ASCR ARRA Update
March 30, 2010

Walt Polansky and Dan Hitchcock
Advanced Scientific Computing Research
American Recovery and Reinvestment Act (Recovery Act)

**ASCAR’s Recovery Act Projects ($154.9M)**

- **Advanced Networking Initiative** ($66.8M)
  - Testbed to demonstrate and build tools for 100Gbps optical networking technologies

- **Leadership Computing Facility Upgrades** ($19.9M)
  - Six-core upgrade to Oak Ridge LCF machine delivered 2.2 Petaflops

- **Advanced Computer Architectures** ($5.2M)
  - Research on next generation technologies

- **Magellan** ($33M)
  - Research to demonstrate and build tools to enable scientists to utilize cloud computing resources for mid-range computing needs

- **SciDAC-e** ($30M)
  - Supplement and leverage existing SciDAC investments to advance the high performance computational capabilities of the BES - Energy Frontier Research Centers; Extra user support for Energy related projects at the Leadership Computing and NERSC facilities; Applied mathematics research in support of DOE electricity grid efforts.
DOE Explores Cloud Computing

- **ASCR Magellan Project Summary**
  - $32M project at NERSC and ALCF
  - ~100 TF/s compute cloud testbed (across sites)
  - Petabyte-scale storage cloud testbed

- **Project Progress**
  - Funding distributed to ANL and LBNL based on peer reviewed proposal
  - ANL and LBNL procured and installed compute and first stage of data hardware
  - Identified experts to review integrated research demonstration topics (Q1)
  - Joint Magellan-ANI PI meeting was conducted at SC09
  - Coordination with ANI on-going
  - First cycles now available

- **Cloud questions to explore on Magellan:**
  - Can a cloud serve DOE’s mid-range computing needs?
    - More efficient than cluster-per-PI model
  - What part of the workload can be served on a cloud?
  - What features (hardware and software) are needed of a “Science Cloud”?
    - (Eucalyptus at ALCF; Linux at NERSC)
  - How does this differ, if at all, from commercial clouds?
NERSC Magellan Cloud Hardware

720 nodes, 5760 cores in 9 Scalable Units (SUs) ➞ 61.9 Teraflops
SU = IBM iDataplex rack with 640 Intel Nehalem cores

18 Login/network nodes
10G Ethernet

Load Balancer

HPSS (15PB)

100-G Router

QDR IB Fabric

14 I/O nodes (shared)
8G FC

NERSC Global Filesystem

1 Petabyte with GPFS

Internet

ASCAC March 30, 2010
Compute
504 Compute Nodes
Nehalem Dual quad-core 2.66GHz
24GB RAM, 500GB Disk
QDR IB link
Totals
4032 Cores, 40TF Peak
12TB RAM, 250TB Disk

Active Storage
~100 Compute/Storage Nodes
~10TB FLASH/SSD Storage
~500TB Disk Storage
Key is flexible and dynamic scheduling of resources

- Runtime provisioning of software images
- Rolling upgrades can improve availability
- Ability to schedule to local or remote cloud for most cost effective cycles
The goal of the ANI project is two-fold:

- Accelerate the commercialization of 100 Gigabit per second (Gbps) networking technologies by deploying a national-scale prototype network that will span four distinct geographic regions, connecting the three major ASCR computing facilities and the New York multi-agency peering point providing transatlantic Research and Education (R&E) connectivity at 100 Gbps.

- Complement the prototype 100 Gbps network with a testbed providing an experimental network research environment at sufficient scale to usefully test experimental approaches to next generation networks and applications.
ANI: Status

- Funds sent to all contractors and grant recipients
- Program review of preliminary design for ChiExpress and design posted on ASCR website: [http://www.sc.doe.gov/ascr/Misc/ASCRRecovery.html](http://www.sc.doe.gov/ascr/Misc/ASCRRecovery.html)
- Project Plan (design document) posted on ASCR website:
  - [http://www.sc.doe.gov/ascr/Misc/ASCRRecovery.html](http://www.sc.doe.gov/ascr/Misc/ASCRRecovery.html)
- Testbed Timeline
  - Feb 2010: Initial ‘table top’ node hardware setup at LBNL
  - Sept 2010: 10 Gbps ‘table top’ testbed available to researchers
  - April 2011: 10/100 Gbps WAN testbed available to researchers
  - Jan 2012: full 100 Gbps WAN testbed available to researchers.
- Upcoming
  - Solicitation for second round of research topics for use of test bed (Q3-Q4).
Advanced Network Initiative Topology
ANI: 100 Gbps Prototype Network
Preliminary Baseline Design

Key:
- Brown: Carrier Layer 1 device
- Blue: ESnet co-managed Layer 1 device
- Green: Site Router
- Red: ANI Router
- Light Blue: Exchange Point Router
- Orange: ESnet co-managed 100G
- Black: Site managed fiber

Assumptions:
- co-managed IRU dark fiber for the following paths:
  - SUNN to NERSC
  - Chicago to ANL
  - AOFA to BNL
- Carriers will provide 100G ethernet circuit

This network provides:
- L3 (routed) service between sites and 100G ANI routers
- L2 (VC) service between sites
- Backbone VCs are hybrid:
  - L2.5 on carrier Link (MPLS)
  - L1 on dark fiber (GMPLS)
Leadership Computing Upgrade
ASCR Deploys World’s Most Powerful Computer for Open Science at ORNL

- ASCR reviewed LCF’s upgrade implementation plan to ensure upgrade activities result in less than 10% unscheduled downtime for current users. (2009)
- ASCR reviewed and approved acceptance test plan including applications to be used in acceptance test. (2009)
- OLCF Completed acceptance test for quad-core to six-core upgrade of Cray XT5 at Oak Ridge (Q1).
  - ASCR reviewed acceptance test results and approved start of operations.
OLCF worked with users to produce scalable, high-performance apps for the petascale.

<table>
<thead>
<tr>
<th>Science Area</th>
<th>Code</th>
<th>Contact</th>
<th>Cores</th>
<th>Total Performance</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Materials</td>
<td>DCA++</td>
<td>Schulthess</td>
<td>213,120</td>
<td>1.9 PF*</td>
<td>2008 Gordon Bell Winner</td>
</tr>
<tr>
<td>Materials</td>
<td>WL-LSMS</td>
<td>Eisenbach</td>
<td>223,232</td>
<td>1.8 PF</td>
<td>2009 Gordon Bell Winner</td>
</tr>
<tr>
<td>Chemistry</td>
<td>NWChem</td>
<td>Apra</td>
<td>224,196</td>
<td>1.4 PF</td>
<td>2009 Gordon Bell Finalist</td>
</tr>
<tr>
<td>Nano Materials</td>
<td>OMEN</td>
<td>Klimeck</td>
<td>222,720</td>
<td>860 TF</td>
<td></td>
</tr>
<tr>
<td>Seismology</td>
<td>SPECFEM3D</td>
<td>Carrington</td>
<td>149,784</td>
<td>165 TF</td>
<td>2008 Gordon Bell Finalist</td>
</tr>
<tr>
<td>Weather</td>
<td>WRF</td>
<td>Michalakes</td>
<td>150,000</td>
<td>50 TF</td>
<td></td>
</tr>
<tr>
<td>Combustion</td>
<td>S3D</td>
<td>Chen</td>
<td>144,000</td>
<td>83 TF</td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>GTC</td>
<td>PPPL</td>
<td>102,000</td>
<td>20 billion Particles / sec</td>
<td>2008 Gordon Bell Winner</td>
</tr>
<tr>
<td>Chemistry</td>
<td>MADNESS</td>
<td>Harrison</td>
<td>140,000</td>
<td>550+ TF</td>
<td></td>
</tr>
</tbody>
</table>
Scientific Progress Resulting from OLCF Upgrade

**Turbulence**
Understanding the statistical geometry of turbulent dispersion of pollutants in the environment.

**Nuclear Energy**
High-fidelity predictive simulation tools for the design of next-generation nuclear reactors to safely increase operating margins.

**Energy Storage**
Understanding the storage and flow of energy in next-generation nanostructured carbon tube supercapacitors

**Fusion Energy**
Substantial progress in the understanding of anomalous electron energy loss in the National Spherical Torus Experiment (NSTX).

**Biofuels**
A comprehensive simulation model of lignocellulosic biomass to understand the bottleneck to sustainable and economical ethanol production.

**Nano Science**
Understanding the atomic and electronic properties of nanostructures in next-generation photovoltaic solar cell materials.
Advanced Architectures

What it is:
- New effort to provide early access to DOE researchers of emerging.
- Enhancement of University of California, Berkeley RAMP effort to provide focused research on flexible simulations of performance of scientific applications on next generation microprocessors.
- Both proposals were in hand and were peer reviewed.

Goal:
- By September 30, 2010, complete initial definition of architectural features and performance levels for a system that will meet the needs of at least one science application that requires extreme scale computing while using energy efficiency.

Progress:
- Funding in place at ORNL for IBM PERCS effort
  - Negotiations underway for prototype testbed
- Grant for RAMP Awarded March 19, 2010

Review:
- Established the Charge for Expert Panel peer review with definitions of ratings (Q1).
- Identify Expert Panel Chair and reviewers (Q3).
SciDAC-e
Applied Mathematics for grids

- Applied mathematics research in support of DOE electricity grid efforts.
  - Robust Optimization for Connectivity and Flows in Dynamic Complex Networks, Lead PI: Balasundaram (Oklahoma State)
  - Reconfiguring Power Systems to Minimize Cascading Failures: Models and Algorithms, Co-PIs: Bienstock (Columbia), Wright (UW-Madison)
  - Approaches for Rare-event Simulation and Decision Making, Lead PI: Shortle (GMU)
  - Analysis and Reduction of Complex Networks under Uncertainty, Marzouk (MIT), Knio (JHU), Ghanem (USC), Najm (SNL)
  - Optimization and Control of the Electric Power Systems, Co-PIs: Meza (LBNL), Thomas (Cornell), Lesieutre (UW-Madison)
  - Advanced Kalman Filter for Real-Time Responsiveness in Complex Systems, Co-PIs: Huang (PNNL), Welch (UNC-Chapel Hill)
  - Extending the Realm of Optimization for Complex Systems: Uncertainty, Competition and Dynamics, Lead PI: Shanbhag (UIUC)

- All awards made and work has begun.
- Plan to conduct programmatic and expert review of progress and results.
SciDAC-e
Computational Postdoctoral Fellows

Progress (10/28/2010)

NERSC (Goal: 8 total)
- 5 hired/accepted, 3 of those started work
- 3 new candidates contacted and in pipeline to be interviewed

ALCF (Goal: 10-11 total)
- 5 hired (3 have already started work)
- 1 in process
- 4 candidates identified for interviews

OLCF (Goal: 10 total)
- 2 hired/accepted
- 4 in process
- 4 candidates interviewed/to be interviewed
- Looking for additional candidates

Overview

- Post Doctoral Fellows to provide extra user support for Extra Energy related projects at the Leadership Computing and NERSC facilities
- Funds became available at the end of FY09

ASCAC March 30, 2010