Scientific Discovery through Advanced Computing Program: SciDAC Update

Steven L. Lee
DOE Program Manager, Office of Science
Advanced Scientific Computing Research

August 14, 2012
Specific goals and objectives for the SciDAC Institutes:

- Tools and resources for lowering the barriers to effectively use state-of-the-art computational systems;
- Procedures for taking on computational grand challenges across different science application areas;
- Procedures for incorporating and demonstrating the value of basic research results from Applied Mathematics and Computer Science; and
- Plans for building up and engaging our nation’s computational science research communities.

FY11 Program Funding – Office of Advanced Scientific Computing Research (ASCR)

- Up to $13M/year for 5 years may be available to support between 1 and 5 SciDAC Institutes
- DOE National Laboratories, Universities, Industry and other organizations may apply

Timeline

- Issued - February 23, 2011
- Letters of Intent (LOI), though not required, are strongly encouraged - March 30, 2011
- Proposal due date – May 2, 2011
- FY11 Awards for 3 SciDAC Institutes completed – July 2011
  - New SciDAC Institutes solicitation for Scientific Data Management, Analysis and Visualization
  - FY12 - Posted Sep 16; LOIs due Oct 12; Proposals due Nov 9; Awarded in Feb 2012
The four SciDAC Institutes are large team projects involving National Laboratory, University and Industry collaborators.

<table>
<thead>
<tr>
<th>FASTMath</th>
<th>QUEST</th>
<th>SDAV</th>
<th>SUPER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Director – Lori Diachin</td>
<td>Director – Habib Najm</td>
<td>Director – Arie Shoshani</td>
<td>Director – Robert Lucas</td>
</tr>
<tr>
<td>Scalable solvers &amp; discretizations</td>
<td>Uncertainty Quantification</td>
<td>Scalable data management, analysis &amp; visualization</td>
<td>Performance tools &amp; code optimization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lawrence Livermore (CA)</th>
<th>Sandia (CA)</th>
<th>Lawrence Berkeley (CA)</th>
<th>Univ of Southern CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argonne (IL)</td>
<td>Los Alamos (NM)</td>
<td>Argonne (IL)</td>
<td>Argonne (IL)</td>
</tr>
<tr>
<td>Lawrence Berkeley (CA)</td>
<td>Duke University (NC)</td>
<td>Lawrence Livermore (CA)</td>
<td>Lawrence Berkeley (CA)</td>
</tr>
<tr>
<td>Sandia (CA &amp; NM)</td>
<td>MIT (MA)</td>
<td>Los Alamos (NM)</td>
<td>Lawrence Livermore (CA)</td>
</tr>
<tr>
<td>RPI (NY)</td>
<td>Univ of Southern CA</td>
<td>Oak Ridge (TN)</td>
<td>Oak Ridge (TN)</td>
</tr>
<tr>
<td>Univ of Texas, Austin (TX)</td>
<td>Sandia (NM)</td>
<td>Univ of CA, San Diego (CA)</td>
<td></td>
</tr>
<tr>
<td>Univ of CA, Davis (CA)</td>
<td>Georgia Tech (GA)</td>
<td>Univ of Maryland (MD)</td>
<td></td>
</tr>
<tr>
<td>Georgia Tech (GA)</td>
<td>North Carolina St Univ (NC)</td>
<td>Univ of North Carolina (NC)</td>
<td></td>
</tr>
<tr>
<td>North Carolina St Univ (NC)</td>
<td>Northwestern (IL)</td>
<td>Univ of Oregon (OR)</td>
<td></td>
</tr>
<tr>
<td>Northwestern (IL)</td>
<td>Ohio State Univ (OH)</td>
<td>Univ of Tenn, Knoxville (TN)</td>
<td></td>
</tr>
<tr>
<td>Ohio State Univ (OH)</td>
<td>Rutgers Univ (NJ)</td>
<td>Univ of Utah (UT)</td>
<td></td>
</tr>
<tr>
<td>Rutgers Univ (NJ)</td>
<td>Univ of Utah (UT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitware, Inc (NY)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FASTMath** - Frameworks, Algorithms & Scalable Technologies for Mathematics  
**QUEST** - Quantification of Uncertainty in Extreme-Scale Computations  
**SDAV** - Scalable Data Management, Analysis & Visualization  
**SUPER** - Institute for Sustained Performance, Energy & Resilience
Scientific computing software must address ever increasing challenges:

- Million to billion way parallelism
- Deeply hierarchical NUMA for multi-core processors
- Fault tolerance
- Data movement constraints
- Heterogeneous, accelerated architectures
- Power constraints

Modeling and simulation is significantly complicated by the change in computing architectures.
FASTMath helps application scientists overcome two fundamental challenges
Director: Lori Diachin at Lawrence Livermore National Laboratory

1. Improve the quality of their simulations
   - Increase accuracy
   - Increase physical fidelity
   - Improve robustness and reliability

2. Adapt computations to make effective use of Leadership Computing Facilities
   - Million way parallelism
   - Multi-/many-core nodes

*FASTMath addresses both challenges by focusing on the interactions among mathematical algorithms, software design, and computer architectures*
FASTMath encompasses three broad topical areas:

**Tools for Problem Discretization**
- Structured grid technologies
- Unstructured grid technologies
- Adaptive mesh refinement
- Complex geometry
- High-order discretizations
- Particle methods
- Time integration

**Solution of Algebraic Systems**
- Iterative solution of linear systems
- Direct solution of linear systems
- Nonlinear systems
- Eigensystems
- Differential variational inequalities

**High Level Integrated Capabilities**
- Adaptivity through the software stack
- Management of field data
- Coupling difference physics domains
- Mesh/particle coupling methods
All FASTMath technologies will focus on performance engineering for multi-/many-core architectures

**MPI Level Parallelism**
- Operate efficiently at $10^5$ to $10^6$ cores
- Architecture-aware and multi-objective load balancing
- Communication avoiding and latency tolerant algorithms
- Synchronization reducing algorithms

**Node Level Parallelism**
- Use of threading techniques
- Multi-core kernels and data ordering
- Exploit compilers, code transformation tools, programming models and run-time systems as they become available

**Data Locality**
- Hierarchical partitioning and local data ordering methods
- Shared efficient data layouts in software packages to prevent re-organization
- Code transformation systems, domain specific language extensions to gain performance while maintaining reusability

Coordinated parallelism between different levels (MPI, node, instruction)

**FASTMath Program Manager:** S. Lee

See [http://www.fastmath-scidac.org](http://www.fastmath-scidac.org)
**SDAV Goals**

- To actively work with application teams to assist them in achieving breakthrough science
- To provide technical solutions in the data management, analysis, and visualization regimes that are broadly used by the computational science community running on Leadership Class machines
- To use existing robust tools to the extent possible and develop/adapt tools on as-as-needed basis

SDAV tools have been developed over many years – robust, well-documented. Tools are being enhanced in several ways:

- Scale tools for high-parallelization levels
- Adapt tools to take advantage of new hybrid hardware (CPUs + GPUs)
- Minimize data movement between nodes
- Adapt tools for in-situ processing and analysis
- Compress and index data for both in-situ and post-processing analysis
SDAV Goals

- To actively work with application teams to assist them in achieving breakthrough science
- To provide technical solutions in the data management, analysis, and visualization regimes that are broadly used by the computational science community running on Leadership Class machines
- To use existing robust tools to the extent possible and develop/adapt tools on an as-needed basis

Data Management

- In-Situ Processing & Code Coupling
  - ADIOS
  - Glean
- Indexing
  - FastBit
- In-Situ Data Compression
  - ISABELLA
- Parallel I/O & File Formats
  - PnetCDF
  - BP-files
  - HDF5

Data Analysis

- Statistical & Data Mining
  - NU-Minebench
- Importance-Driven Analysis
- Domain-knowledge directed
- Geometry-based
- Topological Methods
  - In-Situ Topology
  - Feature-based analysis
  - High-frequency analysis & tracking

Visualization

- Parallel tools
  - VisIt
  - ParaView
- VTK-m framework
- Flow Visualization methods
- Rendering
- Ensembles, Uncertainty & Higher-Dimensional methods

See [http://sdav-scidac.org](http://sdav-scidac.org)

SDAV Program Manager: Lucy Nowell
**QUEST objectives**

- Deliver expertise, advice and state of the art UQ tools on advanced computational architectures
- Shepherd forward QUEST repertoire of UQ theory, algorithms, and software, and enhancing their effectiveness for relevant benchmark problems

**QUEST vision**

1. Well-founded setup of the UQ problem
2. Characterization of the input space given available data
3. Local and global sensitivity analysis
4. Adaptive dimensionality and order reduction
5. Forward and inverse propagation of uncertainty
6. Handling of application code failures, missing data & fault tolerance
7. Model comparison, validation, selection, and averaging

QUEST tools include: DAKOTA, UQTk, QUESO, GPMSA

QUEST Program Manager: S. Lee

See [http://quest-scidac.org](http://quest-scidac.org)
SUPER Goal
Ensure DOE’s computational scientists can successfully exploit the emerging generation of high performance computing (HPC) systems.

Research Activities:
• Performance Portability – Extend performance measurement, modeling and auto-tuning technology to petascale & heterogeneous computing systems
• Energy Efficiency – Investigate application-level energy efficiency techniques
• Resilience – Explore strategies to enable application resilience against faults
• Optimization – Develop strategies to collectively optimize performance, energy efficiency, and resilience

Collaboration:
• Application Engagement – Work on science applications for tool development
• Tool Integration – Create end-to-end, integrated performance tool suite
• Outreach – Web-based & hand-on tutorials for science community impact

SUPER Program Manager: Ceren Susut
See http://super-scidac.org
**Grand Challenge Goal:** Leverage U.S. leadership in advanced computing, modeling & simulation to deploy affordable, user-friendly, accessible platforms for broad use across America’s energy sector

**Educational Colloquiums** – Tuesday, July 31

**Panels** - Wednesday AM, August 1
- DOE Assistant Secretaries Panel on DOE Applied Technology Programs (EERE, NE, NRAP, OE)
- Energy Innovation: Success Stories
- Energy Innovation: Potential and Challenges

**Q&A session with Secretary Chu** – Wednesday Lunch, August 1

**Breakout Sessions** – Wednesday PM through August 2
- DOE Applied Programs
- Current Users – Grand Challenges
- Potential Users – Grand Challenges

**Outcome:** Workshop Report due by November 1
Partnerships for Science

➤ SciDAC Application Partnerships with SC Offices (Randall Laviolette, Ceren Susut)
  • DOE Applied Offices
  • Industry

SciDAC-3 Principal Investigator Meeting
  • Focused on PIs, Institutes and projects funded in this third round of SciDAC
  • September 10 – 12 in Rockville, MD
  • Conference approval is pending

New SciDAC-3 Program is currently comprised of SciDAC Institutes and Application Partnerships over the next 5 years: 2011 – 2016.

• “The overall portfolio & management of Institute awards is expected to cover a significant portion of DOE computational science needs on current and emerging computational systems”
• Basic research programs prepare the way for SciDAC-4 & Extreme-Scale Institutes