Newsbytes and will be releasing a monthly Newsbytes e-mail.

After researching existing technologies available at Argonne, it was decided to use a lab-approved service called MailChimp to create monthly electronic newsletters. Besides the ALCF users, the general public and interested parties can then subscribe themselves or request to be subscribed via the ALCF website. Communications dealing with machine-specific issues will be routed to the "notify" e-mail lists. Communications dealing with conference information, highlighted ALCF stories, and key events and workshops will be directed to the "info" list.

Publication	Frequency	When
ALCF Weekly Update E-mail	Weekly	Thursday
Highlights for ASCR Computing News Roundup	Monthly	1st Week
ALCF Newsbytes Newsletter *	Quarterly	Jan, Apr, Jul, Oct*
ALCF Fact Sheets	As needed	As needed
ALCF Annual Report	Yearly	Mar
ALCF Science Highlights	Yearly	Sep
ALCF Calendar	Yearly	Nov
ALCF User Packets	As needed	As needed
ALCF Industry Brochure	Yearly	June

Table 1.4: Key Publications of the ALCF

* Newsbytes Newsletter is moving to an electronic format and will be issued monthly, beginning February 2013.

Conclusion

As a user facility, ALCF is focused on ensuring the success of all the facility users and customers. To this end, during CY2012 ALCF improved processes, revised the user survey, acted on feedback from the User Advisory Council, and refocused dissemination efforts from print-centric to web-centric writing and design. As such, ALCF continues to effectively support customers, resolve problems, and perform outreach.

This page intentionally left blank.

Section 2. Business Results

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

ALCF Response

For those measures where there are concrete metrics, e.g., system availability, INCITE hours delivered, and capability hours delivered, ALCF has exceeded the metrics. For the reportable areas, such as Mean Time to Interrupt (MTTI), Mean Time to Failure (MTTF), and system utilization, ALCF is on par with the other facilities and has demonstrated acceptable performance.

ALCF tracks hardware and software failures and analyzes their impact on the user jobs and metrics as a significant part of its improvement efforts.

Intrepid (40K-node, 160K-core BG/P)				
	CY 2011		CY 2012	
	Target	Actual	Target	Actual
Scheduled Availability	95%	97.8%	95%	98.5%
Overall Availability	90%	94.0%	90%	95.9%
System MTTI	N/A	10.05 days	N/A	8.55 days
System MTTF	N/A	17.95 days	N/A	21.29 days
INCITE Usage	732M	876.6M	732M	803.8M
Total Usage	N/A	1.20B	N/A	1.20B
INCITE Capability Usage in percentage*	N/A	57.0%	40%	63.1%
INCITE Capability Usage in core-hours*	300M	499.7M	N/A	507.4M
System Utilization	N/A	88.8%	N/A	87.6%

***ALCF INCITE capability metric changed from core-hours in CY2011 to percentage in CY2012**

Table 2.1: Summary of All Metrics Reported in the Business Results Section

2.1 Resource Availability

Overall availability is the percentage of time a system is available to users. Outage time reflects both scheduled and unscheduled outages. For HPC Facilities, scheduled availability is the percentage of time a designated level of resource is available to users, excluding scheduled downtime for maintenance and upgrades. To be considered a scheduled outage, the user *community must be notified of the need for a maintenance event window no less than 24 hours in advance of the outage (emergency fixes). Users will be notified of regularly scheduled maintenance in advance, on a schedule that provides sufficient notification, and no less than 72 hours prior to the event, and preferably as much as seven calendar days prior. If that regularly scheduled maintenance is not needed, users will be informed of the cancellation of that maintenance event in a timely manner. Any interruption of service that does not meet the minimum notification window is categorized as an unscheduled outage. A significant event that delays a return to scheduled production will be counted as an adjacent unscheduled outage. Typically, this would be for a return to service four or more hours later than the scheduled end time. The centers have not yet agreed on a specific definition for this rare scenario.*

This section reports on measures that are indicative of the stability of the system and the quality of the maintenance procedures.

2.1.1 Scheduled and 2.1.2 Overall Availability Summary

Intrepid has been in full production since February 2009. ALCF has agreed, with the DOE Program Manager, to metrics of 90% overall availability and 95% scheduled availability, which is consistent with OLCF and NERSC. Table 2.2 summarizes the availability results:

Intrepid (40K-node, 160K-core BG/P)				
	CY 2011		CY 2012	
	Target	**Actual**	**Target**	**Actual**
Scheduled Availability	95%	97.8%	95%	98.5%
Overall Availability	90%	94.0%	90%	95.9%

Table 2.2: Availability Results

The remainder of this section covers significant availability losses and responses to them, for both scheduled and overall availability data. Details on how the calculations are handled can be found in Appendix A.

Explanation of Significant Availability Losses on Intrepid

This section provides a brief description of the causes of major losses of availability, annotated in Figure 2.1.

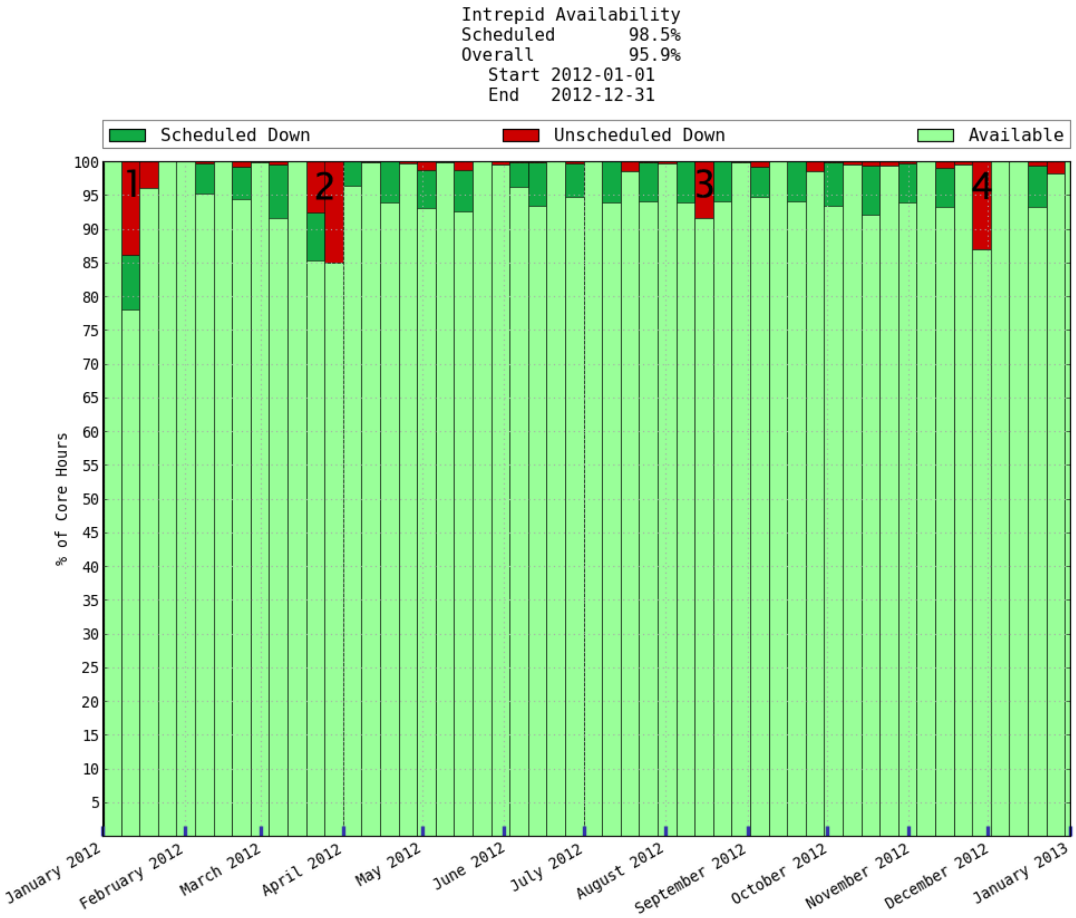


Figure 2.1: Weekly Availability for CY2012

Graph Description: Each bar represents the average of seven days of core hour usage. Each bar accounts for all the time in one of three categories. The pale green portion represents available core hours; the darker green represents scheduled downtime for that week; and red represents unscheduled downtime. As an example, the second bar from the left indicates ALCF had approximately 78% of the core hours available, approximately 8% of the core hours were scheduled down, and approximately 14% unscheduled downtime during the 7-day period beginning Jan 8, 2012. The numeric annotations are the significant losses. Each of these events is described in detail below.

Item 1: This event was actually a series of two events, with two primary causes. The impact stretched over a period of 11 days starting January 8, 2012, because of the two events and some hardware remained down while ALCF was troubleshooting the service card issue. The total impact was equivalent to the entire machine being down for 36.8 hours or 0.42% of Intrepid's annual availability. The first event (ID 2915) was a failure of the swap partition on the Blue Gene service node. This failure happened just prior to a scheduled maintenance, and then went into maintenance, requiring an extension of the maintenance period to complete the scheduled work. The remainder, and vast majority, of this time was due to a catastrophic service card failure (ID 3136,3137; Reservation IDs 5344, 5345, 5346, 5347, 5348, 5349, 5339, 5340, 5341, 5342). This was referenced in ALCF's April 2012 OAR report and was still considered open. IBM has now determined root cause. This is an excerpt from the IBM failure analysis:

It has been determined that the service card design can possibly, while in low current idle state, put one of the two 1.8V regulators locked into a non-current share mode. When this occurs, the regulator turns into a high current load that the other regulator is able to drive. This is in excess of the total current needed and designed for, in the service cards 1.8V power buss. While in this condition, the 1.8V regulators power plane between the regulators is driven into a high temperature condition that eventually break down the Printed Circuit Board (PCB) insulator. The 48V input to the regulator is then shorted to the 1.8V output power plane through the PCB heated residue. This applies high potential onto the 1.8V power buss that is also driven to the service, link, and node cards within the same midplane. This will cause the FPGA control modules on all three card types to be put into an electrical over-load condition and permanently damaging them requiring their replacement. This condition does not occur in normal operation because sufficient current is required keeping both regulators in normal, current share and supply mode.

The only physical way to prevent this failure scenario from occurring with the current service card design would be to power off the racks completely during maintenance, which is not practical. Due to the age of the BG/P product line and the release of the new BG/Q product line, IBM is not going to design a new service card. The agreement reached between IBM and ALCF is as follows: For all past and any future parts damaged as a result of this failure scenario, ALCF receives a credit with IBM for the purchase of future BG/P spare parts equal to the value of the damaged parts. This means ALCF can order whatever spare parts are needed, not just receive one-for-one replacements of the damaged parts. Since ALCF bought a parts cache at the beginning of the project and received two ORNL BG/P racks for parts, the facility already has ample supplies of some parts, such as nodes; but supplies of others that were not damaged, such as Bulk Power Modules, could run low and could be reordered at no cost through this agreement. This is a win-win situation for IBM and for ALCF. ALCF considers this issue resolved.

Item 2: This event was also a combination of events, and took place over a period of 5 days in March. The total impact was equivalent to the entire machine being down for 25.1 hours, or 0.29% of Intrepid's annual availability. The largest of these events (ID 3139; Reservation IDs 5372,5373) was a corruption in the BG/P environmental (temperatures and voltages) tables in the control system database. ALCF eventually exported the records, recreated the tables on known good disk, and then reloaded the data. No data was lost. The other events were a stuck 4K node block (Reservation ID 5374), a control system panic (ID 2975; Reservation ID 5378), and an out-of-memory event on the service node (ID 2981; Reservation ID 5384). ALCF is confident that the database corruption has been eliminated. Stuck blocks do occur on occasion and the impact is usually small. This one was included because it occurred around the same time as the other events. A true root cause for the control system panic or out-of-memory situation was never determined. However, the issues had never occurred before and have not happened since, so they have been classified as one-off anomalous events, and ALCF considers this issue closed.

Item 3: This was event ID 3045. It occurred on August 15 and was equivalent to the entire machine being down for 14.1 hours or 0.16% of Intrepid's annual availability. This was another anomalous event for which no root cause could be determined. The "service-rack" service action was executed, something that had been done hundreds of times before. On this occasion, however, the action caused the control system to crash. There were significant problems getting the control system to restart. Again, this had never happened before and has not happened since, so the issue is considered closed.

Item 4: This was event ID 3123. The impact was the equivalent of the machine being down for 22.1 hours, or 0.25% of Intrepid's annual availability. The actual loss of availability occurred November 28-29, 2012, and was due to cross-linked files in the Intrepid production GPFS file system. In consultation with IBM it was determined that an off-line fsck (file system metadata repair) was required. The availability impact was the time it took us to get ourselves and the users prepared to use the PVFS file system, which provided a temporary measure to allow users to run while we dealt with the GPFS issues. While the availability loss was significant, it was not the only impact. Due to the size of our file system, bugs in GPFS, and insufficient memory in the server running the master process in our early attempts, it took us three weeks to bring GPFS back into production. Periodic on-line fsck scans will be run in the future to check for similar problems, and we are in discussion with IBM to get direct access to GPFS level 3 support. Options for modifying the Mira file systems to have multiple file systems are also being investigated. A positive outcome of this event was that INCITE projects with large remaining balances were given priority and nearly-dedicated access to Intrepid allowing them to use significant portions of their allocation.

2.1.3 System Mean Time To Interrupt (MTTI) and 2.1.4 System Mean Time To Failure (MTTF)

MTTI = Time, on average, to any outage on the system, whether unscheduled or scheduled. Also known as MTBI (Mean Time Between Interrupt).
MTTF = Time, on average, to an unscheduled outage on the system.

ALCF MTTI and MTTF Summary:

A series of discussions occurred among the sites, and all three sites agreed to a common calculation for System MTTI and System MTTF. MTTI and MTTF are reportable values, but no specific metric has been set. Table 2.3 summarizes the current MTTI and MTTF values.

Intrepid (40K-node, 160K-core BG/P)				
	CY 2011		**CY 2012**	
	Target	**Actual**	**Target**	**Actual**
System MTTI	N/A	10.05 days	N/A	8.55 days
System MTTF	N/A	17.95 days	N/A	21.29 days

Table 2.3: MTTI and MTTF Results

ALCF MTTI and MTTF values are in line with the other centers. Overall, system stability is slightly improved in 2012 versus 2011. Unexpected failures were down, though scheduled downtimes were slightly up.

2.2 Resource Utilization

The following sections discuss system allocation and usage, total system utilization percentage, and capability usage. For clarity, usage is defined as resources consumed in units of core-hours. Utilization is the percentage of the available core-hours that were used (i.e., it is a measure of how busy the system was kept).

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

Utilization is a reportable value, but no specific metric has been set, though 80% or higher is generally considered acceptable for a leadership-class system. Table 2.4 summarizes ALCF utilization results and Figure 2.2 shows system utilization over time by program.

Intrepid (40K-node, 160K-core BG/P)				
	CY 2011		**CY 2012**	
	Target	**Actual**	**Target**	**Actual**
System Utilization	N/A	88.8%	N/A	87.6%

Table 2.4: System Utilization Results

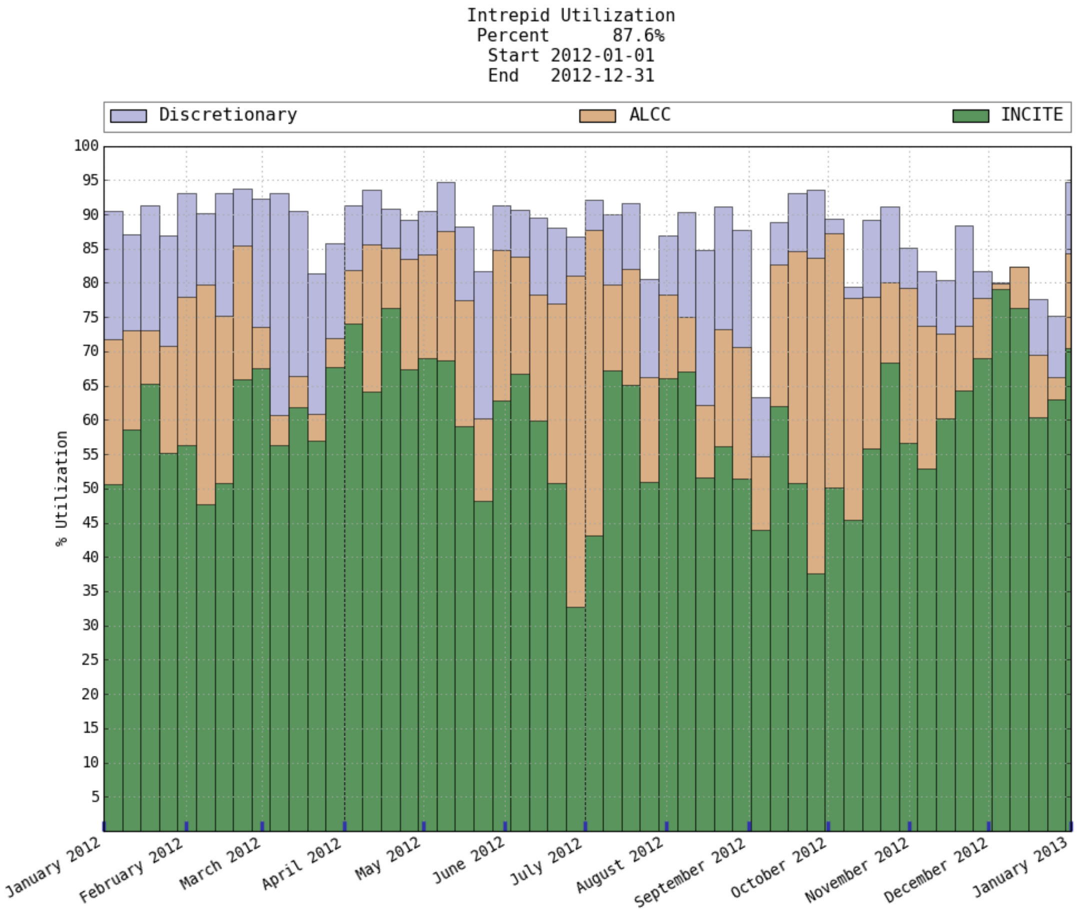


Figure 2.2: System Utilization Over Time By Program

Utilization, though down from last year, was good. Two significant drops in utilization occurred in CY2012, once in September and once in December. The drop in September was due to a bug in the scheduler that was "tickled" as a result of enabling a partition geometry that had never been used before. Specifically, a partition structure that was not strictly hierarchical. A 16K partition was created that was half in and half out of one of the 32K partitions. The drop in utilization in December was due to the time required to get projects running on PVFS during the GPFS outage described above in the availability section.

Table 2.5 shows how Intrepid's system hours were allocated and used by the allocation source. Multiplying the theoretical hours by availability and utilization values determines the hours available. Of the hours available, 60% is allocated to the INCITE program, up to 30% is available for ALCC program allocations, and 10% is available for Director's Discretionary (DD) allocations. The ALCC program runs from July through June, so to arrive at allocated values for the calendar year, half the hours are arbitrarily assigned to each year. The allocated values for the DD allocations appear higher than expected, because they represent a rolling allocation. Since a majority of the DD projects are exploratory investigations, the time allocations are not used in full. The DD allocations are discussed in detail in the Strategic Results section. In CY2012, the ALCF successfully delivered a total of 1.20 billion core-hours across all users, the same as in CY2012.

Intrepid (40K-node, 160K-core BG/P)								
	CY 2011				CY 2012			
	Allocated		Used		Allocated		Used	
	%	Core hours	%	Core hours	%	Core hours	%	Core hours
INCITE	60	732.0M	73.1	876.6M	60	732.0M	66.7	803.8M
ALCC	30	250.4M*	14.1	168.1M	30	216.0M	16.6	199.6M
DD	10	479.3M	12.8	153.4M	10	433.4M	16.7	200.6M
Total	100	1.46B	100	1.20B	100	1.38B	100	1.20B

Table 2.5: Core-hours Allocated and Used by Program

* This number was misreported in the 2011 OAR as 210.4M because of a typographical error. The number reported above is correct.

Summary: For CY2012, the system usage and utilization values are in line with general expectations. Utilization was down slightly due to two significant events. The calculations for utilization are described in Appendix A.

2.3 Capability Utilization

The facility shall describe the agreed definition of capability, the agreed metric, and the operational measures that are taken to support the metric.

ALCF Capability Utilization:
For 2012, the ALCF capability metric changed from delivering a minimum number of INCITE capability core hours to a requirement that a minimum percentage of the INCITE core hours delivered be capability. For comparison purposes, numbers have been added for CY2011 in Table 2.6.

Intrepid (40K-node, 160K-core BG/P)						
	CY 2011			CY 2012		
Capability Usage	**Total Hours**	**Capability Hours**	**Percent Capability**	**Total Hours**	**Capability Hours**	**Percent Capability**
INCITE*	876.6M	499.7M	57.0%	803.8M	507.4M	63.1%
ALCC	168.1M	95.7M	56.9%	199.6M	154.6M	77.4%
Director's Discretionary	153.4M	89.9M	58.6%	200.6M	133.8M	66.7%
TOTAL	1.20B	685.4M	57.2%	1.20B	795.8M	66.1%

Table 2.6: Capability Results

* The 2011 INCITE capability metric was 300M core-hours. The 2012 metric is 40%.

ALCF uses the Cobalt job scheduler to automatically schedule user submitted jobs without human intervention. The scheduling policy on Intrepid is designed to support capability usage, machine utilization, and quick job turnaround. Jobs are divided into queues based on their requested node count. Jobs requesting 20% or more of the machine are routed to prod-capability, where they have access to the entire machine. Smaller jobs are split again, based on their requested wall time. Small jobs with a requested wall time over six hours are restricted to running on just 40% of the machine. Small jobs with a shorter runtime are allowed to run anywhere on the machine. The effect of this policy is to ensure that when a capability job is selected to run, 60% of the machine will be available with a maximum eviction time of six hours. It is worth noting that the requested number of nodes does not always equal the allocated number of nodes. This is because the Blue Gene architecture does not allow the allocation of an arbitrary number of nodes. Hardware and policy constraints mean that only specific "quanta" of nodes can be allocated. When a user requests a job size that is not equal to one of these quanta, the next larger size is allocated. A job is not counted as capability unless the requested job size *is* capability. If the request is for something smaller but gets allocated a "capability sized" partition, that job will not count as capability. Requests for odd-sized jobs represent just a small percentage of the core hours requested over the year.

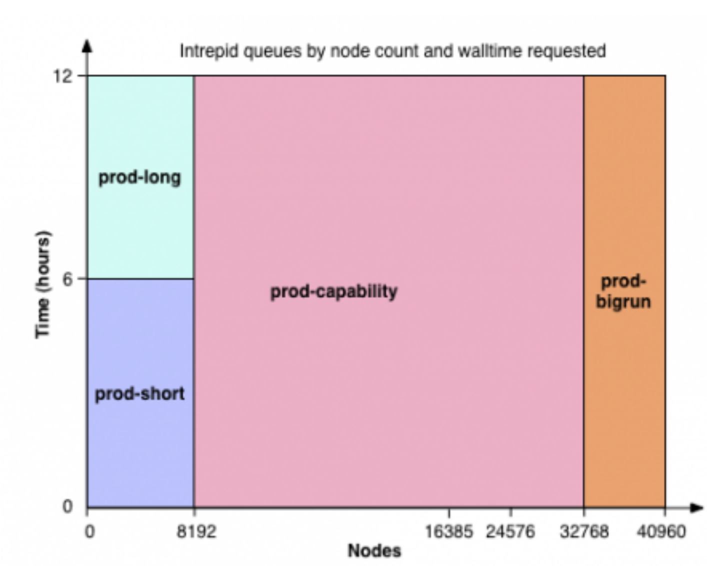


Figure 2.3: Diagram Depicting Queue Assignment Based on Job Size and Run Time.

To improve utilization, ALCF will use any available short job to fill in the sub-partitions that would otherwise be idle while they wait for the partition to drain. This is known as "backfill." Another utilization feature used is called "Big Run Monday." Every Monday, job scores are adjusted manually to maximize utilization. In general, a "best pack" is performed, which considers both job size and run time, with job size generally taking precedence. Additionally, relative job order is maintained where possible, no one project is allowed to dominate the time, and packing is limited to a maximum of 48-60 hours. (Figure 2.3) In this way, the machine is not constantly entering a draining/backfill state when one of the large jobs is run. The weak point of the scheduling policy is turnaround time on prod-long (small, long-duration) jobs. ALCF tries to mitigate this through user education about the better queue wait times for prod-short (small, short) jobs and is looking at dynamically re-sizing the prod-long percentage of machine in response to queued jobs.

Summary: ALCF exceeded the target of delivering 40% of the INCITE cycles to capability jobs. In fact, all three categories (INCITE, ALCC, and DD) exceeded 60%, as shown in Table 2.6.

The remainder of this section presents graphs showing the capability use over time, by INCITE, ALCC, and Discretionary for CY2012, and then a breakdown of how the total core-hours delivered were distributed across job sizes (Figures 2.4 and 2.5).

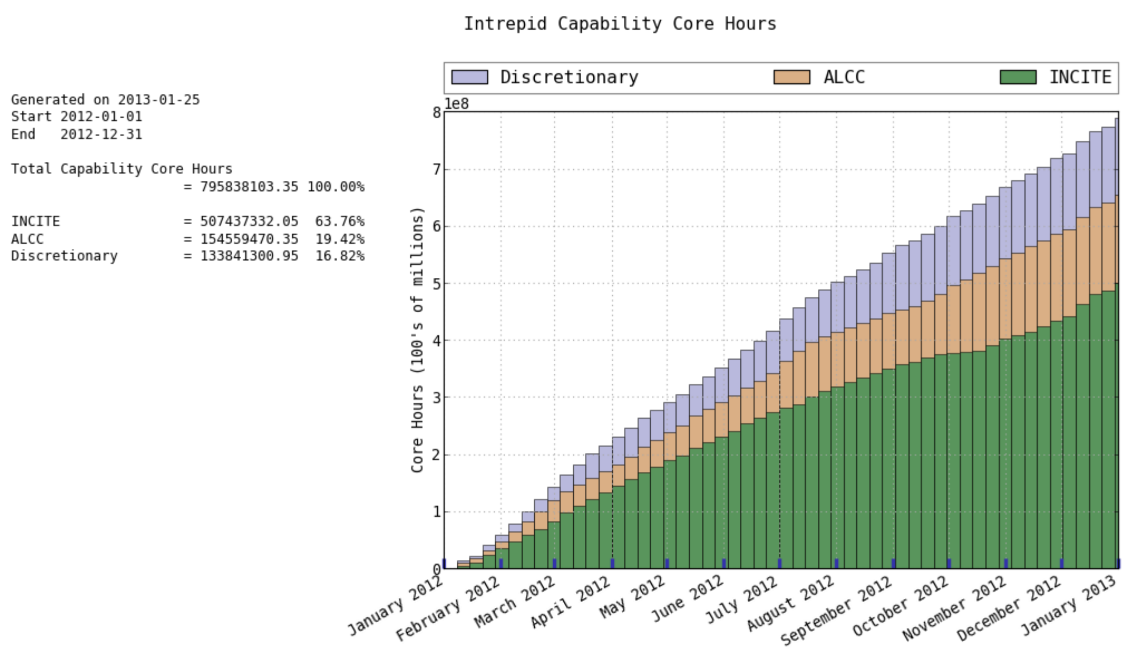


Figure 2.4: Cumulative Capability Usage During CY2012

Figure 2.5 shows job distribution by size of run. Note that larger/capability jobs are at the bottom.

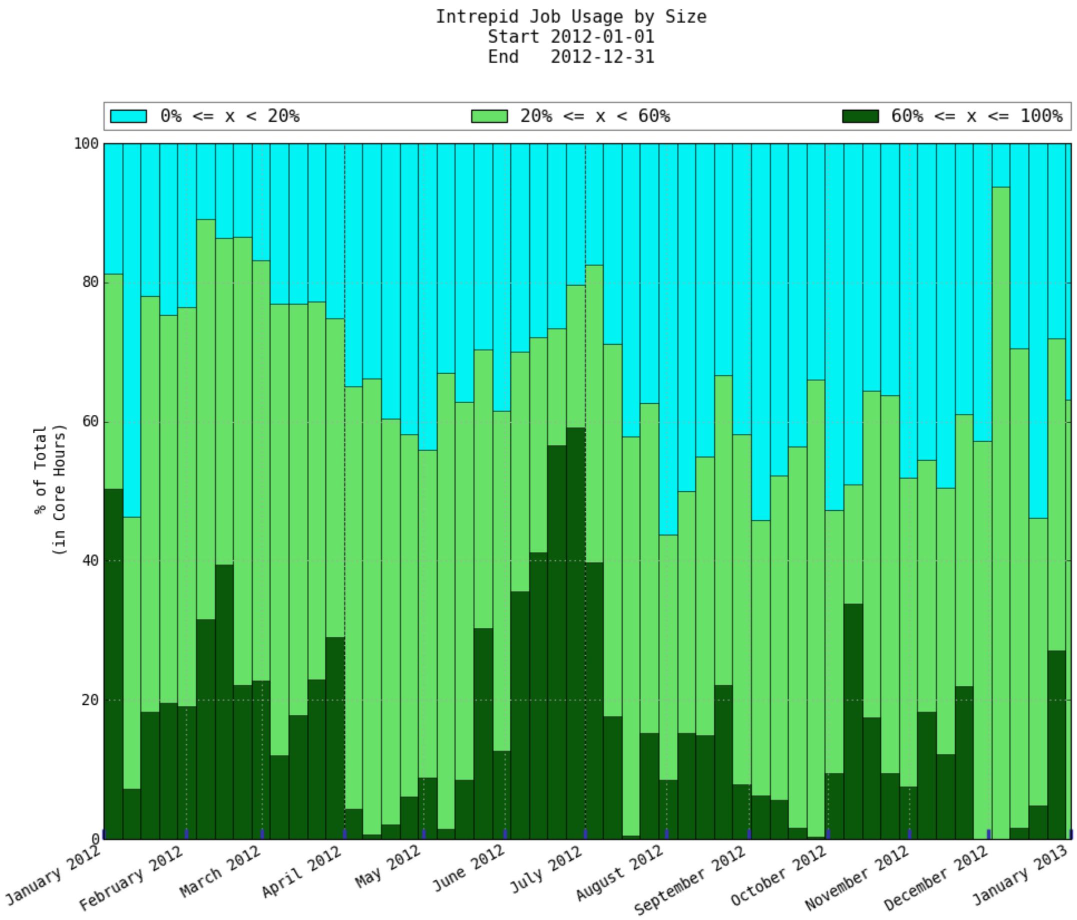


Figure 2.5: Intrepid Jobs Binned by Size

2.4 Management of INCITE Selection Process (LCFs only)

2.4.1 Process

The Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program promotes transformational advances in science and technology through large allocations of computer time, supporting resources, and data storage at the Argonne and Oak Ridge Leadership Computing Facilities (LCFs) for computationally intensive, large-scale research projects. A detailed description of the selection process is available online.[1]

The INCITE awards committee is comprised of the LCF directors, INCITE manager, LCF directors of science, and senior management. The committee identifies the top-ranked proposals by (a) peer-review panel ratings, rankings, and reports and (b) additional considerations, such as the desire to promote use of HPC resources by underrepresented communities.

[1] www.doeleadershipcomputing.org/policies/incite_overview_policies.pdf

Figure 2.6 schematically outlines the decision-making process of the INCITE awards committee. Input from the peer-review panels and computational-readiness reviewers is combined to yield an initial list of projects sorted by panel-ranked order. The INCITE awards committee identifies the top-ranked proposals by (a) peer-review panel rating and reports and (b) additional considerations, such as the desire to promote use of HPC resources by underrepresented communities. A balance is struck to ensure that each awarded project has sufficient allocation to enable all or part of the proposed scientific or technical achievements and to maximize the scientific support provided to each INCITE project.

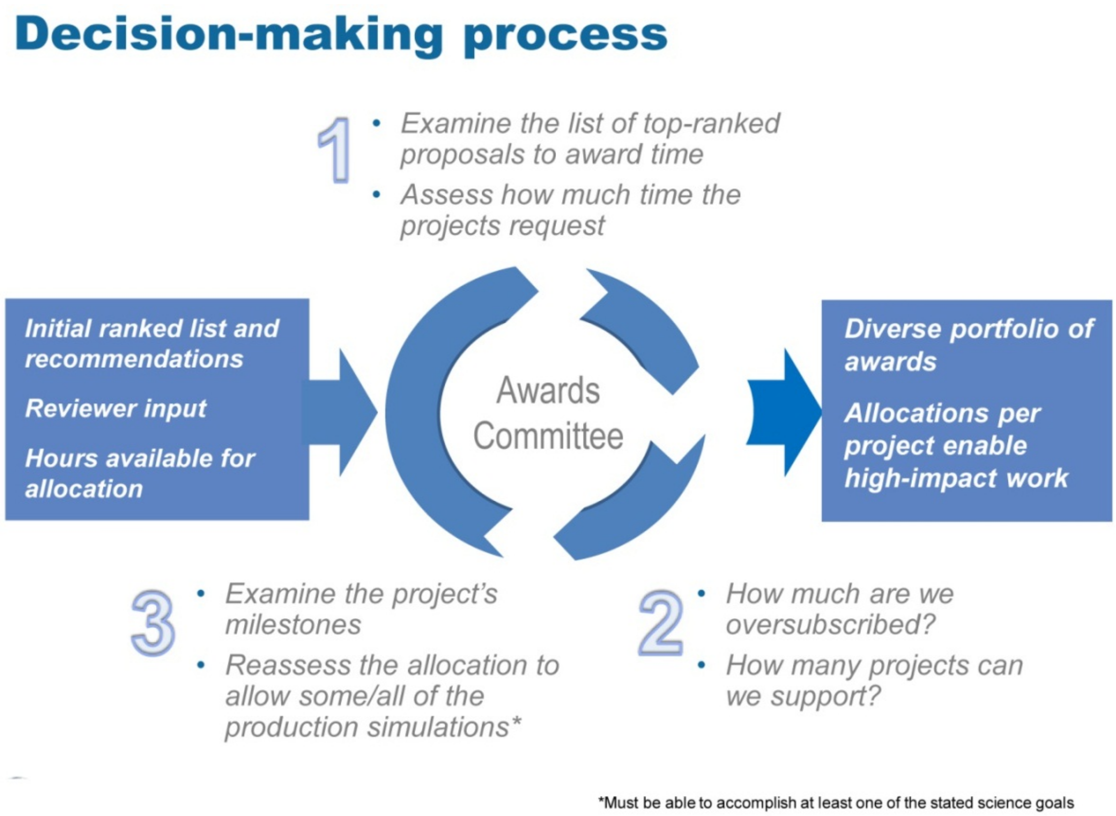


Figure 2.6: INCITE Award Decisions Workflow

The INCITE award period is January through December of the calendar year following the date of the call for proposals. (For example, "2011 INCITE" refers to the call for proposals, reviews, etc., that took place in calendar year 2010 in support of awards for the January through December 2011 time frame). The timeline of activities for the reporting period of the LCF operational assessment are summarized in Table 2.7.

2013 Call for Proposals and Awards Activities	
April 11, 2012	2013 INCITE Call for Proposals Opened
June 27, 2012	2013 INCITE Call for Proposals Closed
September 25, 26, 27, 2012	2013 INCITE Peer-Review Panel (Rockville, MD)
October 22, 2012	Announcement of Awards to Principal Investigators for CY 2013

Table 2.7: 2013 Call for Proposals and Awards Activities

2.4.2 Peer Reviewers

The INCITE manager will convene independent peer-review panels to evaluate each proposal's potential for impact. Proposals will be evaluated on scientific quality, proposed impact, appropriateness of the proposed method or approach, competence of the principal investigator (PI) and proposed research team, computational plan, and reasonableness and appropriateness of the proposed request for computational resources. Scientific review panels are composed of application domain experts from national laboratories, universities, and industry who have a working knowledge of the current computational challenges and opportunities in their fields.

83 science experts participated in the 2013 INCITE Peer-Review Panel. More than half of the reviewers are Society fellows (AAAS, APS, SIAM, IEEE, etc.), agency awardees (e.g., NSF Early Career), Laboratory fellows, National Academy members, and National Society presidents. Forty-one percent of these reviewers also participated in the 2012 INCITE review. Figure 2.7 illustrates the organizational affiliation of the reviewers.

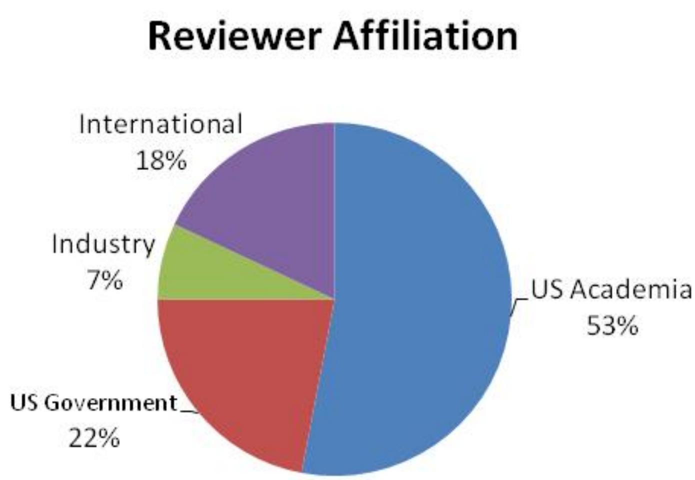


Figure 2.7: 2013 INCITE Reviewer Affiliation

The INCITE manager conducts a survey of the reviewer participants at the INCITE Peer-Review Panel. The questions and responses are summarized in Table 2.8. The scores indicate that, in the opinion of the diverse set of science, engineering, and computer science experts who participated in the one-day review of proposals, the INCITE program represents cutting-edge computational work, the proposals are of high quality, the panel itself was sufficiently diverse, and the reviewers are satisfied with the panel review process.

Survey Question	2012 INCITE	2013 INCITE
"The INCITE proposals discussed in the panel represent some of the most cutting-edge computational work in the field."	4.5	4.5
"The proposals were comprehensive and of appropriate length given the award amount requested."	4.2	4.1
"The science panel was sufficiently diverse to assess the range of research topics being considered."	4.4	4.3
"Please rate your overall satisfaction with the [2012, 2013] INCITE Science Panel review process."	4.8	4.6

Scores of "1" indicate "strongly disagree" and "5" indicate "strongly agree". (Or, for the final question "1" for "very dissatisfied" and "5" for "very satisfied".

Table 2.8: INCITE reviewer survey results (>80% response rate).

2.4.3 Proposal Allocation

Statistics on the applications submitted and awarded are reported in Table 2.9. INCITE grants one-, two-, and three-year awards: Multiyear projects are eligible to submit a renewal application. Projects that have completed the term of their award may submit a new proposal for the next allocation period.

Data Description	2012 INCITE	2013 INCITE
Total number of proposals submitted (percentage accepted)	84 (33%)	123 (33%)
Total number of renewals submitted (percentage accepted)	35 (91%)	20 (100%)
Total number of awards	60	61
Total number of hours requested	5,163M	14,679M
Total number of hours awarded	1,672M	4,678M
Oversubscription	309%	314%

Table 2.9: INCITE Submittal and Award Statistics

Conclusions

ALCF is maximizing the use of its HPC systems and other resources, consistent with its mission. For those measures where there are specific metrics (availability, INCITE hours delivered, and capability hours delivered), ALCF exceeded the required metric. For the reportable areas—Mean Time to Interrupt (MTTI), Mean Time to Failure (MTTF), and utilization—ALCF is on par with OLCF and NERSC, and the values reported are reasonable. These measures are summarized in Table 2.10 below.

Intrepid (40K-node, 160K-core BG/P)				
	CY 2011		CY 2012	
	Target	Actual	Target	Actual
Scheduled Availability	95%	97.8%	95%	98.5%
Overall Availability	90%	94.0%	90%	95.9%
System MTTI	N/A	10.05 days	N/A	8.55 days
System MTTF	N/A	17.95 days	N/A	21.29 days
INCITE Usage	732M	876.6M	732M	803.8M
Total Usage	N/A	1.20B	N/A	1.20B
INCITE Capability Usage in percentage*	N/A	57.0%	40%	63.1%
INCITE Capability Usage in core-hours*	300M	499.7M	N/A	507.4M
System Utilization	N/A	88.8%	N/A	87.6%

***ALCF INCITE capability metric changed from core-hours in CY2011 to percentage in CY2012**

Table 2.10: Summary of All Metrics Reported in the Business Results Section.

ALCF closely tracks hardware and software failures and their impact on user jobs and metrics. This data is used as a significant factor in the selection of troubleshooting efforts and improvement projects.

This page intentionally left blank.

Section 3. Strategic Results

Is the facility enabling scientific achievements consistent with the Department of Energy strategic goals?

ALCF Response

The science accomplishments of INCITE, ALCC, and DD projects clearly demonstrate the impact that ALCF is having in supporting scientific breakthroughs. ALCF staff has worked effectively with project members to adapt their project simulation codes to run efficiently in a high performance computing (HPC) environment and has enabled scientific achievements that would not otherwise have been possible.

In this section, the Facility reports:

- *Science Output;*
- *Scientific Accomplishments; and*
- *Allocation of Facility Director's Reserve Computer Time.*

ALCF Science Overview

A selection of the science output from 2012 is highlighted in section 3.1 covering areas including material science, energy grid stability, astrophysics and large-scale algorithm development. The highlights give a sense of the science accomplishments as well as the ALCF's participation and contributions to the science goals.

While supporting current projects is the key mission of the ALCF, the staff must stay engaged in the larger community for both science domains and high performance computing. In 2012 the Early Science Program (ESP) was a very tight level of engagement. The ALCF collaborated with 16 science projects to ensure their applications were ready for the new scale of science and computation on Mira. The program activities in 2012 are discussed in more detail in the next section of this overview. In addition, active engagement in the communities to pursue new projects, inform people about ALCF capabilities, stay engaged in the most recent developments and develop professionally are critical to developing the ALCF community.

In 2012, the ALCF team had a Gordon Bell finalist at SC12, wrote a chapter on marquis applications in an upcoming book by IBM on exascale computing and wrote an IBM Research Journal article about applications on the BG/Q and hosted a mini-symposium at the 15th SIAM Conference on Parallel Processing for Scientific Computing.

Individually, ALCF staff members had at least 22 publications in journals such as Proceedings of the National Academy of Science, Phys. Rev. Letters and ACM Transactions on Mathematical Software and contributed a book chapter "Visualization of Large-Scale Distributed Data". One staff member was invited to write a tutorial paper on chemical compound spaces for the International Journal of Quantum Chemistry and another was invited to do a review paper for Current Topics in Medicinal Chemistry. Staff member work was published in proceedings of twelve conferences, including three papers in Supercomputing 2012 and two in *Proceedings of the fifth International Workshop on Parallel Programming Models and System Software for High-End Computing, SIAM*. The staff presented 19 invited talks, served on the organizing committees of two workshops, and presented posters at many meetings throughout the year. The work of one member of the team was highlighted in *The New Scientist*.

The ALCF staff have some flexibility to pursue funding to pay for a limited fraction of their time or to bring in postdocs and students to work on relevant projects. The staff are participating in one X-stack project, three SciDAC 3 projects, two Basic Energy Sciences (BES) projects, and six Argonne Laboratory Director's Research and Development projects. Finally, a computational scientist on the team was awarded the Swiss National Science Foundation Assistant Professorship award (4+2 years, funding for 2 postdocs and 1 grad student) to join faculty at Basel University in 2013.

All of these strategic activities serve to attract additional, high-impact computational science projects to the ALCF, prepare users for the transition of their applications to the BG/Q, and ensure that the ALCF staff maintains its deep expertise in computational science, thus providing excellent support of its user community on current and future ALCF resources.

ALCF Early Science in 2012

Over the course of 2012, the ALCF built and accepted Mira, a 10PF BG/Q. As part of the ALCF-2 project (which is not a part of regular operations), ALCF created the Early Science Program (ESP). ALCF-2 had two main goals for the ESP:

- *Gain as much useful information as possible, as early as possible, about how scientific applications would use Mira.*
- *Have a diverse set of science projects that would help get the full system ready for production while getting real science out of the system.*

In 2010 there was an open call for projects and sixteen were selected. ALCF-2 secured funding for post-docs for all of the selected projects. These post-docs would sit at the facility and would be focused on the scientific, algorithmic and performance requirements for running their project's science. The results of these efforts will be captured in technical reports available mid-year 2013. By mid-year 2012, ALCF-2 had hired twelve of its fourteen post-docs; two teams deferred the post-doc option. ALCF computational science staff helped those projects that did not identify a postdoc.

Early Science had a very exciting year. Early in 2012, ALCF staff had access to a small 128-node very early Q system. This experience was crucial as the first racks were deployed at the ALCF. In March and then again in April, ALCF hosted two 'Jumpstart' workshops for the early science teams. About two weeks before the first workshop in March the ESP teams were given access to the first rack (1024 node, 16,384 core). Great effort was focused on making the system available for the projects as soon as possible.

While the projects had access to BG/Q hardware, the ALCF enabled all the projects on Intrepid through the Discretionary program. In 2012, the ASCR program office chose not to allocate all of the ALCC program time and gave the ALCF permission to use the extra time for ESP preparation. Several projects used the Intrepid access to develop new science capabilities in their codes needed for the new scale of science on a machine twenty times as powerful. Other projects significantly reworked their application codes. Besides development time, all of these new efforts require validation and verification. In total, in 2012, ESP projects used 83M core-hours on Intrepid. This is reported as part of the Director's Discretionary pool for 2012.

In the summer of 2012, prior to the acceptance of Mira, the ALCF opened up the resource to ESP projects to begin larger scale tests. The projects were able to share the system with acceptance preparation until the system went for acceptance, a period of several weeks. On December 18 the ALCF opened the fully accepted Mira for Early Science.

In the pre-acceptance and acceptance phases for Mira, the ESP projects were incredibly valuable. The modern application codes found multiple hardware issues in the pre-accepted hardware that were helpful in improvements for acceptance. Additionally, they were extremely valuable digging into the software stack and identifying performance bottlenecks and bugs in the pre-release system software. Examples of these (now fixed) bugs include several bugs within the MPI implementation and low-level communication layer (PAMI), compiler problems, configuration problems and failures in the I/O nodes. Significantly, one ESP project revealed a fabrication flaw in some of the BG/Q chips that only manifested after the chips had experienced many periods of high-intensity use.

Code performance has been very successful but is still at a preliminary stage. Figure 3.1 shows some of the initial speed-ups of applications from BG/P to BG/Q. These numbers are preliminary and will be changing in 2013. Figure 3.2 shows the effort required in enabling the codes to run on BG/Q. Figure 3.3 is an example of some of the technical work done on one of the ESP projects.

Application	Speedup Node to Node	Description
GAMESS	12x	Quantum chemistry
GFMC	8.5x	Green's function monte carlo. Nuclear structure.
GTC-P	15x	Particle-in-cell ITER simulations
HIRAM	>6x	Global cloud resolving climate model for GFDL
HSCD	9.2x	AMR fluid dynamics and combustion
LAMMPS/Raptor	3.1x*	Proton transport in molecular dynamics
MADNESS	3.9x*	Quantum Chemistry
MILC	>6x	Hybrid monte carlo, Krylov methods
NAMD	9x	Molecular dynamics.
Nek	8.5x	Spectral element fluid dynamics
QMCPACK	12x	Quantum Monte Carlo
SORD	12.8x	Earthquake simulation, source imaging, and full 3D tomography

Figure 3.1: PRELIMINARY Intrepid to Mira Application Speedup (Node to Node). 11/2012. The ideal speedup from BG/P to BG/Q is 15x. These measurements are done comparing runs of exactly the same number of cores on each system. The two projects with an * indicate that this speedup comparison is limited for these problem areas. The applications have significantly more capability on BG/Q.

PI/Affiliation	Code(s)	Runs on BG/Q	Magnitude of changes	Nature of Changes
Balaji/GFDL	HIRAM	✔	L	Improve OpenMP implementation, reformulate divergence-damping
Curtiss/ANL	QMCPACK	✔	M	S to port, L to use QPX in key kernels; plan: nested OpenMP
Frouzakis/Swiss Fed Inst of Technology	Nek5000	✔	S	Optimized small matrix-matrix multiply using QPX
Gordon/Iowa State	GAMESS	✔	M	64-bit addressing, thread integral kernels with OpenMP
Habib/ANL, UC	HACC	✔	M	Short-range-force only: tree code
Harrison/ORNL	MADNESS	✔	S	Threading runtime tuning Kernel tuning to use QPX
Jansen/U Colorado	PHASTA	✔	S	Unchanged MPI-only performs well; OpenMP threaded in testing
Jordan/USC	AWP-ODC, SORD	✔	S, M	None, Threading

PI/Affiliation	Code(s)	Runs on BG/Q	Magnitude of changes	Nature of Changes
Khoklov/UC	HSCD	✔	S	Tune OpenMP parameters, link optimized math libs
Lamb/UC	FLASH/RTFlame	✔	S	OpenMP threading
Mackenzie/ Fermilab	MILC, Chroma, CPS	✔	L	Full threading, QPX intrinsics/ assembler, kernel on SPI comm.
Moser/UTexas	PSDNS	✔	S	Compile dependency libs, add OpenMP directives for threading
Pieper/ANL	GFMC	✔	S	Tune no. threads & ranks.
Roux/UC	NAMD, Charm++	✔	L	Threads, PAMI implementation of Charm++
Tang/Princeton	GTC	✔	S	Improve OpenMP implementation
Voth/UC, ANL	NAMD, LAMMPS, RAPTOR	✔	M	OpenMP threads & serial optimizations in RAPTOR/ LAMMPS

Figure 3.2: Effort to port codes to BG/Q, 11/2012. There are three measures of difficulty. 'S' is zero to a few days of effort, modification 0-3% of existing lines of code. 'M' is a few weeks worth of effort, modification 3-10% of existing lines of code. 'L' is a few months worth of effort, modification beyond 10% of existing lines of code. The ranking is based on estimates made by the people who did the work.

In 2012, the work done by the ESP projects was primarily technical. The projects were testing new algorithms, performance, scalability, etc. Figure 3.3 shows an example of technical work by one of the projects significantly improving time-to-solution. These new capabilities will push bio-simulation upward in scale to eventually reach key cellular scale processes involving multi-protein complexes. An example of this is Cytochrome c Oxidase, a key enzyme in the respiratory electron transport chain of mitochondria fundamental in synthesizing ATP (the primary source of energy transfer in biological systems).

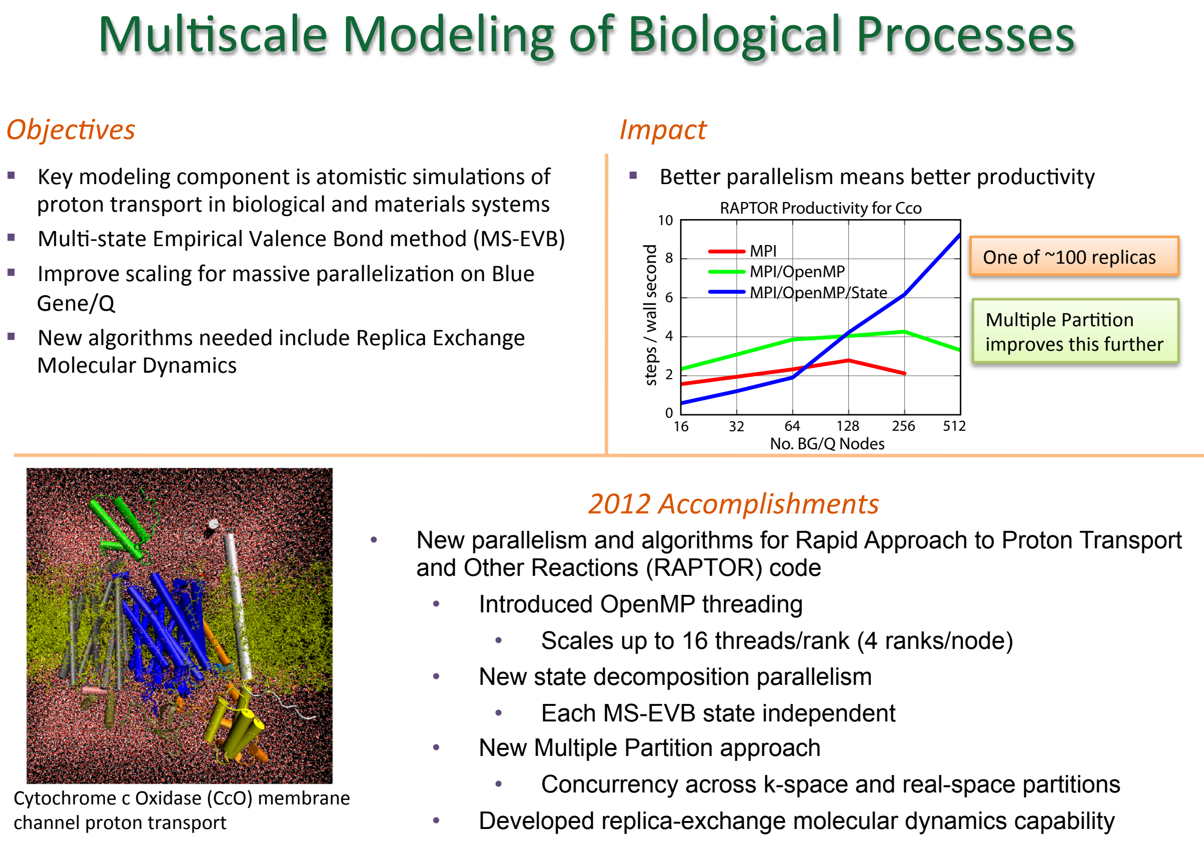


Figure 3.3: An example of some development work by one of the Early Science projects

In 2013, the ESP will be concluding. As the technical work on the codes has been drawing to a close in the early months of 2013, the ALCF will issue a technical report in the late spring. This report will cover the technical work done on the application codes to be ready to run and improve performance. The first draft of information in these reports will be given at the Mira introductory workshop the first week of March. This is a public conference that will cover the transition to BG/Q. All ESP projects will be participating and presenting results. Then, in May the ALCF will host an open ESP PI meeting. This meeting will discuss the science of the projects and any preliminary results that might be available by that point in the year. As the projects do publish science results from their work, the ALCF will eventually issue an updated ESP report that includes both the science and technical stories.

3.1 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. Tracking is done for a period of five years following the project's use of the Facility. This number may include publications in press or accepted, but not submitted or in preparation. This is a reported number, not a metric. In addition, the Facility may report other publications where appropriate.

Publications

During CY2012, at least 180 new refereed and already published papers were based directly on ALCF projects, including nine in *Physical Review Letters*, three in *Nature*, six in *Astrophysical Journal*, four in *Proceedings of the National Academy of Sciences,* and ten in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis (SC12)*. Due to the timeline for project and operational assessment reporting, this publication count does not include all of the fourth quarter of 2012.

	Nature	PNAS	Astrophysical Journal	Physical Review Letters	SC12
INCITE	3	3	5	8	3
ALCC	0	0	0	0	1
DD	0	1	1	1	6

The ALCF reports the publications derived from research done at the facility. The ALCF maintains a database of publications, and the facility website presents a bibliography of publications with abstracts and online access to some papers. This list of papers is based solely on what projects report back to the ALCF in the quarterly or renewal reports. The papers are exclusively refereed articles or proceedings. The ALCF is doing publication searches for unreported publications from previous years, but the efforts have not yet yielded new publications.

3.2 Scientific Accomplishments

The Facility highlights a modest number (at least five) 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. The accomplishment slides should include the allocation, amount used, and a small bar graph indicating size of jobs.

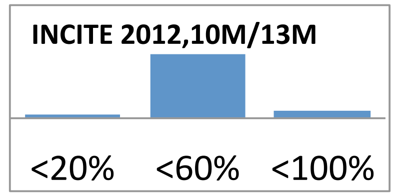
LCFs should include tables/charts comparing time allocated to time used by projects.

ALCF Science Highlights

This section highlights some of the scientific accomplishments achieved using the ALCF resources in the past year.

Fighting Drug Resistant Bacteria

INCITE: Andrew Binkowski, Argonne National Laboratory



The NDM-1 protein became known a few years ago when an infection was found in a person in England. They traced this back to India but the protein was detected globally rather quickly. In 2011, a global study found NDM-1 in about 75% of the drinking water in developed countries. In the United States, it is know that E. Coli, K. pneumoniae, and E. cloacae all have NDM-1 and, globally, drug resistant tuberculosis (among other bacteria).

Researchers from the Argonne National Laboratory Midwest Center for Structural Genomics (MCSG) and Structural Biology Center (SBC) began collaborating with researchers at Texas A&M who had been studying the enzyme, but had been unable to determine the 3-D structure of the protein. In record time, within just a few months of acquiring the protein, researchers at MCSG solved the structure of the enzyme using the Advanced Photon Source at Argonne. They released the first publically available structure for the protein in the Protein Data Bank (PDB). (This structure was with ampicillin but the group is now working with NDM-1 bound to many compounds.)

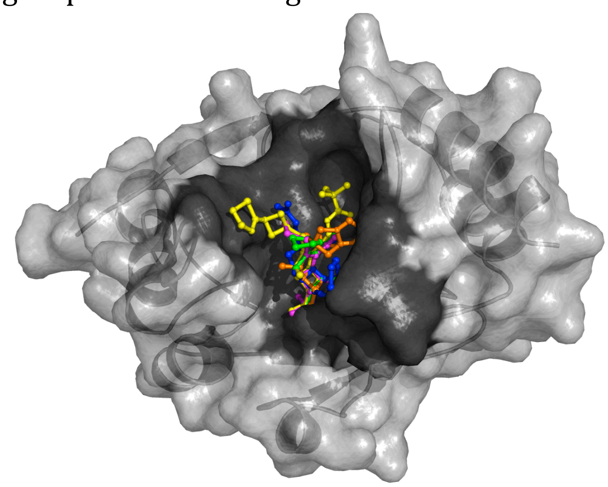


Figure 3.4: NDM-1 enzyme's structure revealed a large cavity (dark gray) capable of binding a variety of known antibiotics (shown in different colors). Once bound the enzyme can cut the antibiotic, destroying the capabilities of the antibiotic.

Once the MCSG team had determined the protein structure, Binkowski, a member of MCSG, used his INCITE allocation at the Argonne Leadership Computing Facility to determine why NDM-1 made bacteria so drug resistant. The computations are a detailed workflow, written specifically for the project, they analyze the structure and simulate how different antibiotic compounds bind to NDMs. They used 10M Intrepid core-hours over the course of one month to analyze nine different compounds. This represents a substantial speed-up over traditional experiments that can take many months to obtain the same answers.

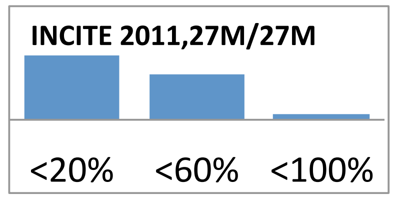
NDM-1 is a member of the beta-lactamase enzyme family characterized by their ability to hydrolyze or "cut" the lactam ring in common antibiotics. The antibiotics would normally kill bacteria by disrupting cell wall synthesis, but once cut they are unable to carry out this function. What makes NDM-1 so dangerous is that its cavity is particularly large and flexible so most antibiotics (even the strongest antibiotics of last resort) can enter the cavity and be rendered ineffective.

IMPACT: Determining the structure and behavior of NDM-1 illustrated a flaw in the design of new antibiotics. The design of an entirely new drug and moving it through the approval process and out to market can take 20 years. To modify an existing drug can reduce that time significantly. By continually producing larger and larger antibiotics, researchers have tried to keep pace with anti-biotic resistant enzymes, but proteins like NDM-1 are adapting faster than the pipeline for drugs can make new drugs available. The computational pipeline at the ALCF is helping to better understand the mechanism of NDM-1, hoping that this will provide insights in to developing a novel strategy to combat this public health threat.

ALCF Contribution: ALCF integrated NAMD with an IBM PAMI version of Charm++. This was the root of a 40% speedup.

Lithium Air Battery

INCITE: Jack Wells, Oak Ridge National Laboratory



Weight-for-weight, a rechargeable lithium-air battery can store five to ten times the energy of a conventional lithium-ion battery. It has the highest energy density of any battery yet devised. Currently, the batteries in electric cars can't compare with gasoline in the amount of energy derived from a given weight of fuel. Cars with lighter, more powerful lithium-air batteries could overcome this challenge. The battery's enhanced energy storage capability could enable the widespread use of electric vehicles.

Realistic computational chemistry models are of crucial importance to developing a viable lithium-air battery. Led by Jack Wells, director of science at the OLCF, a research team modeled how the lithium-air cell would work. Using supercomputing resources allocated by a Department of Energy INCITE award, the researchers studied the cell's basic chemistry, how catalysts affect its performance, and how to make the device more stable. They investigated the use of propylene carbonate (PC) as a potential electrolyte for a rechargeable lithium-air battery to establish the practicality of PC. By employing sophisticated models for both the surface and the electrolyte, they achieved sufficient detail to study the full complexity of propylene carbonate degradation by lithium-air discharge products. The team evaluated this problem using three application codes -- CP2K, CPMD, and VASP -- on ALCF's Intrepid.

For the first time, the scientists demonstrated that propylene-carbonate is not a suitable electrolyte for lithium-air-based batteries. Formerly, PC was believed to be stable, based on the behavior of similar systems. However, the simulations in this project revealed that during discharge, oxygen reduces to peroxide, forming layers of Li_2O_2 that immediately degrade PC. Moreover, special properties of the surface of Li_2O_2 act to enhance its reactivity. The findings also explained the recent failure to identify Li_2O_2 in the discharge products on the cathode surface during experimental research into this technology and determined the major chemical mechanisms behind the charge/discharge processes.

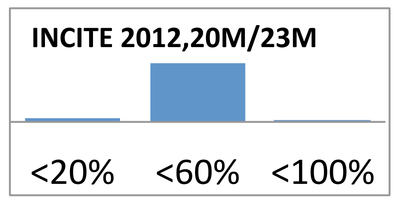
IMPACT: Demonstrates that a material, previously assumed a good candidate for Li/Air batteries, should not be used and revealed characteristics unseen in experiments.

ALCF Contribution: Helped enable ensemble simulations with IBM and ANL researchers.

Publication (cover article): Laino, T. and Curioni, A. (2012), "New Piece in the Puzzle of Lithium/Air Batteries: Computational Study on the Chemical Stability of Propylene Carbonate in the Presence of Lithium Peroxide," *Chem. Eur. J.*, 18: 3421. doi: 10.1002/chem.201290042

Large-Eddy Simulations of Contrail-to-Cirrus Transition

INCITE: Roberto Paoli, CERFACS



Contrails are ice clouds that form by condensation of water vapor exhaust from aircraft engines and develop further in the aircraft wake as they are entrained by the airplane trailing vortices. When contrails spread to form cirrus clouds, they can persist for hours and extend over areas of several square kilometers. These "contrail cirrus," which artificially increase Earth's cloudiness and become almost indistinguishable from natural cirrus, are among the most uncertain contributors to the Earth's radiative forcing, and represent a source of increasing concern for scientists and policymakers as the demand for air travel continues to grow.

Roberto Paoli from CERFACS is leading a team of scientists using resources at the ALCF to study the contrail-to-cirrus transition. The main objective of the project is to identify the key processes that determine the dynamical and microphysical characteristics of the contrail as a function of its age. The simulations were performed on up to 32,768 cores of Intrepid and on a computational domain of 4 Km and up to 2 m resolution in all directions -- the largest simulation to date in this area. It is found that ambient turbulence (shown in Figure 3.5a by means of temperature fluctuations) controls the initial contrail evolution whereas radiative transfer is the main driver later on. The evolution of the 3-D structure of the contrail is illustrated in Fig. 3.5b through an iso-surface of ice mixing ratio: the contrail grows, spreads and shears under the combined action of turbulent diffusion, condensation and local updrafts induced by radiative transfer with the atmosphere.

IMPACT: Better understanding the physics of the formation of contrail cirrus will lead to more complete representation of aviation impact into next-generation climate models.

ALCF Contribution: ALCF staff helped on the activation of a Globus Online endpoint with Globus Connect Multi-User, which speeds up by a factor of six their data transfers from/to Intrepid. Additionally, ALCF tuning improved I/O performance at capability levels by a factor of twenty.

Publication: R. Paoli, O. Thouron, J. Picot and D. Cariolle "Large-eddy simulations of contrail-to-cirrus transition in atmospheric turbulence", *65th Annual Meeting of the APS, Division of Fluid Dynamics*, Vol. 57, No. 17, San Diego, CA, November 18-20, 2012

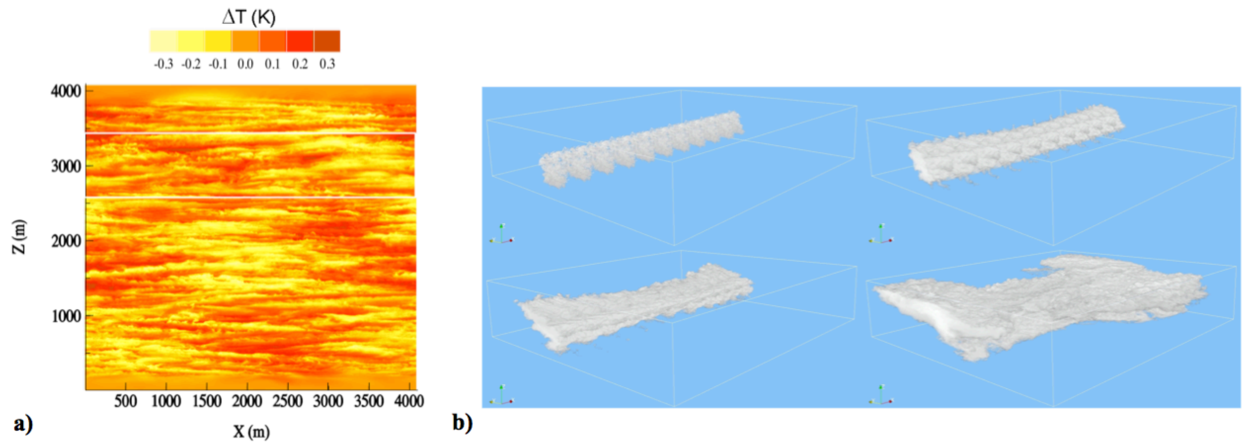
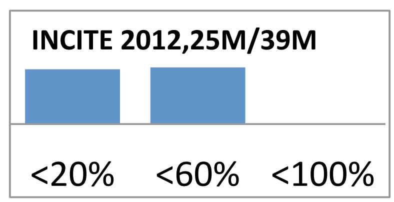


Figure 3.5: a) Potential temperature fluctuations of the atmosphere in a vertical plane in the middle of the computational domain (t=5 min. old contrail). The white boxes indicate the portion of the domain where the contrail is initially located. b) 3-D iso-surface of ice mixing ratio showing the contrail transformation into a young cirrus cloud. From top to bottom, left to right: t=5 min; t=10 min; t=20 min and t=40 min old contrail.

New Approach for Calculating Surface Characteristics
INCITE: Giulia Galli, University of California, Davis



Within this INCITE project first principles molecular dynamics simulations were developed and applied using semi-local, van der Waals corrected, and hybrid electron density functional theory (DFT) to compute vibrational properties of liquid mixtures and liquid/solid interfaces.

The scientific goal is direct comparison with experimental data becoming available from experiments at the Spallation Neutron Source at OLCF. Systems of interest are water/methane/carbon dioxide mixtures and water in contact with oxide and graphitic-like surfaces. The research is enabling the team to separate the effects responsible for observed phenomena that can't be easily probed experimentally, such as interface structure dependence on vibrational properties of liquids and liquid mixtures.

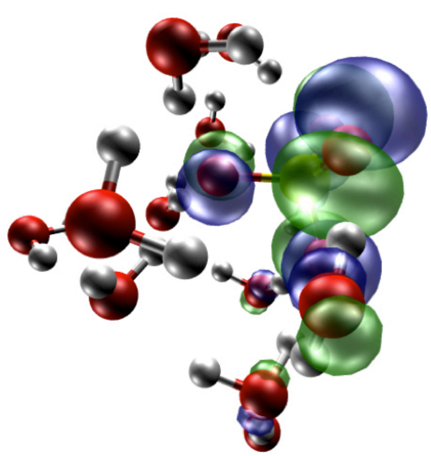


Figure 3.6: Highest occupied molecular orbital of two configurations of a solvated [SO_4]—cluster in water, as obtained using *ab initio* calculations with the hybrid Perdew-Burke_Ernzerhof functional PBE0.

The team completed a series of simulations of the structure and infrared (IR) spectra of liquid water and ice and various surfaces using *ab initio* molecular dynamics and local, generalized gradient, non-local van der Waals corrected, or hybrid DFT. These simulations represent the first calculations of infrared spectra of liquid water carried out by various different DFTs.

One of the main findings is that the use of the hybrid functional has resulted in calculated vibrational spectra that offer dramatically improved agreement with experiment, when compared to the lower level DFT approximations, such as the local density approximation, or semi-local generalized gradient approximated functionals. Furthermore, the hybrid DFT estimated band gap of the liquid also improved when compared to experiment. The van der Waals corrected DFTs also yielded superior results. Therefore, future simulation results using this hybrid DFT approach will be significantly better validated.

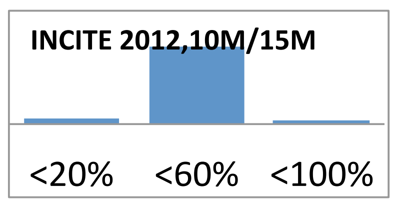
IMPACT: Validation of these new hybrid DFTs is a significant step in gaining predictive first principles capabilities for energy-relevant systems for which experiments to do exist.

ALCF Contribution: Enabled improved throughput on Intrepid by writing scripts for the job management.

Publication: This work has been presented in six publications, three in 2012, and was presented at four conferences in 2012. One publication: "Dispersion interactions and vibrational effects in ice as a function of pressure: A first principles study," E. D. Murray and G. Galli, *Phys. Rev. Lett.,* 2012, 108, pp. 105502.

Petascale Simulations of Inhomogeneous Alfvén Turbulence in the Solar Wind

INCITE: Jean C. Perez, University of New Hampshire



The origin of the solar wind and the heating of the solar corona are two of the most compelling problems in heliospheric physics. A number of observations suggest that Alfvén waves (AWs) and AW turbulence play a key role in the solution to both problems. These mechanisms may explain how plasma in the corona is heated and accelerated away from the sun to produce the solar wind. However, the scientific community remains divided about the viability of the AW turbulence scenario, in large part because they do not have a detailed understanding of the properties of such turbulence in the highly inhomogeneous near-sun region. For example, between the coronal base and a heliocentric distance r of ten solar radii (R⊙), the solar wind outflow velocity U increases dramatically and the plasma density ρ drops by about 5 orders of magnitude. This inhomogeneity leads to the partial reflection of the low-frequency AWs launched by the sun and fundamentally alters the nature of AW turbulence relative to the better-understood homogeneous case.

Jean C. Perez from the Space Science Center at University of New Hampshire is leading a numerical study of Alfvén waves in an inhomogeneous background and the resulting turbulence, which is prevalent in the extended solar atmosphere. The physics governing this turbulence is well represented by the reduced magnetohydrodynamics (MHD) model, modified to account for an inhomogeneous background. These equations describe the plasma as an electrically conducting fluid. Perez's IRMHD code implements a pseudo-spectral numerical solution to the inhomogeneous reduced MHD equations. Perez et al. have completed and analyzed numerical results for a number of resolutions, ranging from 4,096 to 8,192 grid points in the radially outward direction, parallel to the background magnetic field, and 1,282 and 2,562 grid points in the perpendicular directions. The overall geometry is a thin magnetic flux tube, extending from the solar surface to up to $11R_{\odot}$. These results show evidence of a turbulent inertial range—a power-law spectrum typical of turbulent flows. Figure 3.7 shows the overall nature of the flow. Recent and future simulations at higher resolutions (512^2 and 1024^2) will verify if that inertial range is properly resolved.

IMPACT: In addition to being basic unsolved problems in plasma physics, coronal heating and the solar wind drive space weather. This work describes changes in the solar plasma and magnetic field in the vicinity of Earth, which can impact its inhabitants in many ways, including disruption of GPS and other radio signals.

ALCF Contribution: At the 2012 ALCF Winter Workshop, ALCF staff sat with Perez and with Sameer Shende (U. Oregon), an expert on the TAU performance profiling and tracing software. The ALCF analyzed performance of the IRMHD code and identified a way to improve the inter-process communication in a key part of the algorithm for a 2x improvement. This, coupled with adoption of FFT routines more highly optimized for Blue Gene architecture (a 3.5x improvement on smaller scale runs), dramatically reduced runtime of the IRMHD code, allowing a much more complete physics study within the allocated INCITE hours.

Presentation: : J. C. Perez, *Turbulent Cascade in the Solar Wind: Anisotropy and Dissipation*, invited presentation at Centre International d'Ateliers Scientifiques, Observatoire de Paris, September 17-21, 2012. **Presented** at the American Geophysical Union Fall Meeting, December 3-7, 2012.

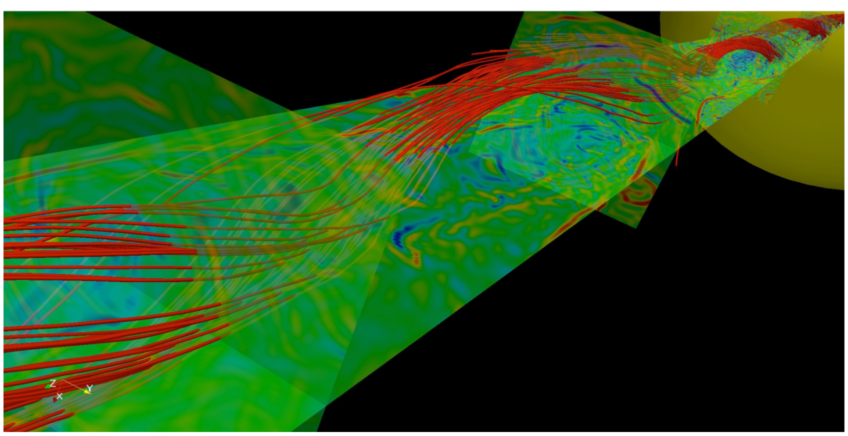
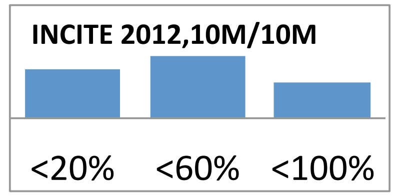


Figure 3.7: Turbulent magnetic field lines (red) inside the simulated magnetic flux tube, representing a coronal hole that expands from a small patch on the solar surface to 5 solar radii. Alfvén waves, launched by convective motions on the photosphere, propagate in the inhomogeneous solar atmosphere producing primary reflected waves that interact non-linearly with the outward waves, driving a turbulent cascade. This cascade continues with secondary reflections in a very complex interplay between the wave reflections and nonlinear interactions.

Stochastic Optimization of Complex Energy Systems

INCITE: Mihai Anitescu, Argonne National Laboratory



The national imperative to use to handle increased levels of renewable sources of energy, such as wind and solar, has resulted in a significant challenge for the operation and design of the U.S. power grid. As opposed to fossil-fuel generation systems, the available amount of renewable energy at any given time is uncertain and unmatched with demand changes, which leads to the risk of not satisfying demand and thus having to "shed load," that is, to cut consumers from the grid. This uncertainty in weather and other risks such as generator and transmission line failures are currently mitigated by using conservative and expensive reserve units, which are coal or natural gas generators that can increase electricity generation on short notice. Such reserves are both economically and environmentally costly.

Mihai Anitescu from Argonne National Laboratory is leading a team of scientists that explores the benefits of using optimization under uncertainty as the paradigm for managing the uncertainty in renewable energy supply in order to reduce reserve requirements and stabilize electricity markets in the next-generation power grid. As a part of this project, stochastic programming – a leading paradigm for optimization under uncertainty – formulations of the decision process that schedules supply and matches demand are solved by a new scalable solver, PIPS. The magnitude of such formulations for the state of Illinois, which contains approximately 2,000 transmission nodes, 2,500 transmission lines, 900 demand nodes, and 300 generation nodes and needs to be considered over 24 successive hourly time periods, can reach billions of variables and constraints once the uncertainty in the supply is taken into account.

Incorporation of new algorithmic research results into PIPS made it possible to efficiently simulate such large systems on Intrepid in an amount of time considered "real-time" by current power industry practices. Strong scaling was obtained by PIPS in solving a 12-hour horizon stochastic problem with 2 billion constraints and 2 billion inequalities. In turn, this has allowed the team to demonstrate that, at least on some configurations, even 20% wind penetration – the U.S. stated policy for the year 2030 – can be accommodated without the need for significant reserve increase if using stochastic optimization, something which would not be achievable with the traditional formulations.

IMPACT: Novel optimization-under-uncertainty formulations of management of renewable energy sources may enable as much as twenty percent wind penetration without significant increase in reserves or decrease in reliability.

ALCF Contribution: ALCF helped adopt and port required libraries and debug low MPI challenges.

Publications: M. Lubin, J. A. J. Hall, C. G. Petra, and M. Anitescu. Parallel distributed-memory simplex for large-scale stochastic LP problems, to be published in COAP.

M. Lubin, K. Martin, C.G. Petra, and Burhaneddin Sandikic. *On parallelizing dual decomposition in stochastic integer programming,* to be published in OR Letters.

C. Petra, O. Shenck, Klaus Gaertner, M.Lubin, "*Scalable stochastic optimization of complex energy systems on high-performance computers",* in review by SIAM Journal on Scientific Computing

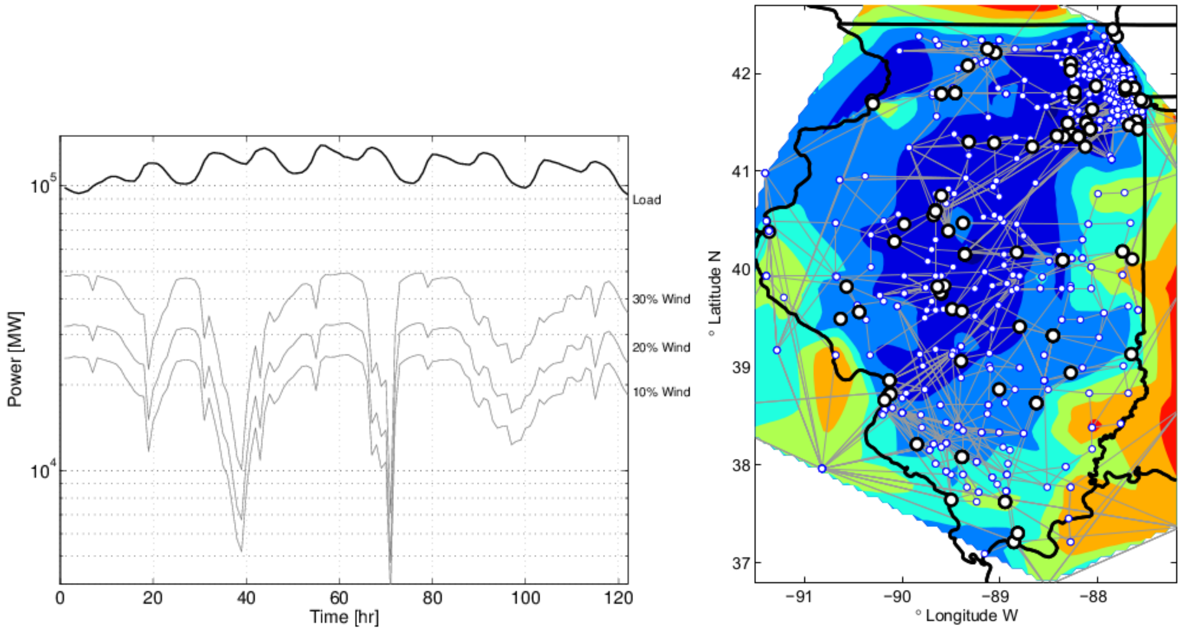
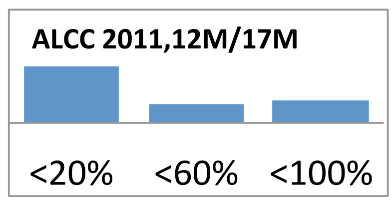


Figure 3.8: (Left) Snapshot of total load and wind supply variability at different adoption levels. (Right) State of Illinois power grid network and prices field.

Toward Crystal Engineering from First Principles

ALCC: James Chelikowsky, University of Texas at Austin



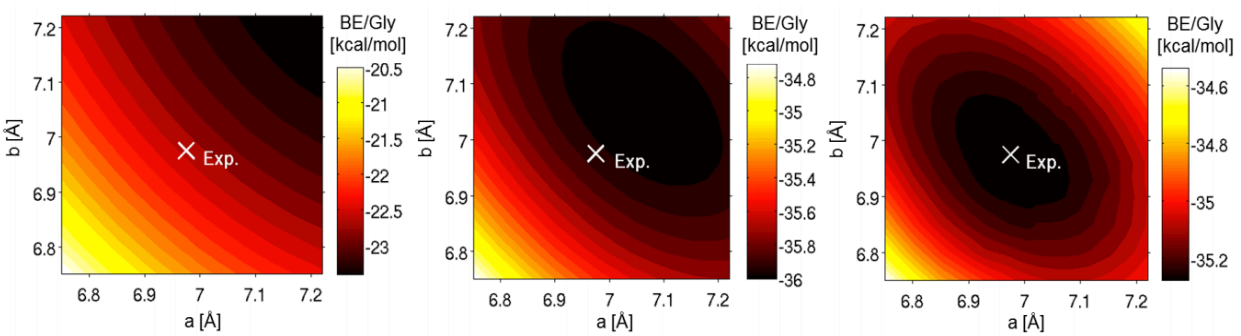
A deeper understanding of the dispersive interactions that govern the structure and properties of molecular crystals will enable computational crystal engineering from first principles. For this purpose, University of Texas at Austin researchers Noa Marom and James Chelikowsky, in collaboration with Alexandre Tkatchenko from the Fritz Haber Institute of the Max Planck Society in Berlin, are employing density functional theory (DFT) in conjunction with newly developed methods for describing dispersion interactions. Their research focuses primarily on demonstrating the capability of their approach for a series of case studies reflecting the wide variety of applications of crystal engineering from biological systems, such as amino acids and antimalarial drugs, to technological applications, such as organic semiconductors for organic electronics and dye-sensitized TiO_2 clusters for solar cells. The systems being studied comprise several hundred atoms, pushing the size limits of fully quantum mechanical calculations. Potential energy surface (PES) exploration for such systems requires massively parallel computing.

To date, the accurate and efficient many-body dispersion (MBD) method has been developed and applied to the prediction of the structure, energetics, and electronic properties of molecular crystals. The MBD method is a twofold improvement over the Tkatchenko-Scheffler (TS) dispersion method. First, depolarization effects arising from the long-range electrostatic screening in the crystalline medium are accounted for via the self-consistent screening equation of classical electrodynamics. Second, the MBD energy is calculated to infinite order by solving the Schrodinger equation for a system of coupled fluctuating dipoles. The MBD contribution is crucial for locating the PES minima for molecular crystals.

IMPACT: The team found good agreement with one of their methods, providing a path forward for their new approach enabling computational crystal engineering.

ALCF Contribution: The Argonne Leadership Computing Facility team helped to write scripts for exploring the PES of molecular crystals using the MBD method, and in adapting evolutionary optimization codes to the IBM Blue Gene/P platform.

Abstract: "Many-body dispersion interactions in molecular crystal polymorphism," N. Marom, R. A. DiStasio Jr., V. Atalla, S. Levchenko, J. R. Chelikowsky, L. Leiserowitz, A. Tkatchenko. http://arxiv.org/abs/1210.5636

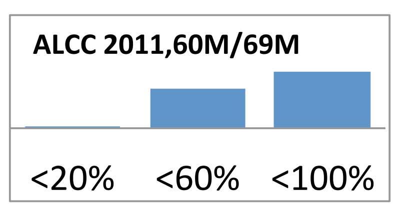


(a) PBE (b) PBE+TS (c) PBE+MBD

Figure 3.9: Potential energy surfaces for the a-b plane of Gamma-Gly using (a) their specific density functional theory (DFT) (using the Perdew-Burke-Ernzerhof (PBE) functional), (b) DFT plus an atom pairwise van der Waals correction (TS), and (c) DFT plus a many-body van der Waals correction (MBD). Only the last correctly explains the experimental lattice parameters (marked by a cross).

Understanding of "Crackle" in Supersonic Jet Noise

ALCC: Parviz Moin, Stanford University



Crackle is an especially irritating component of supersonic jet noise because it is associated with shock-like "N-shaped" acoustic waveforms consisting of sudden strong compressions followed by more gradual expansions. In addition to the irritation factor, crackle may account for up to 30% of the overall sound pressure level along the direction of peak jet noise, when it occurs. Because crackle occurs in the direction of peak jet noise, its elimination has the potential to help meet the U.S. Navy's near-term jet noise reduction goal of 3 dB in the peak noise. The mechanism by which N-shaped waves are generated is not yet fully understood. In particular, there has been some debate about whether these waves are generated directly by the supersonic jet itself, or whether they may be a result of nonlinear acoustic propagation. One of the objectives of the ALCC 2012 allocation was to apply large eddy simulation (LES) to determine the source of crackle in hot supersonic jets issuing from military-style faceted nozzles.

From the simulations performed on Intrepid, a new understanding of the source of crackle noise in heated, supersonic jets is emerging. The team has identified that a region of high pressure along the upper jet shear layer appears at the point where crackle waves are emitted from the jet. This region of high pressure appears to be well organized, extending all the way across the jet shear layer, even connecting to the shock cell in the interior of the jet. Using a phase-averaging technique, the team has been able to identify patterns in how "crackle" might form. These events are rare and so require a significant amount of statistics. Several such crackle events for the base jet noise case were detected and stored for future analysis. Since jet noise crackle is very intermittent, a novel detection scheme and backtracking capability was implemented so that, once a strong pressure spike in the near-field jet is detected, the code backtracks to an earlier restart file and then saves the relevant data for that crackle event.

IMPACT: Identifying the source of the crackle noise is a new discovery. From this new understanding, new nozzle designs can be simulated to minimize the source of the noise.

ALCF Contribution: The ALCF ported the code and worked on algorithmic improvements.

Publication: Nichols, J.W., Lele, S.K., Moin, P., Ham, F.E., Bridges, J.E. (2012). "Large-eddy simulation for supersonic rectangular jet noise prediction: effects of chevrons," 18th AIAA/CEAS Aeroacoustics Conference, chapter DOI 10.2514/6.2012-2212

Report: Nichols, J.W., Lele, S.K. (2012) "Crackle in supersonic heated jets," Center for Turbulence Research Annual Research Briefs, Stanford University.

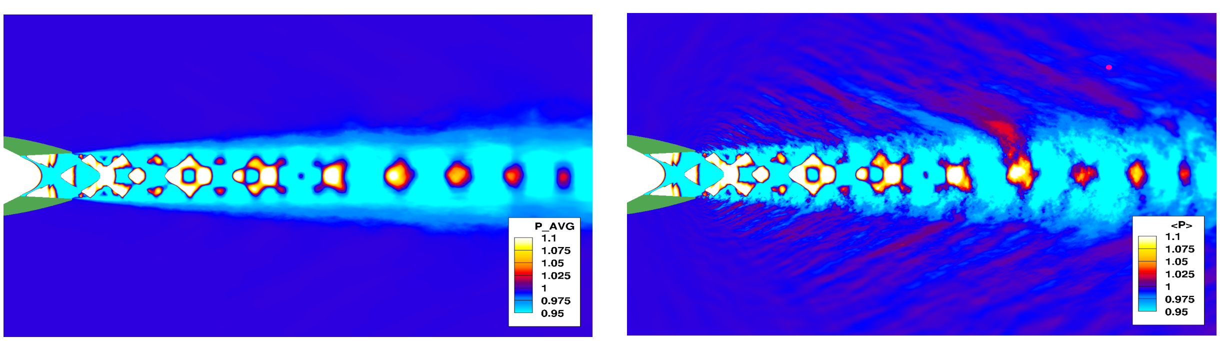


Figure 3.10: (left) shows contours of the LES mean pressure on an axial cross section. The nozzle material is shown in green. Both the sharp throat and the nozzle lip produce shocks that create a double-diamond shock cell pattern downstream. Far downstream, the double diamond cells merge into single regions where the pressure is statistically high. These single-celled shock cells then gradually reduce in intensity downstream. (Right) shows contours of pressure averaged over all of the source fields associated with 22 different crackle events.

Probe the Cosmic Structure of the Dark Universe – Gordon Bell Finalist

Early Science: Salman Habib, Argonne National Laboratory

> **ESP on Mira, 150M**
> ***Usage cannot be reported for this work. It was done on very early systems at ANL, LLNL, and IBM.***

Scientists have long been fascinated by the structure of the universe, hoping ultimately to be able to reveal its fundamental origins. As part of this quest, researchers are now planning the most detailed, largest-scale simulation of structure formation in the universe to date. The simulation targets an improved understanding of the nature of the "dark universe."

Physicists Salman Habib and Katrin Heitmann of Argonne National Laboratory's High Energy Physics (HEP) and Mathematics and Computer Science (MCS) divisions lead the computational cosmology project. Their code is called HACC, an acronym for hardware/hybrid Accelerated Cosmology Codes. HACC is designed for extreme performance, but with great flexibility in mind, combining a variety of algorithms and programming models that make it easily adaptable to different platforms.

To get good performance on the BG/Q (as with most modern machines) it is necessary to create as much data locality as possible, decompose the problem in a balanced way (both between nodes and between threads on each node), and then map the problem of interest onto what the machine does well; for the BG/Q, this is fused multiply-add operations. This means that, for the most expensive part of the simulation, the short-range force calculation, ESP post-doc Hal Finkel created a preprocessing step (the RCB tree) that maximized data locality and created balanced partitions of the overall problem. Then the team used the force kernel, based on a polynomial fit to the short-range force formula and highly optimized by ALCF staff member Vitali Morozov. Polynomial evaluations use mostly fused multiply-add instructions, and so this maps the problem onto what the BG/Q does well.

The team attained nearly 14 PF of sustained performance at 69% of peak and 90% parallel efficiency on more than 1.5 million cores, setting a benchmark for the largest cosmological simulation to date. The team's paper was a Gordon Bell Finalist at SC12.

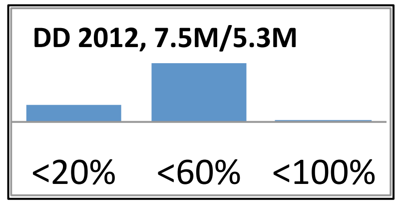
IMPACT: The speedup achieved for the Gordon Bell competition is enabling time-to-solution improvements for all science runs. This technical achievement enabled researchers to perform the largest-scale simulation to date and will, hopefully, improve the understanding of the nature of the "dark universe".

ALCF Contribution: ALCF ESP postdoc Hal Finkel and staff member Vitali Morozov rewrote algorithms and used input from the larger team to optimize kernels for amazing performance. The ALCF integrated GLEAN into multiple applications, including HACC achieving throughput of 170GB/s.

Publication: S. Habib, V. Morozov, H. Finkel, A. Pope, K. Heitmann, K. Kumaran, T. Peterka, J. Insley, D. Daniel, P. Fasel, N. Frontiere, and Z. Lukić. 2012, "The universe at extreme scale: multi-petaflop sky simulation on the BG/Q," In *Proc. of the Int.l Conf. High Performance Computing, Networking, Storage and Analysis (SC '12*), IEEE Computer Society Press, Los Alamitos, CA, USA, Article 4, 11 pages.

Petascale Simulation of Magnetorotational Core-Collapse Supernovae

Discretionary: Sean Couch, University of Chicago



Core-collapse supernovae (CCSNe) are the luminous explosions that herald the death of massive stars. Neutron stars, pulsars, magnetars, and stellar-mass black holes are all born out of these explosions. While their importance in galactic evolution cannot be underestimated, the mechanism of CCSN is not fully understood and remains one of the most important challenges for modern computational astrophysics.

Hubble Fellow Sean Couch conducted a series of 3-D magnetohydrodynamics simulations of the collapse of rotating, magnetic stellar cores using the FLASH multi-physics, adaptive mesh refinement simulation framework. Efforts at the ALCF include approximate treatments for neutrino physics, realistic progenitor rotation and magnetic fields. These simulations allow researchers to predict the spins, kicks, magnetic field strengths and alignments of newly formed neutron stars, pulsars, and magnetars, as well as the dependence of these parameters on originating conditions. Researchers also hope the simulations will reveal the relative importance of neutrino-driven convection and the standing accretion shock instability in driving the stalled supernova shock outward in realistic, rotating and magnetic progenitors.

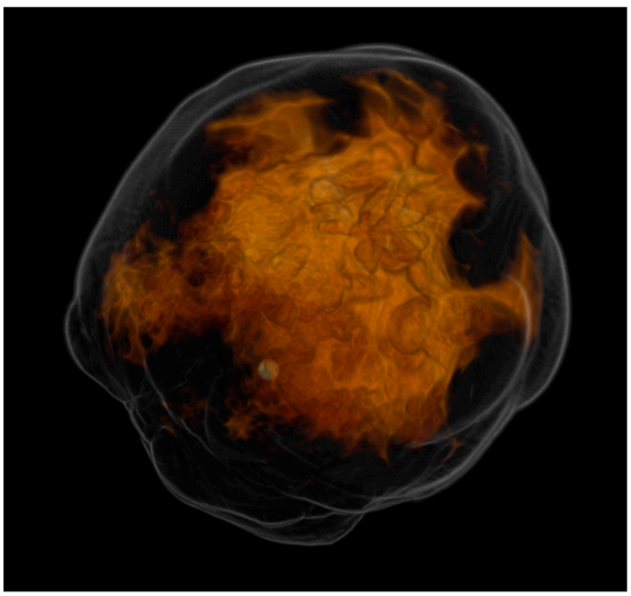


Figure 3.11: Volume rendering of the highest entropy plumes in a 3-D core-collapse supernova simulation. The edge of the nascent neutron star is shown as the faint blue sphere near the center. Entropy, a measure of the thermodynamic disorder, shows the buoyant convective plumes and turbulence that play a critical part in the core-collapse supernova mechanism. The surface of the outgoing shock wave is also shown in faint gray.

In 2012, Couch completed a sequence of 1-D, 2-D and 3-D simulations using 5.3 million core-hours on Intrepid. His studies showed the substantial differences in the behavior of the 2-D and 3-D buoyant plumes adding more evidence that 3-D characteristics make a substantial difference in the simulations and are necessary for realistic simulations. While more realistic simulations will likely require introducing more physics than in today's models, these studies confirm that moving to 3-D is a necessary step.

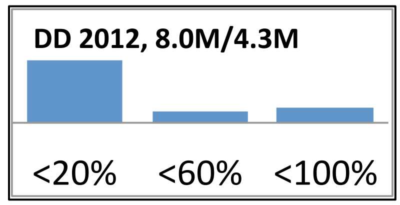
IMPACT: Proving the substantial difference in the characteristics in 2-D and 3-D simulations shows efforts should be targeted to 3-D studies.

ALCF Contribution: Work on FLASH through Early Science has improved time to solution at least 4x through threading and single core improvements.

Publication: Couch, Sean M. "On the Impact of Three Dimensions in Simulations of Neutrino-driven Core-Collapse Supernova Explosions," *Astrophysical Journal*, arXiv:1212.0010

Discretionary Projects Advancing Highly Scalable Algorithms

The DD Program, described in more detail in section 3.3, can be an immensely valuable tool to develop new tools and algorithmic capabilities. Below is a collection of short highlights on some interesting algorithmic advancements in 2012 completed in the DD program. It should be noted that many of the allocation totals are not uniquely for the projects listed. All but one of the projects has several related subprojects.



Cyclops Tensor Framework – Sustained Petaflops for Quantum Chemistry
Discretionary: Jeff Hammond, Argonne National Laboratory

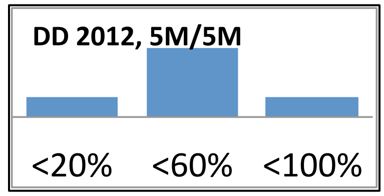
Cyclops Tensor Frameworks (CTF), enabled by PARTS DD on BGP and MADNESS ESP on BG/Q, is doing coupled-cluster (CCSD) twice as fast as NWChem on Cray XE6 (the only system where comparisons would be fair).

CTF is not a new implementation of coupled-cluster, but a completely new algorithm that leverages ten years of work by Jim Demmel in dense linear algebra. The approach is topology-mapped MPI such that it strongly scales on BG/Q in ways never seen with other coupled-cluster methods on large systems. So far, the method has scaled to eight racks i.e., 512K threads, on BG/Q and it will cross 1M-way parallelism very soon. Doing real science, CTF should be capable of running at >25% of theoretical peak -- meaning just one-third of Mira should be able to deliver sustained petaflop quantum chemistry.

IMPACT: This project was initiated by ALCF staff member, Jeff Hammond, as part of the 2011 practicum for Computational Science Graduate Fellowship fellows (Solomonik and Matthews). Beyond the training impact, this new framework will enable much larger scale quantum chemistry than previous algorithms could.

ALCF Contribution: This is an ALCF-driven project.

Publication: Edgar Solomonik, Devin Matthews, Jeff Hammond and James Demmel. Proc. 27th Intl. Parallel and Distributed Processing Symp. (IPDPS), Boston, Massachusetts, May 2013. "Cyclops Tensor Framework: reducing communication and eliminating load imbalance in massively parallel contractions"



Heuristic Static Load-Balancing Algorithm
Discretionary: Yuri Alexeev, Argonne National Laboratory

Achieving an even load balance is a key issue in parallel computing, and increasingly so as we enter the petascale supercomputing era. By Amdahl's law, the scalable component of the total wall time shrinks as the numbers of processors increases, while the load imbalance, together with the constant sequential component, acts to retard the scalability. Improved load balancing can be a simple and effective way to boost the scalability and performance of parallel algorithms. The problem of load balancing is especially important for large systems when the number of tasks is less than or equal to the number of nodes, and/or the tasks have uneven sizes. The solution presented by a simple dynamic load-balancing scheme will have a prohibitively high overhead on large processor counts, motivating the present heuristic static load-balancing (HSLB) method. HSLB has been successfully applied to the Fragment Molecular Orbital (FMO) method, implemented in the quantum chemistry package, GAMESS, and applied to the climate-modeling package, CESM.

The HSLB method consists of four steps. First, we collect benchmarking data related to the compute time of tasks. Second, we solve for the optimal parameters by a least-squares method based on our chosen scalability model. Third, we solve an integer optimization problem in order to obtain an optimal allocation of processors. Fourth, we allocate the optimal number of processors obtained from the optimization to each task. The key steps are the second and third. For step two, we developed a mathematical scalability model in accordance with Amdahl's law. The integer optimization problem in step three is solved using the software package, MINLP, developed at Argonne.

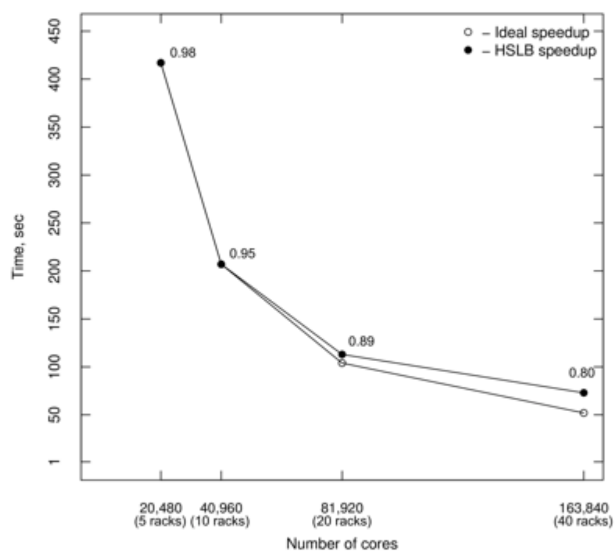


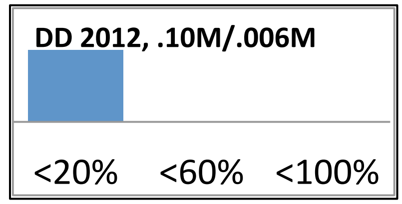
Figure 3.12: Ideal and observed scalability curves based on wall-clock time for the first FMO SCC iteration on Intrepid for 17,767 atoms protein-ligand bio-system.

IMPACT: Many applications stand to benefit from the HSLB approach, and we expect its user-base is expected to grow over time. Currently, HSLB is used in two INCITE projects: Computational Studies of Nucleosome Stability (PI George Schatz) and Attributing Changes in the Risk of Extreme Weather and Climate (PI Michael Wehner).

ALCF Contribution: This is an ALCF driven project.

Publication: Yuri Alexeev, Ashutosh Mahajan, Sven Leyffer, Graham Fletcher, Dmitri Fedorov, "Heuristic Static Load-Balancing Algorithm Applied to the Fragment Molecular Orbital Method," *The International Conference for High Performance Computing, Networking, Storage and Analysis (SC12)*, November 2012, Salt Lake City, Utah, IEEE Computer Society, 2012.

Expressing Graph Algorithms Using Generalized Active Messages
Discretionary: Nick Edmunds, Indiana University (Parallel Boost Graph Library)



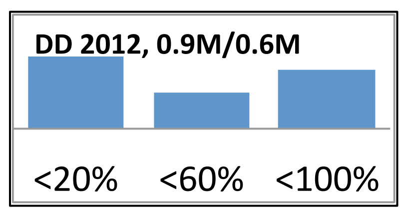
Recently, graph computation has emerged as an important class of high-performance computing application whose characteristics differ markedly from those of traditional, compute-bound, kernels. Libraries such as BLAS, LAPACK, and others have been successful in codifying best practices in numerical computing. The data-driven nature of graph applications necessitates a more complex application stack incorporating runtime optimization. The researchers have developed a method of phrasing graph algorithms as collections of asynchronous, concurrently executing, concise code fragments, which may be invoked both locally, and in remote address spaces. A runtime layer performs a number of dynamic optimizations, including message coalescing, message combining, and software routing. Practical implementations and performance results have been generated for a number of representative algorithms.

IMPACT: A successful active message abstraction for graph algorithms facilitates the growing body of HPC applications that need such algorithms. The abstraction of the capability yields a flexible and easily improvable tool for HPC.

ALCF Contribution: This project was given a DD award.

Publication: "Expressing Graph Algorithms Using Generalized Active Messages" *18th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming, N. Edmonds, A. Lumsdaine*

Scalable Parallel Particle Advection
Discretionary: Michael E. Papka, Argonne National Laboratory



Flow fields are an important product of scientific simulations. One popular flow visualization technique is particle advection, in which seeds are traced through the flow field. One use of these traces is to compute a powerful analysis tool called the Finite-Time Lyapunov Exponent (FTLE) field, but no existing particle tracing algorithms scale to the particle injection frequency required for high-resolution FTLE analysis.

A framework was developed to trace the massive number of particles necessary for FTLE computation. Processes are divided into groups, and are responsible for mutually exclusive spans of time. This pipelining over time intervals reduces overall idle time of processes and decreases I/O overhead. The parallel FTLE framework is capable of advecting hundreds of millions of particles at once, with performance scaling up to tens of thousands of processes.

IMPACT: A scalable algorithm where once there was none.

ALCF Contribution: This work was done through the SciDAC Scalable Data Management, Analysis, and Visualization (SDAV) discretionary allocation.

Publication: Boonthanome Nouanesengsy, Teng-Yok Lee, Kewei Lu, Han-Wei Shen, Tom Peterka, "Parallel Particle Advection and FTLE Computation for Time-Varying Flow Fields," *The International Conference for High Performance Computing, Networking, Storage and Analysis*, November 2012, Salt Lake City, Utah, IEEE Computer Society, 2012.

INCITE and ALCC Project Usage on Intrepid

CY2012 INCITE Usage

During CY2012, 803.8 million core-hours were delivered to INCITE projects, with 732 million core-hours committed to the projects. Of the 31 projects, 21 used all of their allocations, and 23 used over 90% of their allocations. A total of 89 million core-hours were unused by projects (12% of the total INCITE allocations). Several projects had unexpected challenges with their physics of their simulations which caused significant delay in their projects.

Figure 3.13 shows the core-hours allocated to, and used by, each INCITE project during CY2012. Projects that exceed their allocation run in a very low priority mode called "backfill." Jobs for a project with a negative allocation balance can run only if no other job on the system is able to run. Projects with very flexible workloads are able to adapt to the machine load and can be very successful using extra hours. A good example of this is given by the second allocation from the left in Figure 3.13, which was able to use much more than its allocation.

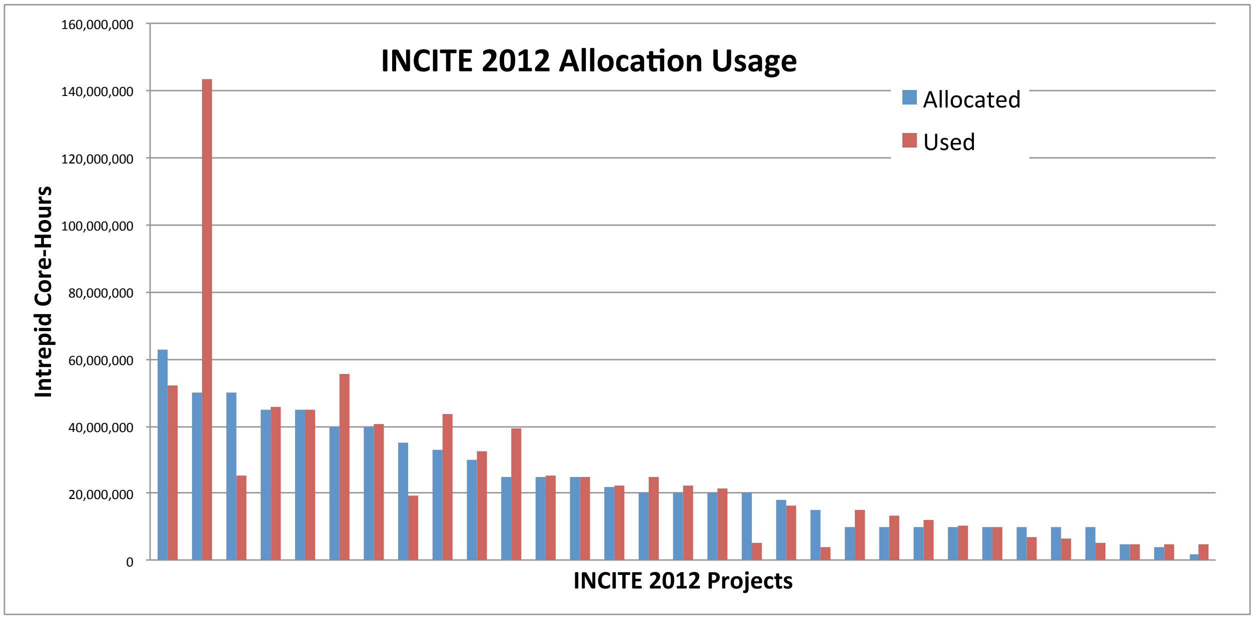


Figure 3.13: Core-hours allocated (blue) and used (red) by INCITE projects in CY2012

Figure 3.14 shows the number of INCITE projects binned by the percent of allocation used.

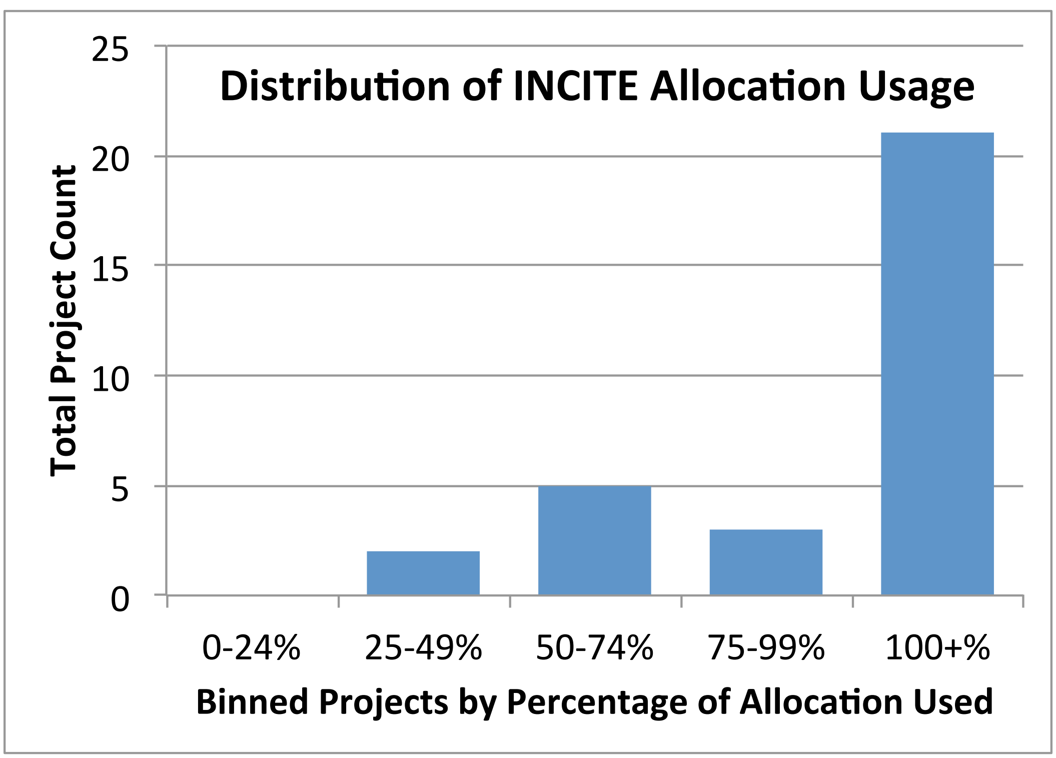


Figure 3.14: CY2012 INCITE Projects Counts by % of Allocation Used

ALCC Usage

The seven 2011/2012 ALCC projects started on July 1, 2011, with allocations lasting one year. Of 185 million core-hours allocated, 177 million core-hours were delivered, leaving 8 million core-hours of unused total time. The total time unused by projects, excluding projects using more than their allocation, is 22 million-core hours. Figure 3.15 shows a breakdown of the allocations compared to usage by project. Two low utilization projects had technical and modeling challenges with their applications that made using their entire allocation within the year difficult. Two other projects had personnel challenges that prevented full use of their allocations.

The ALCF worked closely with many of these projects, particularly those projects that were struggling to use their allocations. While total usage was lower than the total allocated, most projects made good progress through their allocations.

Nine 2012/2013 ALCC projects started on July 1, 2012, with allocations lasting one year. Halfway through the year, projects have used 67 million core-hours of the 247 million core-hours allocated (27.1% of the time has been used).

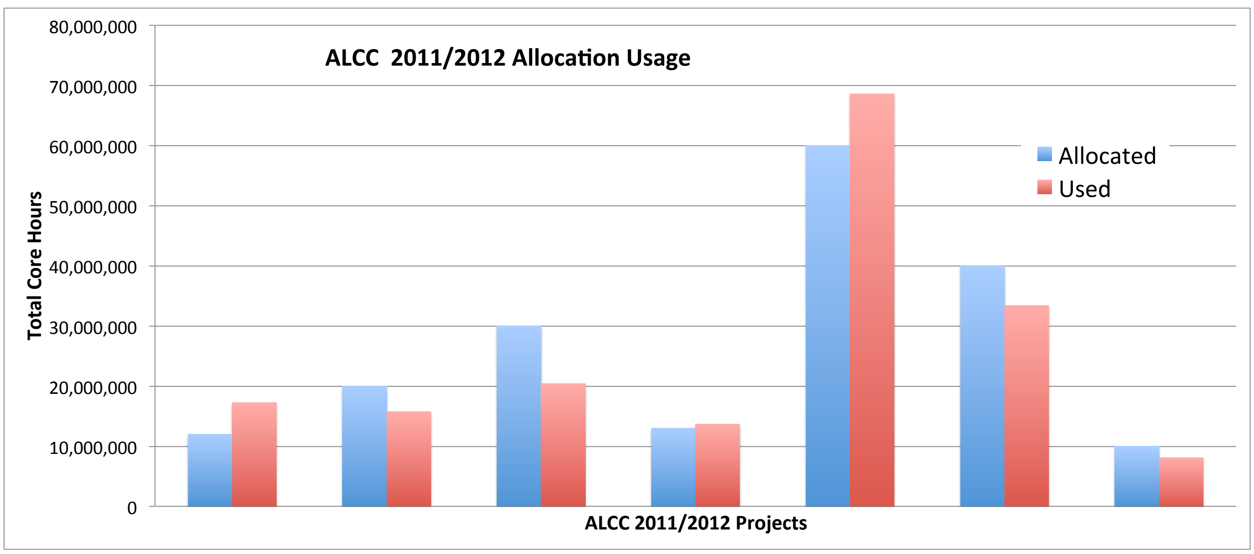


Figure 3.15: Core-Hours Allocated (blue) and Used (red) by 2011/2012 ALCC Projects

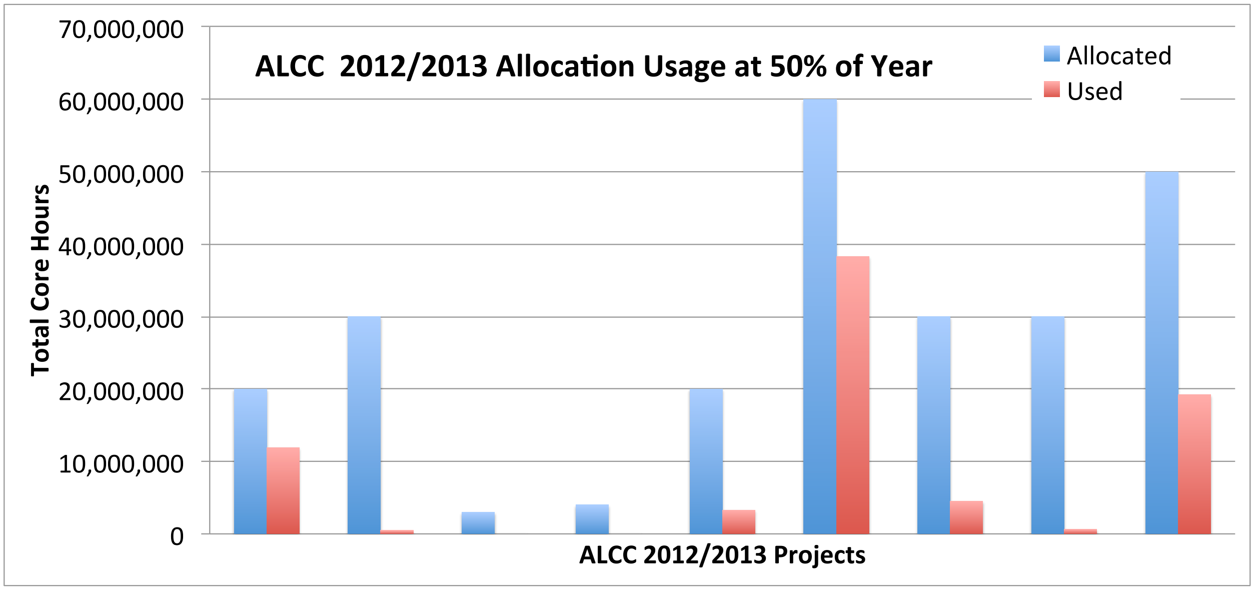


Figure 3.16: Core-Hours Allocated (blue) and Used (red) by 2012/2013 ALCC Projects. The ALCF is halfway through the allocation year.

3.3 Allocation of Facility Director's Reserve Computer Time

In this section we are interested in the strategic rationale behind use of DD time. The Facility should describe how the Director's Reserve is allocated and list the awarded projects, showing the PI name, organization, hours awarded, and project title.

ALCF Director's Reserve Program

The Director's Reserve, or Director's Discretionary (DD) program, serves the HPC community interested in testing science and applications on leadership-class resources. Projects are allocated in five categories: 1) INCITE or ALCC preparation, 2) code support and/or development, 3) science, 4) internal support, and 5) early science.

"INCITE and ALCC preparation" allocations are offered for projects that are preparing proposals. These projects can involve short-term preparation (e.g., run scaling tests for their computational readiness) or longer-term development and testing. "Code support and/or development" allocations support teams porting and supporting specific codes or projects developing new capabilities. This category includes the development, testing, and runs required for competitions such as the Gordon Bell Prize. Projects in this category have been responsible for bringing new capabilities to the ALCF. For example, PARTS has supported multiple libraries and software packages. This effort has fueled multiple, successful INCITE proposals and papers. The ALCF also allocates time to projects that might still be some time away from an INCITE award and needs time to prove a scientific case at scale and have a "science" problem worth pursuing.

Allocations are requested through the ALCF website and reviewed by the ALCF's Director, Director of Science, and Deputy Director of Science at weekly meetings. Input is collected from the ALCF staff where appropriate. Allocations are reviewed on their readiness to use the resources and are awarded time on a quarterly basis. The DD pool is under great demand and often the requested amount cannot be accommodated.

In 2012, the ALCF awarded or maintained allocations to 131 DD projects. Of these, 15 are labeled "Internal/Support," as these projects are used by the ALCF to support the facility. These include preparation for the Blue Gene/Q acceptance test, Catalyst, Performance Engineering and Operations support, and analysis and tracking of system data. These projects used 32.9M core-hours in CY2012.

The Director's Reserve allocated 433.4M core-hours and used 200.6M core-hours from January 1, 2012 to December 31, 2012. Of this 433.4M, the ALCF allocated 237M core-hours for early science projects. In 2012, the program office did not allocate all of the time for the ALCC program and gave the ALCF permission to use this time for early science preparation. Of the 237M allocated for Early Science, 83.2M were used. The ALCF did not pull back time from ESP projects, so the large allocation amount represents the initial allocation for all the projects plus the requests for additional allocations. Those projects using Intrepid were mostly focused on developing new models and new algorithms. In conclusion, after accounting for the 83.2M used by ESP and the 32.9M used for ALCF support, 84.0M core-hours were used by other discretionary projects.

The nature of the DD pool supports over-allocation, but it should be noted that 433.4M core-hours does not represent open allocations for the entire calendar year. A project might have a 1M core-hour allocation that only persists for three months, but that 1M core-hour allocation is counted entirely in the 433.4M core-hour total. Project PIs are experimenting with their ability to scale their applications or the scale of their science. The projects are not guaranteed the allocated time; instead, the time is provided on a first-come, first-served basis. DD projects run at a lower priority than INCITE or ALCC projects, which reduces the amount of time that is available for their use.

Lists of the CY2012 DD projects, including title, PI, institution, and hours allocated, are provided in Appendix B.

To provide an idea of the distribution of the allocations, Figure 3.17 provides a breakdown of the CY2012 usage by the standard INCITE science domains.

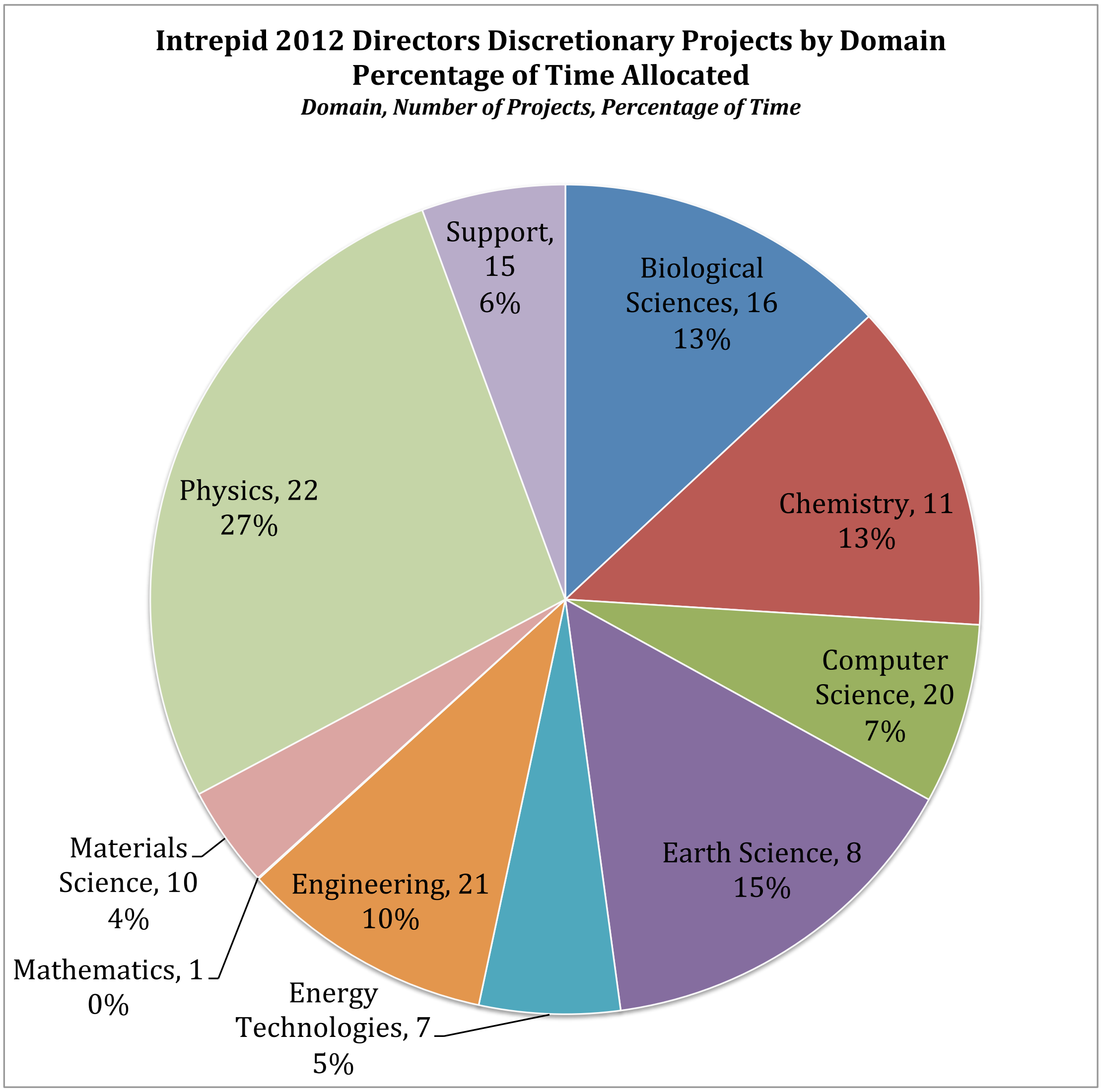


Figure 3.17: Distribution of the CY2012 DD core-hour usage by science domain. A total of 433.4M core-hours were allocated to 131 projects. Of this total, 237M allocated were to Early Science. The data labels are [Domain, Number of Projects, Percentage of Time Allocated].

Conclusions

ALCF continues to enable scientific achievements, consistent with DOE's strategic goals for scientific breakthroughs and foundations of science, through projects carried out on facility machines. Researchers participating in projects using ALCF resources published over 180 papers in the past year. ALCF projects have had success in a variety of fields, using many different computational approaches. They have been able to reach their scientific goals and successfully use their allocations. A number of the projects and PIs have subsequently received awards or have been recognized as achieving significant accomplishments in their fields.

The ALCF delivered 803.8 million core-hours to INCITE projects in CY2012, exceeding the 732 million core-hours committed. The Director's Reserve has been used not only to develop INCITE and ALCC proposals but also to conduct real science of strategic importance and drive development and scaling of key INCITE and ALCC science applications. The excellent ALCF support and the solid, high-performing ALCF resources have enabled the INCITE and ALCC projects to run simulations efficiently on HPC machines and achieve science goals that could not otherwise have been reached.

This page intentionally left blank.

Section 4. Innovation

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

ALCF Response

Listed below are the innovations and best practices carried out at the ALCF from January 1 through December 31, 2012.

4.1 Application Support Improvements

ALCF Library System

The ALCF maintains and makes available to its users a set of mathematics and parallel I/O libraries, performance tools, debuggers, and applications on its Blue Gene production systems.

The libraries, referred to collectively as the ALCF Library System, are ported and optimized for automatic building on the Blue Gene systems. The current set includes: ARPACK, BLAS, CBLAS, FFTW2, FFTW3, LAPACK, METIS, PARMETIS, PNETCDF, SCALAPACK, SILO, SZIP, ZLIB. Additional libraries may be added based on user needs.

Each library is built in two versions; one using the IBM XL compilers and the other using the GNU compilers. Compiler flags, including optimization levels, are set globally by default but may be tuned on an individual basis. The build system includes a framework for running automated library tests after each rebuild.

On Mira, the BLAS library build has been customized. The library is formed from a combination of routines from IBM's ESSL library along with Netlib-derived BLAS versions. Routines that are not tuned in ESSL can then be tuned locally from their source.

In addition, a compatibility build is available for BLAS which supports calling BLAS with 64-bit integer arguments. C wrappers to BLAS are automatically generated which adjust their arguments so that 64-bit integer arguments can be correctly passed on to 32-bit ESSL or other BLAS. A requirement is that the arguments' values fit in the acceptable range for 32-bit integers.

The automatic builds allow for easy updates of all the libraries after driver or compiler upgrades, or after source code changes to individual libraries. Having the libraries in a central location allows easy access for application developers.

ALCF Library Instrumentation and Tracking

The ALCF has also devised a method to track the use of the ALCF Library System by user and their application, which helps the ALCF to determine what libraries to support. The instrumentation was implemented in CY12 and the tracking methodology is currently being tested on Vesta, the BG/Q test and development system. The methodology has no impact on the performance of the library and is described below.

The ALCF software build system incorporates a library instrumentation tool. This tool annotates a library with descriptions of its object file locations, symbols that have been defined, and any desired additional text that can be used to identify the library.

Vesta execution is initiated by ALCF's TATU (Task Authorization and Tracking Utilities) package, which invokes the "tracklib" tool. tracklib extracts the library instrumentation from the executable, allowing it to discover which libraries and object files are used by the executable. This information is associated with the job's Cobalt ID and stored in a DB2 database. The database entry is signed with a cryptographic key to ensure integrity of the data collection.

Usage statistics can be determined for the instrumented libraries by querying the database, The tracking method has been implemented on Mira's testing and development system, Vesta.

Better Compiler Support
In order to provide ALCF users with the best possible programming environment, and to help optimize the performance of the users' applications, ALCF staff have become actively involved in the development of a BG/Q backend for the open source LLVM compiler toolchain. The advantages of this investment include:

- The ability to provide a modern, optimizing C/C++ compiler (Clang) that ALCF can ensure will remain up-to-date over the lifetime of the machine. Importantly, this can be done independent of the future business decisions made by IBM. In addition, many external libraries are already ported to Clang, and Clang already supports many common language extensions, thus easing the porting work that ALCF users must perform in order to compile their applications for Mira.

- The ability to develop optimizations and diagnostics specifically relevant to the users' applications. Regarding optimizations, for example, a BG/Q-specific software prefetching optimization is being developed and integrated with LLVM that promises to assist scientific applications with strided (and other non-trivial) memory access patterns. Regarding diagnostics, for example, ALCF worked with other members of the Clang community to develop the ability to provide MPI-specific type-mismatch warning messages. This capability is now available in Clang, and is used by the latest version of MPICH (developed at Argonne by the MCS Division), and the necessary changes have been made available within the Clang environment and provided to ALCF users. These diagnostic messages have already been used to find bugs in the HACC cosmology code.

Over the lifetime of ALCF-2, and continuing through ALCF-3, ALCF will use this ability to develop and deploy the LLVM compiler toolchain to respond to users' needs, make ALCF resources easier to use, and provide as much continuity as possible as the facility transitions toward exascale architectures.

Optimized FFT Libraries

ALCF is working with Spiral Gen to produce optimized FFT libraries for Mira. These libraries are expected to perform better than existing vendor-supplied (ESSL) and open source (FFTW) versions. Spiral Gen is working closely with a number of early science application teams (QMCPACK, HACC, CPMD) to create optimized routines for their needs.

Improving Debugging at Large Scale: Collaboration with Allinea

Allinea DDT has proven its ability to debug applications at the full scale of the largest machines. Its performance at scale is built on a tree-based communication architecture. Multiple I/O nodes could be used via DDT's tree, but the actual limitations within the individual I/O nodes required attention. This architecture offered an effective path to achieve groundbreaking debugging process counts on the IBM Blue Gene.

In a multi-year project, Allinea addressed performance bottlenecks specific to the IBM Blue Gene series architecture where debugging operations are required to be invoked on I/O nodes rather than on the compute nodes.

The first phase of the work reduced memory consumption of the debugger to a fraction of the initial quantity by sharing file-system based internal tables that each daemon was previously creating for its one target MPI process. This enabled up to 128 processes per I/O node to be debugged from a previous limit of 32.

The second phase of the work reduced both memory consumption and time by allowing daemons to control multiple MPI processes. This not only made possible debugging 256 processes per I/O node, but also improved performance substantially. The net result was a debugger capable of debugging the whole of the Intrepid BG/P system.

Further improvements were made to prepare for the IBM BG/Q, including replacing the IBM gdbserver with a native multiplexed client capable of communicating directly with the system tool daemons, to improve performance. The net result is an interactive debugging tool that is able to handle the 512 MPI processes per I/O node quickly, and able (albeit with linear slowdown) to reach as high as the system maximum of 8,192 MPI processes per I/O node.

This project enabled full scale and interactive debugging of Mira for MPI and hybrid MPI, and OpenMP codes. Furthermore the non-interactive offline mode of DDT also scales to the full system: it is able to pick out segmentation faults extremely quickly and provides a viable method of debugging at extreme scale (see Figure 4.1). This allows users to submit jobs to run overnight and provides detailed crash reports the following day.

Figure 4.1: Output from Offline Mode of Debugging a 16,384 Core Crash on Intrepid (Stack Trace Deliberately Anonymized).

4.2 User Services Improvements

Early Start on Documentation for Users

For the ALCF-2 project, ALCF started the steady state of documentation planning, creation, and maintenance earlier than ALCF-1. This ensured that users have documentation that meets their needs when Mira, and the other systems of ALCF-2 go into production. ALCF convened a cross-organizational planning committee on documentation that included representatives from the Catalysts, User Services, Operations, and Performance Engineering teams.

The committee classified user groups, created documentation types, examined the best internal and external examples of documentation, developed a project plan, and connected subject matter experts with prioritized sections of the restructured documentation. This team also reached out to users early on in the process to test the approach ALCF was taking.

Better Communication and Planning

Taking ideas from the ALCF technical help desk, User Services implemented a media help desk. An in-house communications staff on a rotating, weekly basis staffs this resource. The outcomes were: 1) a specific person is assigned responsibility for recognizing critical media needs, 2) ALCF staff and leadership have a reliable method to request media needs, and 3) this has allowed other communications staff to concentrate on facility writing or editing projects and writing ALCF science success accounts.

In collaboration with the ALCF administrative team, User Services developed a checklist system for conference and workshop logistics to clearly define what needs to be done and a chain of command for decision-making. This is called the Decision Maker's checklist. Each workshop or conference is assigned one or more decision makers and a proxy. The decision maker is the lead organizer for the event and the proxy is a member of the User Services team who serves in a facilitator role.

4.3 Operations Improvements

Test Harness
A particularly noteworthy innovation is ALCF's procedure for testing the facility machines. It is possible to test any of the ALCF Blue Gene computers, via a test harness, for stability, functionality, and variability, and to find regressions in many different places. The job of the harness is to check out, build, submit, and check certain specified jobs on the systems. These procedures will test the full life of a job and allow for a much more stable system. The harness interface also allows for easy use and simple reporting. The test harness has been used for acceptance of Mira, Cetus, and Vesta.

ISSF Power Monitoring
In 2012, ALCF took a two-day outage in order to install power monitoring. The Blue Genes are per power panel (8 or 9 racks per panel); all other infrastructure is per circuit, where a circuit typically feeds a rack PDU. Since then, this monitoring has been integrated into the building management system, allowing the ALCF to monitor "real time" power usage as well as trend it over time.

Challenger
Debug jobs tend to be small, often involve only a single node, are short in duration, and are run with the expectation of a nearly immediate turnaround. This directly conflicts with the principle use of Intrepid, that is, large, long, capability jobs where turnaround time is not critical. The initial solution was to dedicate one of Intrepid's 40 racks to debugging, and to put a queue with different policies ("prod-devel"). While this solution worked, it also blocked several torus configurations including full machine jobs, which required reservations. To resolve this, ALCF purchased an additional BG/P rack, called Challenger, which has been given maximum I/O (16:1 compute to I/O node ratio). The queues are tailored to debugging and the environment is maintained to match Intrepid's. While a causal relationship cannot be proven, utilization on Intrepid has increased since the Challenger install, and user reception has been very positive.

GPFS Configuration Improvements
As noted in previous OARs, ALCF tracks outages and job interrupts, including sources, at a detailed level and uses this information to drive the facility's primary improvement/debugging projects. During 2012, GPFS was the number one source of job interrupts, often by a factor of 3x. The GPFS administrator was able to determine that the problem was due to overloading the GPFS cluster manager. The ALCF purchased two redundant, dedicated cluster managers and reconfigured the GPFS filesystem. Since then, job interrupts due to GPFS outages have greatly reduced, and the job Mean Time To Interrupt metric has been greatly improved. Note that this is a measure of job MTTI, not the required facility MTTI reported above. Job MTTI gives us much finer grained information for tracking issues.

4.4 Participation in DOE HPC Financing Improvement Team

In conjunction with participants from ASCR, ALCF, NERSC, and OLCF, ALCF remains an ongoing coordinator and participant in the Financing Improvement Team's best practices group. All parties have participated and shared information in order to improve leasing and negotiating practices, with the end goal of obtaining the best rates possible for all the facilities. General areas of interest included lease options, requests for proposals, evaluation templates, processes, and best practices.

Within the area of leases, a number of options continued to be discussed and evaluated. These options include:

- Buyout – The ability to pay off the lease principal at any time, to save subsequent interest. A mechanism for pricing this option was developed and used for evaluating different lease options for ALCF-2.
- Buy down – The ability to pay ahead or buy down lease principal at any time, to save interest without having to renegotiate the lease.
- Discount point payment – The ability to pay a fee to be traded for a lower interest rate on a lease. A mechanism for pricing this option was developed and used for evaluating different lease options for ALCF-2.
- Flexible payment schedule – The ability to change payment schedules and amounts depending on revised funding profiles. This could include the ability to skip a lease payment or two without defaulting on the lease.
- Adding on to the lease – The ability to add equipment to an existing lease without the need to negotiate a new agreement.

Templates for evaluating and comparing different leases were enhanced and refined. One example is the template that provides the capability to measure the cost and benefit of paying an upfront fee to reduce the interest rate paid on the lease. Vendor lists continue to be updated and shared among the members of the group.

Shared best practices include specifics such as:

- Saving time and lining up competitors through pre-approval by setting up a "master lease" prior to the specific lease proposal.
- Separating equipment purchases from the financing (which is competitively bid).
- Obtaining a financing package from the equipment provider to serve as a competitive starting point to negotiate lease rates (provides lower risk since a maximum rate is locked at this point).
- Using consistent methods and market data for evaluating interest rate ranges and associated risks.

- Locking a fixed interest rate early, when doing so reduced interest rate risk by more than the market premium.

- Shared analysis tools for pricing and comparing diverse proposals.

A prescient best practice arrived at in 2012 involved discussion of pricing for various lease options. After considering many methods and much discussion, the group opted for lease bids and pricing based on U.S. Treasury rates, plus a spread. An alternative option using LIBOR (London Interbank Offer Rate) was ruled out on the basis that the index was self reported by banks and could be prone to unpredictable swings as well as potential manipulation. Subsequent to the group's decision, it was discovered that banks had been manipulating the LIBOR rate, which has since resulted in sanctions and penalties.

Argonne also broke new ground by providing a detailed financial information package to prospective lenders comprised of many materials often found in due diligence processes. The package provided information not available in the past in the hope it would increase confidence in ALCF as a borrower and result in lower lease rates. The lease rate for ALCF-2 came in lower than anticipated, which may be attributed in part to this financial information package.

During the ALCF-2 financing process, the team evaluated options from both the hardware supplier and multiple independent financiers. The ALCF-2 financing team discovered and exploited cost saving synergies available from the hardware supplier.

The Financing Improvement Team best practices group continues to share information and seeks to continually improve practices among the participants.

Conclusions

ALCF has performed innovations and improved operations in many areas in the past year.

- The center has improved the user experience with application software improvements, such as an optimized central repository for mathematics and I/O libraries and tools, compiler support, and scaling the DDT debugger.
- ALCF has improved communication and planning with the set up of a media helpdesk and a checklist to work out workshop logistics. A jumpstart on documentation for Mira during the project phase should also help the users. Systems operations have been improved by dedicating a single BG/P rack for debugging and reducing downtime due to GPFS. The test harness improved the efficiency for running QA (Quality Assurance) tests after hardware and software changes.
- Finally, collaborations with other facilities included participating in a DOE-sponsored HPC Financing Improvement Team.

This page intentionally left blank.

Section 5. Risk Management

Is the facility effectively managing risk?

ALCF Response

The ALCF has clearly demonstrated successful risk management in the past year for both project and operation risks. The risk management strategy is documented in the ALCF Risk Management Plan (RMP), which is reviewed and updated regularly to incorporate new ideas and best practices from other facilities. Risk management is a part of the ALCF culture, and the RMP processes have been incorporated into both its normal operations and all projects, such as the newly completed ALCF-2 project, the acquisition of a 10PF Blue Gene/Q system called Mira. Risks (proposed, open, and retired) are tracked, along with their triggers and mitigations (proposed, in progress, and completed), in a risk register managed by the Argonne Risk Manager. All risk ratings in this report are post-mitigation ratings. The ALCF currently has 34 open risks, with 1 high operational risk - funding uncertainty, which is managed by careful planning with the DOE program office and the continuation of austerity measures. The major risks tracked for the past year are listed below, with the risks that occurred and their mitigations described in more detail, along with new and retired risks, as well as the major risks that will be tracked in 2013.

Discuss how the facility uses its RMP in day-to-day operations, how often the RMP is reviewed or consulted, and what happens when a risk occurs. For this review the focus is on Operational risks, not Project risks.

The Facility should highlight various risks to include:

- *Major risks that were tracked for the review year;*
- *Any risks that occurred and the effectiveness of their mitigations;*
- *A discussion of risks that were retired during the current year;*
- *The mechanism used to track risks and trigger warnings;*
- *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.*

Note: this is a high level look at the risks, not a deep dive into the risk registry.

5.1. ALCF Risk Management

The ALCF uses the documented risk management processes, first implemented in June 2006 and outlined in its Risk Management Plan (RMP), for both operations and project risk management. The ALCF reviews and updates its RMP annually, and it is also updated during the year if changes at the ALCF necessitate an update (e.g. changes to Argonne risk management), as well as to incorporate new risk management techniques when they are adopted by the facility. The RMP is consulted at all monthly and individual risk meetings. Details of the RMP, including the attributes of each risk managed by the ALCF, have been detailed in past reports and will not be discussed further here. Risks are tracked in a risk register using the commercial management tool PertMaster, which integrates with the Primavera project management tool used to manage all large ALCF projects.

The ALCF currently has 34 open risks in its operations risk register. These risks include general facility risks, such as funding uncertainties, staffing issues, and safety concerns, as well as specific risks such as system component failures, availability of resources, and cost of electricity. Risk mitigation costs on the project side are developed using a bottom-up cost analysis, then used to set the contingency pool utilizing the PertMaster integration with Primavera. On the operations side, the costs are estimated by subject matter experts and used to inform management reserves. One of the benefits of the deep dives implemented in 2011 continues to be the dissemination and discussion of the basis of costs estimations, allowing the team to refine the mitigation costs within the risk register.

Completion of the ALCF-2 project
ALCF keeps project risks and operation risks in separate registers within the same tool. This allows us to build reports easily, as well as manage similar risks that have different impact between operations and the project. With the CD_4 approval of the ALCF-2 project in January 2013, 12 risks, primarily related to technology stability and performance, and facility infrastructure risks, are in the process of being transferred from the project risk register to the operations risk register. These risks will be updated to reflect operational impacts during the ALCF-2 Transition to Operations period (currently underway), so they are not considered open in the register at this time and are not covered in this report. They will be included as new risks in the 2014 OAR.

Risk Management in Day to Day Operations
Beyond the formal monthly and individual risk meetings, the ALCF has many informal risk discussions. Risks are identified and evaluated, and mitigation actions developed, for all changes at the facility, from installing a new piece of hardware, to changing the scheduling policy, to upgrading software. If the risks identified are short-term or minor, the risks are not added to the registry. Otherwise, if new significant risks are identified during the activity planning, they are added to the registry and reviewed at the next monthly risk meeting. Other tools beyond the risk register are used for managing risks in day-to-day operations. An example is the use of Work Planning and Controls (WPC) and Job Hazard Questionnaires (JHQ). Both of these are used to manage risks for activities where safety is a potential concern. JHQ are used for all staff and all contractors, and cover all work, both routine and non-routine. WPC are primarily used for any non-routine work and are developed in consultation with safety and subject matter experts. During planning meetings for non-routine activities, staff review the planned actions and evaluate for safety concerns. If a potential risk is identified, detailed discussions with the safety experts are scheduled, and procedures for mitigating the risks are developed, then documented in the WPC. The WPC is then used during the activity to direct the work. Another example of a tool used for operational risk management is the OARTool. This tool, and its backend database, assists the team with identifying potential risks and triggers, developing possible mitigations, developing data for calculating impacts and probabilities, and tracking outcomes of the risk management itself. The OARTool is used for risk management on a

weekly basis by the operations team. For details on the tool and its database, please see past OAR reports.

Beyond the operations of the machine, risk management is used in such diverse ways as evaluating and managing INCITE proposal risks (risk of too few proposals, risk of a lack of diversity across science domains, risk of too few capability proposals, etc.), safety risks in staff offices, project leasing risks (including the opportunity risk that interest rates could be lower than plan), support risks, etc.

Major Risks Tracked for the Review Year

Since Q4 of FY2010, the ALCF has expected several eventful years. This is due to Mira's deployment schedule in FY2012 and the planned growth of both ALCF staff and budget in order to bring the facility to full strength. As such, the ALCF was monitoring, and continues to monitor, a large number of major risks for the facility. All risk ratings shown are post-mitigation ratings. There were eleven major operation risks tracked for the CY2012, one with a risk rating of *High*, and ten with a risk rating of *Moderate.* Of these eleven risks, four were encountered and managed, and four risks were retired with the acceptance of Mira. The risks are color coded in the following way to assist with reading the table:

- Red risks were encountered and remain moderate or high risks.
- Orange risks were not encountered but remain moderate or high risks.
- Green risks are now retired.

ID	Title	Encountered	Rating	Notes
1059	Funding/Budget Uncertainties	Yes	High	Uncertainty and facility need for growth to deploy and accept Mira combined to make this a significant challenge in the past year. This risk remains a major concern for both this year and the out years as the facility prepares for ALCF-3.
25	Staffing Challenges	Yes	Mod	With the ongoing budget uncertainties and difficulties with retaining staff in the face of a recovering economy, this will continue to be a concern.
1049	Staff Retention	Yes	Mod	Between the salary freeze and budget concerns and the growth in high paying industry jobs for system administrators and programmers with HPC expertise, the ALCF has lost several staff members. This remains a concern.
1060	System Stability problems	Yes	Mod	Issues with the scratch file system have increased. With PVFS retired as a mitigation file system, the

				GPFS scratch file system is a single point of failure for Intrepid. The cost to mitigation the single point of failure for Intrepid is significant and not appropriate for a system that will be decommissioned within a year. Stopgap measures will be used for the remainder of Intrepid's active life. Options for mitigations for Mira are being explored. Stability remains a concern for Intrepid as it ages.
31	Facility Power Interruptions	No	Mod	This concern remains due to Argonne's aging infrastructure. Risk will be reevaluated this year as Mira enters production and Intrepid is decommissioned.
1056	System Stability Issues Due to Upgrades	No	Mod	Remains a concern as Intrepid ages.
1061	System Performance Issues	No	Mod	Remains a concern as Intrepid ages.
994	Difficulties Obtaining Financing for ALCF-2	No	Mod	Separating financing from build contract and competing the procurement worked well. Able to lock in a very favorable interest rate. With the full acceptance of Mira, this risk is retired.
995	Interest Rates on Leases Could be Higher	No	Mod	Able to lock in a very favorable interest rate. With the full acceptance of Mira, this risk is retired.
1062	Facility Problems with ALCF-2 Deployment	No	Mod	With the acceptance of Mira, this risk is retired as it is covered by other general facility problem risks for all ALCF resources.
1063	IBM May Back Out of Mira Contract	No	Mod	With the acceptance of Mira, this risk is retired.

Table 5.1: Major Risks Tracked for CY2012

Risks Encountered in the Review Year and Their Mitigations

The top risks encountered in the last 12 months are discussed below, along with the risk owner, its probability and impacts, a description of the actual problem that occurred, and the management of the risk.

1. Funding/Budget Uncertainties

1059: Funding/Budget Uncertainties	
Risk Owner	Mike Papka
Probability	High
Impact	Schedule: Low; Cost: Low; Perf: High
Risk Rating	High
Primary Management Strategies	Implement austerity measures. Work closely with DOE sponsors to manage expectations and scope.

Description

The Office of Science might not increase the ALCF budget as planned, or could reduce the ALCF budget below previous funding levels. An extended or full-year Continuing Resolution (CR) could prevent the ALCF from receiving planned funding. These scenarios could result in the inability to pay leases, contracts, and staff, as well as the inability to deploy future machines.

Evaluation

During the past year, Funding/Budget Uncertainties was the ALCF's highest risk, and it was also one of the risks that were encountered. The ALCF was supposed to be in a growth phase and planned a large down payment with the acceptance of Mira. A reduction of incoming funds carried a high impact.

Management

In conjunction with the DOE ASCR Budget Deep Dive, ALCF prepared full-year CR and reduced budget scenarios. In order to assure adequate funds to deploy and accept Mira, ALCF continued moderate austerity measures to provide maximum flexibility for the coming fiscal year. Hiring remained slowed to manage possible out year reduction in funds.

ALCF was able to obtain very favorable interest rates on the Mira lease; this reduced funding risk in the out years as the facility prepares for ALCF-3. However, funds in FY2013 were reduced from the original plan of record by 15%, and budget information for FY2014 and beyond indicate a further reduction in funds from the plan of record. Austerity measures remain in place and may be increased depending on the budget through FY2013.

2. Staffing Challenges and 3. Staff Retention

25: Staffing Challenges; 1049 Staff Retention	
Risk Owner	Mike Papka
Probability	Moderate
Impact	Schedule: Low; Cost: Low; Perf: Mod
Risk Rating	Moderate
Primary Management Strategies	Prioritized staffing needs, re-planned work, and re-tasked staff.

Description (Updated)

ALCF was, and remains, in a time of necessary growth, as it staffs up to operate Mira and begin preparations for ALCF-3. An aggressive staff ramp was originally planned for FY2010 through FY2012. This was extended due to budget reductions. ALCF risk evaluation identified two key risks associated with this ramp up, and both occurred in FY2012 as a result of the budget and funding uncertainties. The risks have been combined for this discussion, as they are related. The risks are:

- 25: Funds unavailable to hire to plan
- 1049: Unable to retain staff due to funding uncertainties, recovering economy and the increased demand for staff with compute expertise, and heavy workloads.

Evaluation

Funding uncertainties continue to lead to hiring delays. In addition, the recovering economy is opening more jobs in the industry for ALCF staff. In the past year, as a result, three ALCF staff left for higher paying jobs in industry. The hiring rate is not keeping up with the attrition. Because the facility was already understaffed, the existing staff continues to be overworked.

Management

It can be very challenging to hire experienced HPC staff. Because of this, the ALCF risk management team started several years ago preparing to execute mitigations prior to these risks occurring. When they occurred, the ALCF was able to successfully continue supporting existing projects and continue to deploy the new system even while understaffed. The ALCF has continued to use mitigations to manage this risk over the past year. Staff hires were prioritized, and once the austerity measures were implemented, top priority staff (replacement staff and key staff for the ALCF-2 project) were hired first. The ALCF has re-planned work as possible, delaying planned improvements and lower priority work. Staff have continued to be re-tasked, dropping lower priority tasks, expanding job descriptions, and, where possible, sharing staff with other divisions.

With careful and judicious management of this risk, the ALCF has successfully run the facility, and completed the deployment and acceptance of Mira and the two Test & Development systems Cetus and Vesta. The facility continues to remain understaffed. These risks will remain a concern for the next few years, and the ALCF will continue to carefully manage them.

4. System Stability Problems

1060: System Stability Problems	
Risk Owner	William Allcock
Probability	Moderate
Impact	Schedule: Mod; Cost: Mod; Perf: Low
Risk Rating	Moderate
Primary Management Strategies	Automatic monitoring for fault tracking. Weekly analysis of small issues to determine whether larger ones are developing. Work closely with vendors when problems appear. Dynamic mitigation development.

Description

Risk 1060 is a general risk that is used to track the health of the system and manage stability problems as they are encountered. One such problem was encountered late in CY2012 when corrupted files were found on the Intrepid high performance scratch file system during a migration of files to new disks installed to increase file space.

Evaluation

To address a previous need (insufficient disk space), ALCF added additional disks to increase available disk space on the Intrepid high performance scratch file system. During the migration of files to the new disks, some corrupted files were discovered. Because of the nature of the corruption (cross-linked files), it was not possible to simply delete the files from the scratch space. While the ALCF does not promise to protect files in scratch, the large quantity of important files in that space prevented a simple solution of rebuilding the file system. Instead, a series of file system checks (fscks) were run to locate and remove all cross-linked files. Unfortunately, there were issues with fsck and the version of GPFS currently in use, and the fscks were taking an extremely long time to run, then failing days into the process. Ultimately, the GPFS version was upgraded, the fscks were able to complete much quicker, all cross-linked files were located and removed, and the file system has stabilized. The management of the risk, as discussed below limited impact to the facility.

Management

A number of mitigation activities played into the resolution of the problem and ALCF's ability to limit the impact of the outage. A key one was that ALCF was able to utilize its alternative file system PVFS to allow a small number of projects to continue their work. PVFS was installed on the system as a backup mitigation for GPFS. It proved to be very valuable, although the ability to utilize it fully was limited both by its small size and limitations on what codes can be run out of it. In addition, ALCF opened up the ability to run out of the home directory space (typically not allowed, to protect the smaller, lower performance file system from being overwhelmed by jobs with large, rapid output). ALCF quickly prioritized projects and then worked with key projects (INCITE projects that had not fully utilized their time and ones with small amounts of I/O) to move as many of them as possible to PVFS or the home file system, and to quickly pull copies of their input decks from backups. Within a day, jobs were running again on Intrepid. Finally, as a risk mitigation strategy for a number of the technical risks, ALCF maintains a very close relationship with its vendors, and that played a major part in the resolution of this issue. Both IBM and DDN provide deep support, with direct access to the developers, and close to around the clock assistance. This would not have been possible without the partnerships that ALCF has developed with its vendors.

ALCF is exploring methods to reduce the probability of that this type of problem will occur again. Because Intrepid is nearing end-of-life, ALCF will not look into reducing the risk by modifying the single point of failure high performance scratch file system for it, but various options are being explored for Mira. ALCF is also talking with other sites to find out how they manage this issue and exploring possible implementations of their methods (e.g. regular online and offline fscks).

Retired Risks

The risks in the table below are risks that have been retired in the past year. Some were retired because the threat has been managed and/or no longer exists. Others were new short-term risks that were developed, monitored and tracked, but were retired within the same year.

ID	Title	Rating	Management Strategies	Notes
994	Difficulties Obtaining Financing for ALCF-2	Mod	Used a competitive lease bid process; participated in the DOE IFT (Integrated Finance Team); Requested IBM finance bid as part of build contract to use as a fallback plan.	Able to lock in a very favorable interest rate. With the full acceptance of Mira, this risk is retired.
995	Interest Rates on Leases Could be Higher	Mod	Competitive procurement of financing; reduce scope; budget reserves based on interest rate models.	Able to lock in a very favorable interest rate. With the full acceptance of Mira, this risk is retired.
996	Interest Rates on Leases Could be Lower (Opportunity)	Mod	Competitive procurement of financing.	Able to lock in a very favorable interest rate. With the full acceptance of Mira, this risk is retired. The savings in the out years will assist with managing budget uncertainties.
1062	Facility Problems with ALCF-2 Deployment	Mod	Monitor LLNL deployment; vendor facility visits; test early; independent analysis; prepare safety and contingency plans.	With the acceptance of Mira, this risk is retired as it is covered by other general facility problem risks for all ALCF resources.
1063	IBM May Back Out of Mira Contract	Mod	Run Intrepid and existing hardware for longer while locating another system; Verify whether sufficient management reserves exist to increase power/cooling, or to reduce scope; Utilize new infrastructure hardware with new system.	With the acceptance of Mira, this risk is retired.
1066	Unable to Recruit Enough Projects for Intrepid and Mira Production in 2013	Mod	Early call for interest (RFI); contacting RFI respondents to work with them to prepare them; DD allocations to strong candidates; Intensify outreach.	Mitigations resulted in oversubscription for CY2013 INCITE. Risk retired with full allocations across both machines.
1072	Phase 5 Acceptance Could Be Delayed	Mod	Work with leasing company to push out the lock date as far as possible; Work closely	Mitigations resulted in a lock date with high probability of reaching it,

	Beyond the Lock Date of the Loan and the Lease Rate will Increase.		with vendor to increase odds of reaching acceptance prior to the lock date.	and acceptance well within the deadline.

Table 5.2: Retired risks in 2012

New and Re-characterized Risks Since the Last Review

The ALCF risk culture leads to the identification of new risks and the re-characterization of existing risks on a regular basis. In the past year, seven new risks were identified, developed, and monitored and tracked as part of the Steady State risk register. Two of them, 1066 and 1072 (shown in green in the table), were both created and retired in CY2012. Because Intrepid is nearing end of life, the only changes to existing steady state risks were minor ones such as changing owners and setting the end of the risks to match the planned end of life for Intrepid.

ID	Title	Rating	Management Strategies	Notes
1066	Unable to Recruit Enough Projects for Intrepid and Mira Production in 2013	Retired	Early call for interest (RFI); contacting RFI respondents to work with them to prepare them; DD allocations to strong candidates; Intensify outreach.	Mitigations resulted in oversubscription for CY2013 INCITE. Risk retired with full allocations across both machines.
1067	Total Project Count Grows with Mira and Intrepid Beyond Staffing Levels	Mod	Work with code teams to develop strong, large projects; Reduce number of non-INCITE projects staff are exploring; Reduce support expectations.	New risk, based on concerns caused by the expectation of full INCITE allocations on both systems, as well as the greatly increased number of core hours available for allocation in the coming years.
1069	Eureka Hard Drive Failure Rate May Become Excessive and Replacements More Challenging to Locate	Low	Replace all hard drives with new modern hard drives (at ALCF cost); Vendor replaces hard drives with whatever equipment they can find (covered under maintenance).	Eureka is aging and the hard drives are failing more frequently. For now, this situation is being covered through maintenance. With the ALCF-2 GPU cluster, Tukey, coming online in 2013, the facility will likely continue to use maintenance and move critical users to Tukey sooner.
1072	Phase 5 Acceptance is Delayed Beyond the Lock Date of 6/30/13	Retired	Work with leasing company to push out the lock date as far as possible; Work closely	Mitigations resulted in a lock date with high probability of reaching

	and the Lease Rate Will Increase		with vendor to increase odds of reaching acceptance prior to the lock date.	it, and acceptance well within the deadline.
1075	Not enough Early Science Program project time	Mod	Work with projects to effectively utilize the system to accomplish science goals as quickly as possible; balance production date with time for ESP work.	Demand for INCITE and DD time on Mira was exceptionally high, increasing pressure to reduce Early Science period. ALCF needs to balance these conflicting demands.
1076	If the Interim Supercomputing Support Facility (ISSF) where ALCF-1 Resources are Currently Housed is Decommissioned, ALCF May Not Have an Appropriate Facility to Host its Disaster Recovery Resources	Low	Raise concerns with upper lab management (completed); Explore alternative locations on site. Explore alternative locations offsite (e.g., another lab to host ALCF disaster recovery resources); Accept the risk of a major loss of service or data.	An unexpected request by a project at Argonne came last year to take over the high bay where the ISSF is located; this would require the ISSF to be decommissioned once Intrepid was decommissioned. The decision has not yet been made to decommission the facility, so ALCF is currently exploring potential alternative locations.
1077	Cost of Decommissioning and Removing Intrepid Could Exceed Expectations	Low	Approach LLNL to see if they would like any part of Intrepid (completed); Explore non-standard options; Use management reserves.	ALCF has developed a list of potential methods of decommissioning Intrepid and is exploring the options. LLNL has indicated that they have no need for any of the Intrepid components.

Table 5.3: New and Re-characterized Risks from 2012

Projected Major Operating Risks for the Next Year

The current top operating risks projected for CY2013 are listed below, along with the current risk rating, and management strategies for the risk. ALCF expects to add more risks to this list as the time to begin operations for the ALCF-2 resources draws closer.

ID	Title	Rating	Management Strategies
1059	Funding/Budget Uncertainties	High	Careful planning, in conjunction with program office, for handling CR, leasing costs, and hires; maintain austerity measures; forward pay to reduce overall leasing costs.
25	Staffing Challenges	Mod	Continue to re-plan work, multi-purpose existing staff, and share staff with other divisions.
1049	Staff Retention	Mod	Look into non-salary based compensation (bonuses, office improvements, work from home); cross-train team members.
31	Facility Power Interruptions	Mod	Continued high-quality communications with Facilities Management and Services (FMS); Lab-wide improvements for the aging Argonne infrastructure are underway. ALCF will continue to monitor FMS efforts and provide data for improvement choices. This risk will be expanded to include the Theory and Computing Science (TCS) building.
1056	System Stability Issues Due to Upgrades	Mod	Continue to maintain close vendor relationships; develop rollback plans; deep tests on T&D systems; monitoring and notification.
1060	System Stability Problems	Mod	Continue to maintain close vendor relationships; monitoring and notification.
1061	System Performance Issues	Mod	Baseline; regression testing; monitoring.

Table 5.4: Projected Operating Risks for CY2013

Conclusions

The ALCF uses a proven risk management strategy that is documented in its RMP. This document is regularly reviewed and updated to reflect the dynamic nature of risk management, as well as new lessons learned and best practices captured from other facilities. Risk management is a part of the ALCF culture, and applies equally to all staff from the senior management to the summer students. A formal risk assessment is performed for every major activity, and informal ones are used for smaller activities within the ALCF. Risks are monitored and tracked using the commercial risk management tool PertMaster. Over the past year, seven risks were retired, seven new ones were added, and many open risks updated. ALCF expects to add to this risk register in CY2013 as the ALCF-2 resources are transitioned to operations. Beyond this, many tools are used to manage risks at the ALCF, particularly in the area of safety. The ALCF's effective risk management plan has contributed to the successful management of a number of significant risks that were encountered in the past year.

This page intentionally left blank.

Section 6: Summary of the Proposed Metric Values for future OARs

Are the performance metrics used for the review year and proposed for future years sufficient and reasonable for assessing Operational performance?

ALCF Response

The ALCF and the DOE have agreed to the 2013 metrics and targets as proposed in the April 2012 OAR report. These metrics and targets are reasonable measures of facility performance that are consistent with metrics and targets used at other facilities. For 2014, the proposed metrics and targets for the current production resources remain the same as for 2013. For the new BG/Q system, Mira, expected to go into production some time in 2013, its 2014 proposed targets will be the standard second year of operation targets, with the exception of the capability metric, which was modified to be a two-tier metric beginning with the 2013 metric. This metric ultimately provides a similar 20% of the system capability threshold, while accommodating the new partition sizes of Mira.

The facility should provide a summary table of the metrics and targets agreed upon for the review of calendar year 2013 and include the target and actual values of similar metrics used for 2012 for comparison. The facility should also provide metrics and targets under consideration for the review of CY 2014. Those will be finalized later in the year. The facility should discuss the rationale and use of proposed metrics and targets. This is also a place where a facility can suggest any long-term changes in the metrics and targets used for Operational Assessments.

6.1 Overview

The 2013 metrics and targets are reasonable measures of facility performance that are consistent with metrics and targets used at other facilities. For 2014, the proposed metrics and targets for the current production resources remain the same as for 2013 and are covered in Section 6.2. For the new BG/Q system, Mira, expected to go into production some time in 2013, its proposed 2014 targets will be the standard second year of operation targets, with the exception of the capability metric, which will be similar to the agreed upon 2013 Mira capability metric. ALCF is using a two-tier capability target for Mira. In 2013, it is 15% of the INCITE jobs running on 16.7% to 33.2% of the system (131,072 to 245,760 cores) and 5% of the INCITE jobs running on 33.3% - 100% of the system (245,761 cores or larger). The summed metric is 30% of the INCITE jobs will be run on 20% of the machine or larger. For 2014, it will be 20% of INCITE jobs running on 16.7% - 33.2% of the system (131,072 to 245,760 cores) and 10% of the INCITE jobs running on 33.3% - 100% of the system (245,761 cores or larger).

6.2 ALCF 2013 OA Performance Metrics

The OA performance metrics, 2012 targets and actuals, and agreed upon 2013 targets are presented in Table 6.1.

Area	Metric	2012 Targets	2012 Actuals	2013 Targets
User Results	User Survey – **Overall Satisfaction**	3.5/5.0	4.4/5.0	3.5/5.0
	User Survey – **User Support**	3.5/5.0	4.5/5.0	3.5/5.0
	User Survey – **Problem Resolution**	3.5/5.0	4.5/5.0	3.5/5.0
	User Survey – **Response Rate**	25%	33.2%	25%
	% User Problems Addressed Within Three Working Days	80%	94.4%	80%
Business Results	Intrepid Overall Availability	90%	95.9%	90%
	Intrepid Scheduled Availability	95%	98.5%	95%
	Intrepid INCITE Capability Usage (20%-100% of system)	40%	63.1%	40%
	Mira Overall Availability	-	-	80%
	Mira Scheduled Availability	-	-	85%
	Mira INCITE Capability Usage A (16.7%-33.2% of system)	-	-	15%
	Mira INCITE Capability Usage B (33.3%-100% of system)	-	-	5%

Table 6.1: Performance Metrics - 2012 Targets, 2012 Actuals, and Agreed Upon 2013 Targets

6.3 ALCF Proposed 2014 OA Performance Metrics

The OA performance metrics, agreed upon 2013 targets, and 2014 proposed targets are shown in Table 6.2. Additionally, the 2014 OA non-targeted, reportable-only metrics are shown in Table 6.3.

Area	Metric	2013 Targets	Proposed 2014 Targets
User Results	User Survey – **Overall Satisfaction**	3.5/5.0	3.5/5.0
	User Survey – **User Support**	3.5/5.0	3.5/5.0
	User Survey – **Problem Resolution**	3.5/5.0	3.5/5.0
	User Survey – **Response Rate**	25%	25%
	% User Problems Addressed Within Three Working Days	80%	80%
Business Results	Intrepid Overall Availability	90%	-
	Intrepid Scheduled Availability	95%	-
	Intrepid INCITE Capability Usage (20%-100% of system)	40%	-
	Mira Overall Availability	80%	90%
	Mira Scheduled Availability	85%	95%
	Mira INCITE Capability Usage A (16.7%-33.2% of system)	15%	20%
	Mira INCITE Capability Usage B (33.3%-100% of system)	5%	10%

Table 6.2: Performance Metrics - Agreed Upon 2013 and Proposed 2014 Targets

Area	Metric (no targets)
User Support Results	Summarizes training events and provides examples of in-depth collaborations between Facility staff and the User Community.
Business Results	Reports MTTI, MTTF, Utilization, and Usage for the past CY.
INCITE management	Reports reviewer survey responses and the proposal allocation results (# of proposals, # of awards, % awarded, # hrs requested/award, oversubscription)
Science Results	Track and report the number of publications written annually, projects are tracked for 5 years after award. Report on at least five significant scientific accomplishments, and the DD awards.
Innovation	Report on innovations that have improved operations.
Risk Management	Report on major risks.

Table 6.3: 2014 Reportable Only Metrics

Conclusions

The agreed upon 2013 metrics and targets are reasonable measures of facility performance that are consistent with metrics and targets used at other facilities. For 2014, the proposed metrics and targets for the current production resources remain the same as for 2012. For the new BG/Q system, Mira, the proposed 2014 targets will be the standard second year of operation targets, consistent with metrics and targets used at other facilities, with the exception of the capability metric. The agreed upon two-tier capability metric will provide 20% of the INCITE jobs will run on an average of 20% of the system. Achieving the agreed upon 2013 and the proposed 2014 targets will indicate that the Facility is performing up to stakeholder expectations, and ALCF anticipates being able to meet all metric targets.

This page intentionally left blank.

Appendix A – Calculations

Scheduled Availability Calculation Details

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

A significant event that delays a return to scheduled production will be counted as an adjacent unscheduled outage. Typically, this would be for a return to service four or more hours later than the scheduled end time. The centers have not yet agreed on a specific definition for this rare scenario.

Formula:

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

Overall Availability Calculation Details

Overall availability is the percentage of time a system is available to users. Outage time reflects both scheduled and unscheduled outages.

Formula:

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

ALCF Calculations

A simple example should make this clear. If on a particular day, there were 14 hours of scheduled maintenance, and two hours where the machine was down due to unexpected failures, there were 8 hours actually available (24 -14 – 2 = 8) resulting in 33.3% overall availability (8/24). There were 10 hours scheduled to be available (24-14 = 10),but there were actually 8 hours available, resulting in 80% scheduled availability (8/10).

To implement the above, ALCF tracks availability at the core-second level. Some hardware platforms have to shut down in order to do hardware maintenance, so the availability tends to be binary, either the whole system is available or it is not. This is not the case with Blue Gene. The Blue Gene is capable of taking individual node cards (32 nodes) off line for maintenance, while the rest of the machine continues to run. However, ALCF also takes into account scheduling policy. By policy, jobs smaller than 512 nodes (2,048 cores) are not allowed to run, and that is the smallest number of nodes that will be allocated. Therefore, if a single node were to fail for exactly one hour, it would be recorded as 2,048 cores * 3600 seconds = 7,372,800 core-seconds of down time. ALCF has only one production file system. Therefore, if it is down, the entire machine is considered to be down. There is an exception to the above. Sometimes, jobs can run successfully even when hardware is "considered down." Examples are test jobs run during a maintenance outage, or a job that was running during a file system outage that didn't attempt any I/O while the file system was down, and therefore, was able to complete successfully. When this happens, ALCF credits back the core-seconds for the job that occurred during the downtime. This is done to prevent reporting greater than 100% utilization.

To produce the actual numbers, ALCF takes the downtime data and calculates the scheduled and overall availability on a daily basis. The grand averages for a period are a straight average of the daily results. To produce the bar graph, the overall availability and the scheduled availability daily values are arithmetically averaged over 7-day intervals, and each bar in the graph represents one of those averages. So, for instance, the first bar in the chart is the average of days Jan 1 – Jan 7, the second data point is the average of Jan 8 – Jan 14, etc. If the number of days is not an even multiple of 7, the last data point is handled as follows: If there are more than half (4 or more) of the data points, a final data point is calculated from those values and plotted. If not (3 or fewer), those values are included in the previous data point, which becomes an average of between 8 and 11 data points. This is to avoid significant deviations of the last point due to a small average.

MTTI Calculation Details

Time, on average, to any outage on the system, whether unscheduled or scheduled. Also known as MTBI (Mean Time Between Interrupt).

Formula:

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

where time in period is start time – end time

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

MTTF Calculation Details

Time, on average, to an unscheduled outage on the system.

Formula:

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

where time in period is start time – end time

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

ALCF MTTI/MTTF Calculations

Calculating these values is fairly straightforward. ALCF finds any availability loss as described in the availability section that is for the whole machine; determines how long the loss lasted by wall-time, and whether it was scheduled or not; and then plugs all such losses into the guidance formulas.

***ALCF Utilization Calculation Detail*:** The Cobalt job scheduler writes out job records in the PBS format. Each night at midnight, a script runs and processes the day's records and imports the data into the internal accounting package, called clusterbank. Clusterbank records the time, date, duration, user, project and various other system parameters for every job run in the facility. Projects have attributes associated with them (INCITE, Discretionary, type of science, etc.). To calculate the utilization, queries are run against clusterbank to determine the daily total hours delivered to the various attribute classes and the total hours delivered. Jobs that cross day boundaries have the hours appropriately apportioned to the days. Combining this data with the availability data described in the availability section, the following value is computed on a daily basis:

$$Utilization = \frac{Core\ hours\ consumed}{Total\ core\ hours\ that\ were\ available} * 100$$

The daily values are then averaged as described in the availability section and plotted. The darker black vertical line marks the calendar year boundary.

Capability Calculation Detail: There is little calculation involved with the capability numbers. The data for everything except the job usage by size graph is simply the sum of the core-hours for qualifying jobs, with the plots showing daily values. Each bar in the job usage by size graph covers one week of data. The data is summed by type and then divided by the total for the week to determine the percentage.

This page intentionally left blank.

Appendix B – ALCF Director's Discretionary Projects
January 1, 2012 – December 31, 2012

Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
AbInitioC12_esp	Argonne National Laboratory	Steven C Pieper	Ab-initio Reaction Calculations for Carbon-12	Physics	7,500,000
accel_code_perf	Lawrence Berkeley National Laboratory	Ji Qiang	Performance Studies of the IMPACT-Z Parallel Particle-In-Cell Accelerator Modeling Code	Physics	150,000
Acceptance	ALCF	Susan Coghlan	Acceptance Tests	Support	5,000,000
ACESIII	University of Florida	Erik Deumens	Port ACES III and SIAL	Chemistry	500,000
AlchemicalDrugDesign	New York University	Michel A. Cuendet	Alchemical Drug Design by Ensemble Monte Carlo Molecular Dynamics Simulation	Biological Sciences	1,000,000
Allinea	Allinea Software	Kalyan Kumaran	Improved debugging memory usage for BG/P	Computer Science	300,000
AR-1-2012	Toyota Technological Institute at Chicago	Yang Shen	Molecular Dynamics Simulations for Exploring Androgen Receptor Antagonism, Drug-Resistant Mutations, and Antagonist Design	Biological Sciences	1,000,000
ART	University of Chicago, Fermilab	Nick Gnedin	SURPh: Simulating Universe with Realistic Physics	Physics	100,000
AURORA	ALCF	Ramesh Balakrishnan	A Computational Study of a proprietary parallel, arbitrary cell type code used by G.E. Global Research	Engineering	1,000,000
Autoignition_esp	Swiss Fed. Inst. Tech.	Christos Frouzakis	Direct Numerical Simulation of Autoignition in a Jet in a Cross-Flow	Engineering	5,000,000
BGQtools_esp	ALCF	Kalyan Kumaran	Enabling Petascale Science on BG/Q: Tools, Libraries, Programming Models, & Other System Software	Computer Science	4,000,000
BifurcationALCF	University of Canterbury	Timothy David	Simulation of cell coupling in arterial bifurcation	Biological Sciences	500,000
BrownLNCC	Brown University	Leopold Grinberg	Multi-scale simulations of deformable blood vessels	Biological Sciences	500,000
Bulk_Properties_esp	Iowa State U.	Mark Gordon	High Accuracy Predictions of the Bulk Properties of Water	Chemistry	22,000,000
Catalyst	ALCF	Katherine Riley	Catalyst	Support	5,000,000
CFDAnisotropic_esp	U. Colorado-Boulder	Kenneth Jansen	Petascale, Adaptive CFD	Engineering	12,000,000
CharmRTS	University of Illinois at Urbana Champaign	Laxmikant V. Kale	Charm++ and its applications	Computer Science	1,500,000
CNL_WSU	Washington State University	Soumik Banerjee	Molecular modeling of nanoscale transport pertinent to energy conversion and storage	Energy Technologies	250,000
CobaltDevel	ALCF	Narayan Desai	Cobalt Development	Support	1
ComPASSatBGP	Fermilab	Panagiotis Spentzouris	Community Petascale Project for Accelerator Science and Simulation Codes on BG/P	Physics	500,000
CompTurbFlow	Stanford University	Curtis Hamman	Complex Turbulent Flows	Engineering	250,000
Cosmology	Argonne National Laboratory	Stephen Kuhlmann	Computational Cosmology	Physics	800,000

Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
Crystal_Eng_disc	The University of Texas at Austin, The Fritz-Haber Institute of the Max Planck Society (Berlin), ALCF	J. R. Chelikowsky, N. Marom, A. Tkatchenko, L. Ghiringhelli, O. A. von Lilienfeld, J. Hammond	Toward Crystal Engineering From First Principles	Materials Science	2,000,000
CRYSTALLISATION	Univ College Dublin	Niall English	Massively parallel molecular simulation studies of nanoscale crystal formation	Physics	1,000,000
CTJSNIA	Stony Brook University	Alan Calder	Research Into the Systematics of Type Ia Supernovae	Physics	120,000
Cumulus_romps	Lawrence Berkeley National Laboratory	David Romps	Large-eddy simulations of atmospheric convection	Earth Science	700,000
DarkUniverse_esp	Los Alamos National Laboratory	Salman Habib	Cosmic Structure Probes of the Dark Universe	Physics	26,000,000
DFT-defects-UO2	Argonne National Laboratory	John J. Low, Marius Stan	Modeling of Defects in Materials for Energy Applications	Energy Technologies	500,000
DirectNoise	GE	Umesh Paliath	Enabling Green Energy and Propulsion Systems via Direct Noise Computation	Engineering	5,000,000
disk_study	ALCF	Justin Binns	Intrepid disk I/O analysis	Support	750,000
DNS-velvor	ALCF	Ammar Abdilghanie	Development of a novel methodology for petascale simulation of low Mach combustion	Engineering	200,000
Ducts	IIT (Chicago), ANL (Chicago) and KTH (Stockholm)	Hassan Nagib	Aspect ratio effects in turbulent duct flows studied through direct numerical simulation	Engineering	500,000
DynamicDevel	ALCF	William Scullin	Development and Improvement of Dynamic Programming and Application Environments at Scale	Computer Science	500,000
EESS_Interface	Argonne National Laboratory	Larry Curtiss	First Principles Calculations of Interfaces in Electrical Energy Storage Systems	Energy Technologies	10,000,000
eon	University of Texas at Austin	Graeme Henkelman	Long time simulations of catalytic reaction dynamics	Chemistry	1,000,000
Epigenetics	University of Utah	Nadeem Vellore	Nanoscale Biomolecular simulation for understanding of peptide recognition by epigenetic protein	Biological Sciences	300,000
ExaFMM-2012	Mechanical Engineering Dept, Boston University	Lorena A. Barba	Porting ExaFMM to BG/P architecture and turbulence validation studies	Engineering	1,000,000
Fathom	Argonne National Laboratory	Tim Tautges	Fathom: geometry, mesh generation, and related technologies	Computer Science	500,000
Femtomagnetism	Indiana State University	Guoping Zhang	First-principles calculation of laser induced ultrafast magnetism	Physics	2,500,000
fmo4bgp	ALCF	Graham Fletcher	Fragment Molecular Orbital Calculations on BG/P	Chemistry	2,000,000
fmobench	ALCF	Yuri Alexseev	FMO benchmarking and tuning	Biological Sciences	500,000

Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
FusionFS	Illinois Institute of Technology	Ioan Raicu	Distributed File Systems for Exascale Computing	Computer Science	900,000
GB2012	The University of Texas at Austin	George Biros	Parallel N-body algorithms for computational statistical mechanics	Biological Sciences	1,000,000
GFDL_esp	GFDL	Venkatramani Balaji; Tim Williams; Chris Kerr	Climate-Weather Modeling Studies Using a Prototype Global Cloud-System Resolving Model	Earth Science	7,500,000
graph500	Indiana University	Andrew Lumsdaine	Graph500 benchmark run on Intrepid	Computer Science	1,250,000
GroundMotion_esp	USC	Thomas Jordan	Using Multi-scale Dynamic Rupture Models to Improve Ground Motion Estimates	Earth Science	17,500,000
HEDhydro_RMRT	University of Michigan	Eric Johnsen	High-fidelity high-energy-density hydrodynamics simulations of shocks interacting with material discontinuities	Engineering	200,000
HEIGHTS-3D	Argonne National Laboratory	Ahmed Hassanein	Advanced 2D and 3D Laser-Produced and Electrically Driven Plasma Processes Modeling for EUV Lithography and other advanced plasma applications	Physics	300,000
Heron	NREL	John Michalakes, NREL	Wind turbine array fluid dynamic and aero-elastic simulations	Energy Technologies	1,000,000
hmd_cs	ETH Zurich	Martin Stiebritz	Oxygen-sensitivity problem of hydrogenases	Energy Technologies	1,000,000
HohokamWaterMgmt	Argonne National Laboratory	John T. Murphy	Hohokam Water Management	Earth Science	25,000
HSCD_esp	University of Chicago	Alexei Khokhlov	High-Speed Combustion and Detonation (HSCD)	Chemistry	15,000,000
i-SNSDCs	University of California, Los Angeles	Anastassia N. Alexandrova	Multi-scale modeling of catalytic interfaces based on 2-D sub-nano surface-deposited clusters	Chemistry	2,000,000
IBM-performance	ALCF	Kalyan Kumaran	BG/P performance runs carried out by IBM employees	Support	500,000
intrepid_namd_scale	University of Iowa	Adrian H. Elcock	MD Simulations of Bacterial Cytoplasm	Chemistry	273,000
IOFSL	Argonne National Laboratory	Rob Ross	I/O Forwarding Scalability Layer	Computer Science	1,500,000
ITERG8_GlobalPIC	Princeton University, PICSciE and PPPL (Princeton Plasma Physics Laboratory)	William M. Tang	Global Gyrokinetic Particle-in-Cell Investigations of ITER-relevant Fusion Plasmas	Physics	5,000,000
LatticeQCD_esp	Fermilab	Paul Mackenzie	Lattice Quantum Chromodynamics	Physics	17,500,000
LDAV	Argonne National Laboratory	Joseph Insley	LDAV Visualization Contest	Computer Science	500,000
Leap2Petascale2012	ALCF	Richard Coffey	Leap to Petascale 2012 project	Support	10,000
Licrys	University of Chicago and Argonne National Laboratory (IME)	Juan J. de Pablo	Liquid Crystal Based Functional Materials	Materials Science	2,000,000

Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
LLVM	ALCF	Hal Finkel	LLVM Compiler Tools for the BG/Q	Computer Science	50,000
LowPriority	ALCF	Katherine Riley	Low Priority	Support	5,000,000
MADNESS_MPQC_esp	Oak Ridge National Laboratory	Robert Harrison	Accurate Numerical Simulations Of Chemical Phenomena Involved in Energy Production and Storage with MADNESS and MPQC	Chemistry	7,500,000
Mat_Design_esp	Argonne National Laboratory	Larry Curtiss	Materials Design and Discovery: Catalysis and Energy Storage	Materials Science	8,000,000
MCA	PNNL	Yousu Chen	Massive Contingency Analysis	Energy Technologies	1,000,000
MembraneChannels	University of California - San Diego	Igor Tsigelny	Simulation and modeling of membranes interactions with unstructured proteins and computational design of membrane channels for absorption of specified ions	Biological Sciences	2,000,000
MFiX	DOE	Aytekin Gel	Achieving Scalable, Clean & Efficient Coal Gasifier Designs with High Performance Computing	Engineering	200,000
MiningSysLogs	ALCF	Michael E. Papka	Data Mining ALCF System Logs to better understand System Faults, Usage and Application Characteristics	Computer Science	100,000
MLearning4CCS	ALCF	O. Anatole von Lilienfeld	Machine Learning for the exploration of chemical compound space	Chemistry	1,000,000
MolSaltMix	Sandia National Laboratories, Albuquerque, NM	Sai Jayaraman	Design of low melting heat-transfer-fluids via adaptive screening in composition space based on alchemical interpolations between replicas.	Materials Science	700,000
MRCCSN	University of Chicago, Fermilab	Sean M. Couch	3D Simulations of Magnetorotational Core-Collapse Supernovae	Physics	7,500,000
MRDFT	ALCF	Alvaro Vazquez-Mayagoitia	Wavelet basis set in Density Functional Theory Methods for Photoelectric Materials	Materials Science	900,000
MTCScienceApps	Argonne National Laboratory	Michael Wilde	Many Task Computing Science Applications	Computer Science	6,700,000
MultiscaleMolSim_esp	University of Chicago	Gregory Voth	Multiscale Molecular Simulations at the Petascale	Biological Sciences	17,500,000
NAMD_esp	University of Chicago	Benoit Roux	NAMD - The Engine for Large-Scale Classical MD Simulations of Biomolecular Systems Based on a Polarizable Force Field	Biological Sciences	22,500,000
NEK5000	Argonne National Laboratory	Paul Fischer	NEK5000	Physics	1,000,000
NekCEM	Argonne National Laboratory	Misun Min	Electromagnetics	Physics	1,300,000
NekLBM	Argonne National Laboratory	Misun Min	Lattice Boltzmann Simulations for Fluids	Engineering	400,000

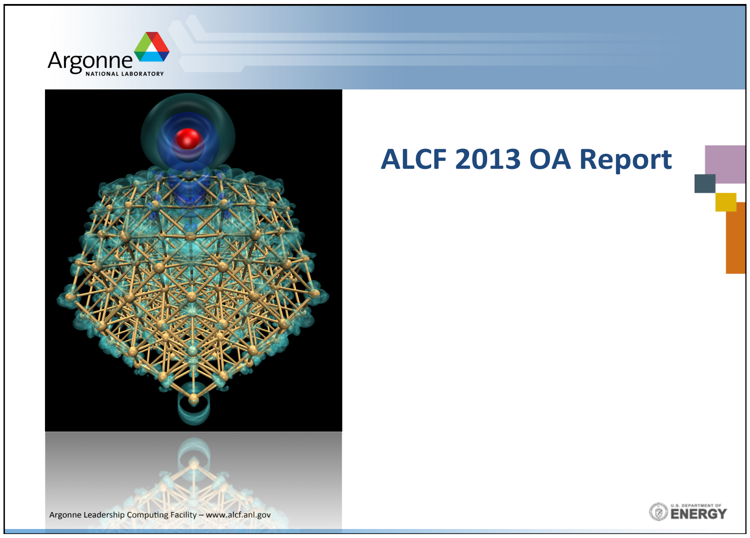
Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
NEURO_EVOLVE	Indian Institute of Science / University College London	Dr. Biswa Sengupta	The role of constraints in the design of the nervous system	Biological Sciences	80,000
neuronq	Yale University	Michael Hines	High Performance Neuron simulations on the Blue Gene/Q	Biological Sciences	1,000,000
NRCM_DD	Argonne National Laboratory	V. Rao Kotamarthi	Dynamic Downscaling of Climate Models	Earth Science	11,250,000
nu-xanalytics	Northwestern University	Alok Choudhary	Scalable data analytics and I/O	Computer Science	250,000
nwchem4bgp	University of Chicago	Jeff Hammond	Testing and tuning NWChem for Blue Gene/P and studies of nonlinear optical properties of conjugated chromophores	Chemistry	4,000,000
OnSite_Workshop	ALCF	Richard Coffey	Project for ALCF Workshops Onsite	Support	20,000
OpenFOAM-ALCF	ALCF	Ramesh Balakrishnan	OpenFOAM based Computational Fluid Dynamics Simulations At the Argonne Leadership Computing Facility	Engineering	1,500,000
Operations	ALCF	Bill Allcock	Systems administration tasks	Support	0
Operations_Test	ALCF	Cheetah Goletz	Operations Infrastructure Testing	Support	500,000
ORNLAstro	Oak Ridge National Laboratory	Bronson Messer	Porting and Scaling of CHIMERA and GenASiS	Physics	1,500,000
Oxygen_defects	Argonne National Laboratory	Olle Heinonen	First-principle investigations of oxygen defects in metal-oxide-metal heterostructures	Materials Science	500,000
ParallelSpectral	University of Michigan	Benson Muite	Parallel Pseudo Spectral Methods	Mathematics	200,000
PARTS	ALCF	Jeff Hammond	PArallel Run-Time Systems	Computer Science	8,000,000
PBGL	Indiana University	Andrew Lumsdaine	Parallel Boost Graph Library	Computer Science	100,000
Pepag	University of Chicago and Argonne National Laboratory (IME)	Juan J. de Pablo	Macromolecular folding and aggregation	Biological Sciences	2,000,000
Performance	ALCF	Kalyan Kumaran	Performance	Support	5,000,000
PHASTA_GB	U Colorado / ANL	Ray Loy	PHASTA-GLEAN Gordon Bell submission	Engineering	1,500,000
PlasmaMicroturb_esp	PPPL	William Tang	Global Simulation of Plasma Microturbulence at the Petascale & Beyond	Physics	17,500,000
pocket_druggability	University of Kansas	John Karanicolas	Predicting druggability of protein interaction sites by pocket optimization	Biological Sciences	500,000
prec_sense	Argonne National Laboratory	Laura Zamboni	Sensitivity and uncertainty of precipitation of the GFDL high resolution model	Earth Science	2,500,000
qe_parallel_ph	Argonne National Laboratory	William Parker	Phonon Parallelization in Quantum ESPRESSO	Materials Science	1,000,000
QLG-Turbulence	William & Mary	George Vahala	Quantum Lattice Algorithm for Quantum Turbulence	Engineering	700,000

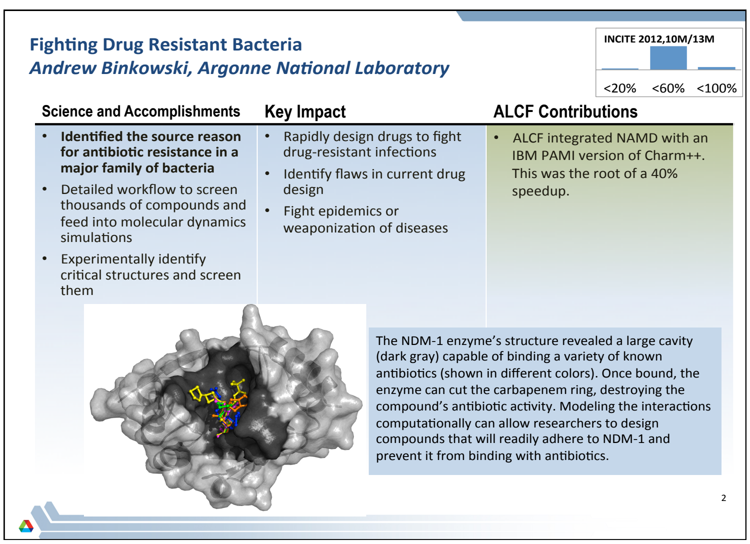
Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
qmc	University of Illinois at Urbana-Champaign	Jeongnim Kim	QMC development on BG	Physics	495,000
QuantumWater	Princeton University, PICSciE and PPPL (Princeton Plasma Physics Laboratory)	Robert A. DiStasio Jr.	Quantum Effects in Atomistic Simulations of Water	Physics	500,000
Reactor_Core_alcc	Argonne National Laboratory	Michael Smith	Scalable, Explicit Geometry, Whole Core Nuclear Reactor Simulations	Energy Technologies	10,000,000
Regional_Climate	National Center for Atmospheric Research	Greg Holland	Simulating Regional Climate at Convection Permitting Resolution	Earth Science	18,000,000
Repast	Argonne National Laboratory	Michael North	Exascale Agent-based Modeling System	Computer Science	1,000,000
RTInstability	SUNY at Stony Brook	James Glimm	Large Eddy Simulation of turbulent mixing by Rayleigh-Taylor instability	Engineering	1,500,000
SDAV	Argonne National Laboratory	Michael E. Papka	SciDAC Scalable Data Management Analysis and Visualization	Computer Science	900,000
SecCG	California Institute of Technology	Thomas F. Miller, III	Coarse-grained simulations of Sec-facilitated protein translocation and membrane integration	Biological Sciences	6,000,000
Shell_Oil	Shell International E&P, Inc.	Detlef Hohl	Modeling oil properties with molecular dynamics	Materials Science	300,000
SPaSM_materials	Los Alamos National Laboratory	Timothy C. Germann	SPaSM molecular dynamics simulations of material dynamics	Materials Science	700,000
Spray_and_soot	Stanford University	Dr. Colleen Kaul	DNS of Spray-Soot-Turbulence-Chemistry Interactions	Chemistry	750,000
StellarTurbulence	Los Alamos National Laboratory	Casey A. Meakin	An Investigation of Turbulence and Mixing in Stellar Interiors: Code Performance Study for INCITE	Physics	500,000
SU_Climate	Argonne National Laboratory	Laura Zamboni	Sensitivity and uncertainty of precipitation of a climate model	Earth Science	6,908,880
SupernovaVandV	University Of Chicago	Don Lamb	Toward Exascale Computing of Type Ia and Ib, c Supernovae: V&V of Current Models	Physics	1,000,000
TAU	ParaTools, Inc.	Sameer Shende	Parallel Performance Evaluation Using the TAU Performance System	Computer Science	600,000
TestHarness	ALCF	Eric Pershey	TestHarness	Support	2,000,000
TotalView	TotalView Technologies	Peter Thompson	TotalView Debugger on Blue Gene P	Computer Science	250,000
turb	Jet Propulsion Laboratory, California Institute of Technology	Georgios Matheou	Direct numerical simulation of stationary homogeneous sheared turbulence	Engineering	400,000

Project Name	PI Institution	PI Name	Short Title	Science Field (Short)	Allocation Amount
Turb_mix	Texas A&M University	Diego Donzis	Simulations of turbulence structure and turbulent mixing	Engineering	300,000
TurbChannelFlow_esp	U. Texas	Robert Moser	Petascale Direct Numerical Simulations of Turbulent Channel Flow	Engineering	5,000,000
TurbMix	Georgia Institute of Technology	P.K Yeung	Simulations of turbulence structure and turbulent mixing	Engineering	200,000
TurbNuclComb_esp	University of Chicago	Don Lamb	Petascale Simulations of Turbulent Nuclear Combustion	Physics	25,000,000
TurbulentMultimix	Stanford University	Sanjiva Lele	Turbulent Multi-Material Mixing in the Richtmyer-Meshkov Instability	Engineering	5,000,000
ucl_qmc	University College London	Dario Alfe	Quantum Monte Carlo methods for solids and liquids	Materials Science	1,000,000
UltraVis	Ultrascale Visualization Institute	Robert Ross	UltraVis Institute Research for Extreme Scale Visualization	Computer Science	1,600,000
UNRES-DEV	Baker Laboratory of Chemistry, Cornell University, Ithaca, NY 14853-1301	Harold A. Scheraga	Optimization for memory and development of the UNRES software for coarse-grained molecular dynamics simulations of proteins	Biological Sciences	200,000
User_Services	ALCF	Richard Coffey	User Services	Support	0
USO_InterimDD	ALCF	Richard Coffey	User Services Organization Interim Project	Support	10,000
visualization	ALCF	Michael E. Papka	Visualization and Analysis Research and Development for ALCF	Support	500,000
Wall_Turb_DNS	NASA Langley Research Center	Gary N Coleman	DNS of wall-bounded turbulence	Engineering	1,000,000
				Total DD	**433,441,881**

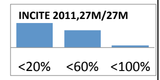
This page intentionally left blank.

Appendix C – Strategic Results Slides





Lithium Air Battery
Jack Wells, Oak Ridge National Laboratory

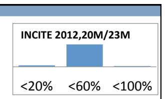


Science and Accomplishments	Key Impact	ALCF Contributions
• Study the use of propylene carbonate (PC) as a potential electrolyte • ***Simulations found that PC is not stable and chemical reactions decrease efficacy*** • Simulations identified why these results had not been seen in experiment • Excluded PC as a viable candidate for Lithium/Air batteries	• Eliminate from consideration a material previously assumed to be a good electrolyte candidate for Li/Air batteries • Design a battery that weight-for-weight stores 10x the energy of lithium ion batteries	• Helped enable ensemble simulations with IBM and ANL researchers

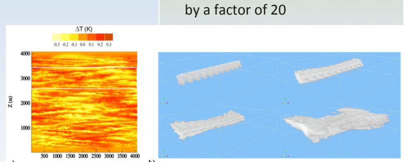
Publication (cover article): Laino, T., and Curioni, A., (2012), "A New Piece in the Puzzle of Lithium/Air Batteries: Computational Study on the Chemical Stability of Propylene Carbonate in the Presence of Lithium Peroxide," *Chem. Eur. J.*, 18: 3421. doi: 10.1002/chem.201290042

3

Large-Eddy Simulations of Contrail-to-Cirrus Transition
Roberto Paoli, CERFACS



Science and Accomplishments	Key Impact	ALCF Contributions
• Contrails are ice clouds formed from the exhaust of aircraft engines and become indistinguishable from natural clouds • The goal is to identify the key processes that determine contrail behavior • **Found that ambient turbulence controls initial formation and radiative transfer becomes main driver**	• Address aviation impact on climate	• ALCF staff helped on the activation of a Globus Online endpoint with Globus Connect Multi-User, increasing transfer rates 6x • ALCF tuning improved I/O performance at capability levels by a factor of 20



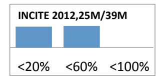
a) Potential temperature fluctuations of the atmosphere in a vertical plane in the middle of the computational domain (t=5 min. old contrail). The white boxes indicate the portion of the domain where the contrail is initially located. b) 3-D iso-surface of ice mixing ratio showing the contrail transformation into a young cirrus cloud. From top to bottom, left to right: t=5 min; t=10 min; t=20 min and t=40 min old contrail.

Publication: R. Paoli, O. Thouron, J. Picot, and D. Cariolle "Large-eddy simulations of contrail-to-cirrus transition in atmospheric turbulence," *65th Annual Meeting of the APS, Division of Fluid Dynamics*, Vol. 57, No. 17, San Diego, CA, Nov. 18-20, 2012

4

New Approach for Calculating Surface Characteristics

Giulia Galli, University of California, Davis



Science and Accomplishments

- First principle molecular dynamics simulations improving the accuracy with non-local van der Waals corrections and DFT
- Targeting direct comparison with experiments from the Spallation Neutron Source at ORNL
- Completed a series of simulations of the structure and IR spectra of liquid water and ice
- **The new technique has resulted in calculated vibrational spectra and estimates of the band gap that offer dramatically improved agreement with experiment, when compared to the lower level DFT approximations**

Key Impact

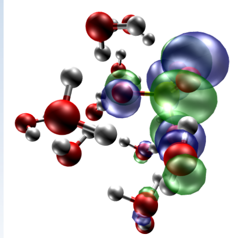
- Validated surface phenomena modeling capability to address effects that are difficult to probe experimentally

In 2012, this work appeared in 3 publications and was presented at 4 conferences.

Publication: "Dispersion interactions and vibrational effects in ice as a function of pressure: A first principles study," E. D. Murray and G. Galli, *Phys. Rev. Lett.*, 2012, 108, pp. 105502.

ALCF Contributions

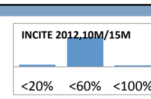
- Enabled improved throughput on Intrepid by writing scripts for the job management



Highest occupied molecular orbital of two configurations of a solvated $[SO_4]$—cluster in water, as obtained using *ab initio* calculations with the new hybrid Perdew-Burke-Ernzerhof functional PBE0.

5

Petascale Simulations of Inhomogeneous Alfvén Turbulence in the Solar Wind - *Jean C. Perez, University of New Hampshire*



Science and Accomplishments

- Alfvén waves (AW) are considered a key player in the origin of solar wind and the heating of the corona
- Resolving the AW role is difficult because there is little understanding of turbulence in the near-Sun region
- **Simulations here show evidence of a turbulent inertial range – a power-law spectrum typical of turbulent flows.**

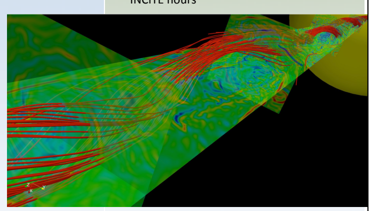
Presentation: J. C. Perez, *Turbulent Cascade in the Solar Wind: Anisotropy and Dissipation*, invited presentation at Centre International d'Ateliers Scientifiques, Observatoire de Paris, September 17-21, 2012.
Presented at the American Geophysical Union Fall Meeting, December 3-7, 2012.

Key Impact

- Progress towards solving two fundamental plasma physics problems concerning phenomena that drive solar magnetic disruptions on Earth

ALCF Contributions

- 2x overall improvement by improving inter-process communications
- 3.5x overall improvement (measured at smaller scale) with optimized FFT
- Improved run time allowed a much more complete physics study within the allocated INCITE hours



Snapshot of turbulent magnetic field lines (red) inside the simulated magnetic flux tube, representing a coronal hole that expands from a small patch on the solar surface to 5 solar radii. AW, launched by convective motions on the photosphere, propagate in the inhomogeneous solar atmosphere producing primary reflected waves that interact non-linearly with the outward waves, driving a turbulent cascade. This cascade continues with secondary reflections in a very complex interplay between the wave reflections and nonlinear interactions. Selected slices across the simulation domain show contours of plasma current, indicating the generation of small scale structures where the turbulent energy ultimately dissipates, heating the ambient plasma.

6

Stochastic Optimization of Complex Energy Systems
Mihai Anitescu, Argonne National Laboratory

INCITE 2012,10M/10M

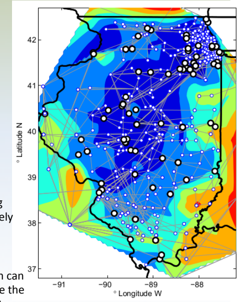
<20% <60% <100%

Science and Accomplishments
- Renewable energy sources supplies are uncertain and unmatched with demand
- Current systems mitigate this by using conservative and expensive reserve units
- Stochastic programming formulations of the decision process which schedules supply and matches demand are solved using PIPS
- New algorithms in PIPS enabled efficient simulations in "real-time"
- ***Demonstrated that even 20% wind penetration can be accommodated without significant reserve increase if using stochastic optimization***

Key Impact
- Addresses the problem of incorporating some renewable energy sources into the U.S. power grid without increasing reserve units or degrading grid performance

Publications: COAP and OR Letters (Accepted). SIAM Journal on Scientific Computing (In Review)

ALCF Contributions
- ALCF helped adopt and port required libraries and debug low level MPI challenges



The magnitude of the stochastic programming formulation for Illinois containing approximately 2,000 transmission nodes, 2,500 transmission lines, 900 demand nodes, and 300 generation nodes and needs to be considered over 24 successive hourly time periods. The simulation can reach billions of variables and constraints once the uncertainty in the supply is taken into account.

7

Toward Crystal Engineering from First Principles
James Chelikowsky, University of Texas at Austin

ALCC 2011,12M/17M

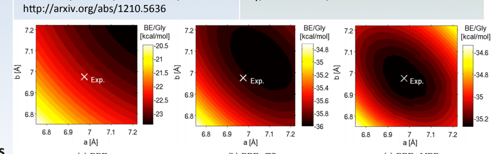
<20% <60% <100%

Science and Accomplishments
- Prove new methods for understanding the structure and properties of molecular crystals
- The many-body dispersion (MBD) method was applied to the prediction of the structure, energetics, and electronic properties of molecular crystals
- **The team has shown good agreement with their method proving a path forward for their approach**
- **The new method addresses two key deficiencies in previous methods**

Key Impact
- Improved computational capability for developing products such as antimalarial drugs, organic electronics, and more efficient solar cells

ALCF Contributions
- The ALCF wrote detailed workflow scripts for exploring the parameter spaces

Abstract: "Many-body dispersion interactions in molecular crystal polymorphism," N. Marom, R. A. DiStasio Jr., V. Atalla, S. Levchenko, J. R. Chelikowsky, L. Leiserowitz, A. Tkatchenko. http://arxiv.org/abs/1210.5636

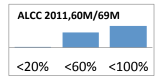


Potential energy surfaces for the a-b plane of Gamma-Gly using (a) their specific density functional theory (DFT) (using the Perdew-Burke-Ernzerhof (PBE) functional), (b) DFT plus an atom pairwise van der Waals correction (TS), and (c) DFT plus a many-body van der Waals correction (MBD). Only the last correctly explains the experimental lattice parameters (marked by a cross).

8

Understanding of "Crackle" in Supersonic Jet Noise

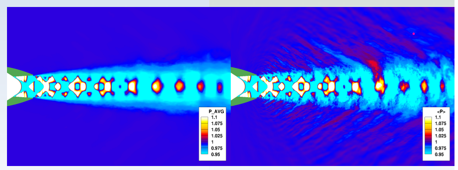
Parviz Moin, Stanford University



Science and Accomplishments	Key Impact	ALCF Contributions
• Crackle jet noise may account for up to 30% of the overall sound along the direction of peak noise • It occurs intermittently so is a hard phenomenon to capture • **The team has identified a region and pattern in the jet flow that is the source of the crackle noise** • **Application can identify the start of an event and backtrack the simulation for further study**	• Design new jet engine nozzles to manufacture quieter jet engines • Increase safety for those who work near jet engines	• The ALCF ported the code and worked on algorithmic improvements

Publication: Nichols, J.W., Lele, S.K., Moin, P., Ham, F.E., Bridges, J.E. (2012). "Large-eddy simulation for supersonic rectangular jet noise prediction: effects of chevrons." 18th AIAA/CEAS Aeroacoustics Conference, chapter DOI 10.2514/6.2012-2212

Report: Nichols, J.W., Lele, S.K. (2012) "Crackle in supersonic heated jets," Center for Turbulence Research Annual Research Briefs, Stanford University



(Left) shows contours of the LES mean pressure on an axial cross section. The nozzle material is shown in green. Both the sharp throat and the nozzle lip produce shocks that create a double-diamond shock cell pattern downstream. Far downstream, the double diamond cells merge into single regions where the pressure is statistically high. These single-celled shock cells then gradually reduce in intensity downstream. (Right) shows contours of pressure averaged over all of the source fields associated with 22 different crackle events.

9

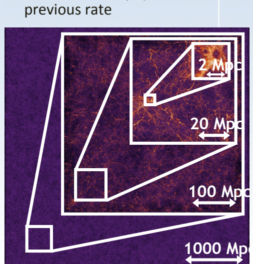
Probe the Cosmic Structure of the Dark Universe

Salman Habib, Argonne National Laboratory

Gordon Bell Finalist

ESP on Mira, 150M
Usage cannot be reported for this work. It was done on very early systems at ANL, LLNL, and IBM.

Accomplishments	Key Impact	ALCF Contributions
• **Largest cosmology simulation ever run** • 69% of peak on Mira and SC12 Gordon Bell Finalist • Achieved ~90% of peak bandwidth for I/O, 18x previous rate	• Time-to-solution improvements applicable to all science runs on Mira	• Designed, implemented, and tuned the tree algorithm needed for load balancing and data locality • Tuning of the key short and long range forces • Designed a new parallel I/O file format for visualization-specific outputs, and work on the *in-situ* analysis



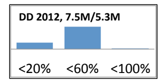
Shown is a zoom-in to small scales, the largest image shows the full simulation. These snapshots are taken when the Universe was 2.7 billion years old, 25% into the simulation time.

Publication: S. Habib, V. Morozov, H. Finkel, A. Pope, K. Heitmann, K. Kumaran, T. Peterka, J. Insley, D. Daniel, P. Fasel, N. Frontiere, and Z. Lukić. 2012, "The universe at extreme scale: multi-petaflop sky simulation on the BG/Q," In *Proc. of the Intl. Conf. on High Performance Computing, Networking, Storage and Analysis (SC '12)*, IEEE Computer Society Press, Los Alamitos, CA, USA, Article 4, 11 pages.

10

Petascale Simulation of Magnetorotational Core-Collapse Supernovae

Sean Couch, University of Chicago



Science and Accomplishments

- Incorporated approximate treatments for neutrino physics, realistic rotation, and magnetic fields into simulations
- **Revealed substantial differences in the behavior of 2-D and 3-D buoyant plumes**

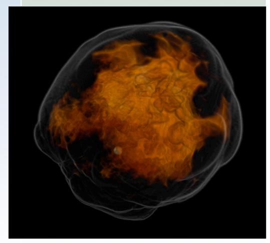
Publication: Couch, Sean M. "On the Impact of Three Dimensions in Simulations of Neutrino-Driven Core-Collapse Supernova Explosions," *Astrophysical Journal*, arXiv:1212.0010

Key Impact

- Proved 3-D simulations are required to model the mechanism of core-collapse supernovae

ALCF Contributions

- Early Science Program improvements improved time to solution by 4x through threading and single core improvements

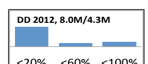


Volume rendering of the highest entropy plumes in a 3-D core-collapse supernova simulation. The edge of the nascent neutron star is shown as the faint blue sphere near the center. Entropy, a measure of the thermodynamic disorder, shows the buoyant convective plumes and turbulence that play a critical part in the core-collapse supernova mechanism. The surface of the outgoing shock wave is also shown in faint gray.

11

Discretionary Projects Advancing Highly Scalable Algorithms

Cyclops Tensor Framework – Sustained Petaflops for Quantum Chemistry
Jeff Hammond, Argonne National Laboratory

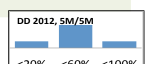


- Developed and implemented a new algorithm for coupled cluster calculations
- Capable of delivering sustained petaflop quantum chemistry on 1/3 of Mira

IPDPS, May 2013

Larger scale quantum chemistry than previous algorithms

Heuristic Static Load-Balancing Algorithm
Yuri Alexeev, Argonne National Laboratory

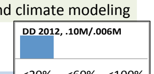


- Load balancing technique for uneven tasks or tasks less then number of nodes
- Improved scaling proven up to 160K cores of Intrepid

SC12

Improved method for load balancing already applicable to quantum chemistry and climate modeling

Expressing Graph Algorithms Using Generalized Active Messages
Nick Edmunds, Indiana University

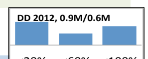


- Graph computation on HPC platforms has growing demand with few tools
- Data-driven nature of graph computation need a complex software stack

18th ACM SIGPLAN

A successful, scalable active message abstraction has been introduced

Scalable Data Particle Advection
Michael E. Papka, Argonne National Laboratory



- Particle advection is a powerful, difficult to scale, analysis tool
- Finite-Time Lyapunov Exponent is a powerful analysis but no previous particle tools scale

SC12

First method that can scale and handle the particle rate for key analysis technique

This page intentionally left blank.

Argonne Leadership Computing Facility
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 240
Argonne, IL 60439-4847

www.anl.gov

U.S. DEPARTMENT OF
ENERGY

Argonne National Laboratory
is a U.S. Department of Energy laboratory
managed by UChicago Argonne, LLC.

Contact ▸ Argonne Leadership Computing Facility
www.alcf.anl.gov | info@alcf.anl.gov | (877) 737-8615