The Launch of Scientific Discovery through Advanced Computing

Presentation to Advanced Scientific Computing Advisory Committee

Washington, DC
October 25-26, 2001

http://www.science.doe.gov/SciDAC/

Stephen Eckstrand
Former Director, SciDAC
SciDAC Goals

An **integrated** program to:

(1) Create a new generation of Scientific Simulation Codes that take full advantage of the extraordinary computing capabilities of terascale computers.

(2) Create the Mathematical and Computing Systems Software to enable these Scientific Simulation Codes to use terascale computers effectively and efficiently.

(3) Create a Collaboratory Software Environment to enable geographically separated scientists to work together as an integrated team and to facilitate remote access to both facilities and data.
Just within the past few years, advanced scientific computing has reached a point where it is on a par with laboratory experiment and mathematical theory as a major tool for scientific discovery.

Major increases in computing power will enable a new generation of simulation codes, based on reliable experimental and theoretical inputs, to lead the way to greatly increased scientific understanding, which in turn will lead to new theoretical and experimental discoveries.
SciDAC-Integrating Science and Computing

Program Management Strategies

- Development of next generation modeling and simulation codes, mathematical algorithms, systems software, along with computing and network resource planning are interdependent activities that must be carefully managed.
  - Coordinate planning and management of SciDAC activities across the Office of Science (SciDAC director and a SciDAC Coordinating Committee)
  - Establish integrated scientific challenge teams (applications scientists, applied mathematicians, and computer scientists) with clear deliverables
  - Provide reliable access to adequate supercomputing and network resources
SciDAC Coordinating Committee

- Dave Bader, BER
- Bill Kirchhoff, BES
- Walter Stevens, BES
- Steve Eckstrand, FES
- Arnold Kritz, FES
- Robin Staffin, HENP
- Vicky White, HENP

- Fred Johnson, ASCR
- Buff Miner, ASCR
- Thomas Ndousse, ASCR
- Kimberly Rasar, ASCR
- Chuck Romine, ASCR
- Mary Anne Scott, ASCR
- George Seweryniak, ASCR
SciDAC Program History

- Prepare Budget Requests 1999-2000
- Prepare Solicitations (6) 12/2000
- Conduct Merit Reviews 4-5/2001
- Make Funding Recommendations 6/2001
- Prepare Grant and Contract Awards 6-7/2001
- Manage Start of Individual Projects 7-9/2001
- Prepare for PI Meeting 10-12/2001
Initial Awards Focus on **Software**

- **Scientific Applications**
  - Climate Simulation
  - Computational Chemistry
  - Fusion – 5 topics
  - High Energy/Nuclear Physics – 5 topics

- **Collaboratories**
  - Four projects

- **Middleware & Network Research**
  - Six projects

- **Computer Science**
  - Scalable Systems Software
  - Common Component Architecture
  - Performance Science and Engineering
  - Scientific Data Management

- **Applied Mathematics**
  - PDE Linear/Nonlinear Solvers and Libraries
  - Structured Grids/AMR
  - Unstructured Grids
Initial Awards Support Teams

SciDAC-Integrating Science and Computing

Twenty-four projects funded at $500K-$3M per year
First Year Funding Distribution

SciDAC Funding ($M)

- MICS: $37
- BER: $8
- BES: $3
- HENP: $7
- FES: $2

Legend:
- MICS
- BER
- BES
- HENP
- FES
SciDAC - Integrating Science and Computing

Distribution of Funding between National Labs and Universities

SciDAC - MICS Awards: $37M

SciDAC - Total Awards: $57M

53%

47%

55%

45%

☑ Univ. & Other Res. Inst.
☑ Lab.
Thirteen National Labs Involved

- Ames Laboratory
- Argonne National Laboratory
- Brookhaven National Laboratory
- Fermi National Accelerator Laboratory
- Lawrence Berkeley National Laboratory
- Lawrence Livermore National Laboratory
- Los Alamos National Laboratory
- Oak Ridge National Laboratory
- Pacific Northwest National Laboratory
- Princeton Plasma Physics Laboratory
- Thomas Jefferson Laboratory
- Sandia National Laboratory
- Stanford Linear Accelerator
Fifty-three Universities in 26 States and Province of Quebec Involved

SciDAC-Integrating Science and Computing

Auburn University; Boston University; California Institute of Technology; Carnegie Mellon University; Clemson University; Colorado State University; Florida Atlantic University; Georgia Institute of Technology; Indiana University; Iowa State University; Massachusetts Institute of Technology; Michigan State University; New York University; North Carolina State University; Northwestern University; Ohio State University; Oklahoma State University; Old Dominion University; Princeton University; Rensselaer Polytechnic Institute; Rice University; Rollins College; Rutgers University; Scripps Institute (UCSD); Stanford University; State University New York at Stony Brook; Stevens Institute of Technology; University Corporation for Atmospheric Research; University of Arizona; University of California Berkeley; University of California Davis; University of California Los Angeles; University of California San Diego; University of California Santa Barbara; University of California Santa Cruz; University of Chicago; University of Colorado; University of Delaware; University of Georgia; University of Illinois at Urbana-Champaign; University of Iowa; University of Maryland; University of Michigan; University of North Carolina; University of Quebec (Canada); University of Southern California; University of Tennessee; University of Texas at Austin; University of Utah; University of Washington; University of Wisconsin - Madison; Utah State University; and Wellesley College
# Computational Chemistry Awards

SciDAC-Integrating Science and Computing

Two Projects on modeling of combustion, one project on heavy element chemistry, and eight projects on chemical structures and reactivity - $1.9M

<table>
<thead>
<tr>
<th>Project Description</th>
<th>Institution</th>
<th>Principal Investigators</th>
<th>Funding Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terascale High-Fidelity Simulations of Turbulent Combustion with Detailed Chemistry</td>
<td>Sandia National Laboratory (Habib Najm, SNL)</td>
<td></td>
<td>$300K/Yr</td>
</tr>
<tr>
<td>Reliable Electronic Structure Calculations for Heavy Element Chemistry: Molecules Containing Actinides, Lanthanides, and Transition Metals</td>
<td>Arnaud Trouve, Univ. of Maryland Hong Geun Im, Univ. of Michigan Christopher J. Rutland, Univ. of Wisconsin-Madison</td>
<td></td>
<td>$227K/Yr</td>
</tr>
<tr>
<td>Reliable Electronic Structure Calculations for Heavy Element Chemistry: Molecules Containing Actinides, Lanthanides, and Transition Metals</td>
<td>Walter Ermler, Stevens Institute Russ Pitzer, Ohio State Univ.</td>
<td></td>
<td>$198K/Yr</td>
</tr>
<tr>
<td>Research Area</td>
<td>Investigator(s)</td>
<td>Funding (Yr)</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Advanced Methods for Electronic Structure</td>
<td>Robert Harrison, PNNL</td>
<td>$375K/Yr</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Martin Head-Gordon, LBNL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advancing Multi-Reference Methods in Electronic Structure Theory</td>
<td>Mark S. Gordon, Ames Laboratory</td>
<td>$250K/Yr</td>
<td></td>
</tr>
<tr>
<td>New Coupled-Cluster Methods for Molecular Potential Energy Surfaces</td>
<td>Piotr Piecuch, Michigan State</td>
<td>$110K/Yr</td>
<td></td>
</tr>
<tr>
<td>Accurate Properties for Open-Shell States of Large Molecules</td>
<td>Peter Taylor, UCSD</td>
<td>$121K/Yr</td>
<td></td>
</tr>
<tr>
<td>Explicitly Correlated Coupled Cluster &amp; Bruecker Methods for Computations of</td>
<td>Fritz Schaefer, Univ. of Georgia</td>
<td>$66K/Yr</td>
<td></td>
</tr>
<tr>
<td>Properties of Chemical Accuracy for Open Shell Systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear Scaling electronic Structure Methods with Periodic Boundary</td>
<td>Gustavo E. Scuseria, Rice Univ.</td>
<td>$100K/Yr</td>
<td></td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Software for the Calculation of Thermochemistry, Kinetics and</td>
<td>Albert F. Wagner, Argonne National Laboratory</td>
<td>$85K/Yr</td>
<td></td>
</tr>
<tr>
<td>Dynamics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical Chemical Dynamics Studies of Elementary Combustion Reactions</td>
<td>Donald L. Thompson, Oklahoma State Univ.</td>
<td>$95K/Yr</td>
<td></td>
</tr>
</tbody>
</table>
Combustion

Models for predicting combustion processes including full chemistry and the interaction of chemistry with fluid dynamics

Laser induced fluorescence imaging of the interaction of chemistry with fluid dynamics
Profiles of species CH and OH with time

Participants …

Sandia National Laboratories
University of Maryland
University of Michigan
University of Wisconsin

3 millisec after vortex reaches flame surface
5 millisec after vortex reaches flame surface
Predicting Molecular Structure and Reactivity

SciDAC-Integrating Science and Computing

Chemically accurate calculations of the properties of molecules with an emphasis on excited states and reactive radicals important in combustion

Participants …

Pacific Northwest National Laboratory
Lawrence Berkeley National Laboratory
Argonne National Laboratory
Ames Laboratory

Michigan State University
University of California, San Diego
University of Georgia
Rice University
Oklahoma State University
Calculation of properties of lanthanide and actinide compounds for which experiments are impractical and to aid in the interpretation of experiments

Participants …

Pacific Northwest National Laboratory

Stevens Institute of Technology
Ohio State University

Uranium Carbonate Complex
\((\text{UO}_2)_3(\text{CO}_3)_6\)
Fusion Energy Sciences Awards

SciDAC-Integrating Science and Computing

Five Projects to develop key elements of an integrated model for the simulation of fusion plasma systems - $3.0M

<table>
<thead>
<tr>
<th>Project Description</th>
<th>Institutions</th>
<th>Funding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnetic Reconnection: Applications to Sawtooth Oscillations, Error Field Induced Islands &amp; the Dynamo Effect</td>
<td>Univ. Of Iowa, Univ. of Chicago, Univ. of Texas (Amitava Bhattacharjee, Univ. of Iowa)</td>
<td>$750K/Yr</td>
</tr>
<tr>
<td>Terascale Computational Atomic Physics for the Edge Region in Controlled Fusion Plasmas</td>
<td>Auburn Univ., Rollins Coll., ORNL (Mitch Pindzola, Auburn)</td>
<td>$300K/Yr</td>
</tr>
<tr>
<td>Numerical Computation of Wave-Plasma Interactions in Multi-dimensional Systems</td>
<td>ORNL, PPPL, MIT, Lodestar CompX (Don Batchelor, ORNL)</td>
<td>$1.0M/Yr</td>
</tr>
<tr>
<td>Center for Extended MHD Modeling</td>
<td>PPPL, SAIC, MIT, Utah State Univ., Univ. of Wisconsin, Univ. of Colorado, NYU (Steve Jardin, PPPL)</td>
<td>$500K/Yr</td>
</tr>
<tr>
<td>Plasma Microturbulence Project</td>
<td>LLNL, PPPL, UCLA, Univ. of Maryland, General Atomics, Univ. of Colorado (Bill Nevins, LLNL)</td>
<td>$450K/Yr</td>
</tr>
</tbody>
</table>
SciDAC-Integrating Science and Computing

Terascale Atomic Physics
Auburn University
Rollins College
Oak Ridge National Laboratory

Substorm in the Magnetotail
(a) time=0
Current J,
0.1707
0.1448
0.1189
0.0930
0.0671
0.0412
0.0153
(b) time=100

Electron Probability Distribution during Atom-Molecule Collision

Magnetic Reconnection Code
University of Iowa
University of Chicago
University of Texas
Fusion Energy Sciences Projects

SciDAC-Integrating Science and Computing

Computation of Wave Plasma Interactions

ORNL, PPPL, MIT, Lodestar, CompX

Extended MHD Modeling

PPPL, SAIC, U. Wisconsin, NYU, U. Colorado, MIT, Utah State U., GA, LANL, U. Texas

M=1, N=1 Plasma Instability

Plasma Microturbulence


Turbulent Eddies in Plasma
High Energy/Nuclear Physics Awards

Four Scientific Simulation project awards were made and one Collaborative Pilot award (jointly with ASCR) for a total of $6.9M of HENP funding

<table>
<thead>
<tr>
<th>National Computational Infrastructure for Lattice Gauge Theory</th>
<th>UCSB</th>
<th>BU</th>
<th>MIT</th>
<th>Columbia</th>
<th>FNAL</th>
<th>BNL</th>
<th>TJNAF</th>
<th>U. Arizona</th>
<th>U. Utah</th>
<th>UIUC</th>
<th>$1.86M/Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Computing for 21st Century Accelerator Science and Technology</td>
<td>LBNL</td>
<td>SLAC</td>
<td>FNAL</td>
<td>SNL</td>
<td>UCLA</td>
<td>Stanford</td>
<td>U. Maryland</td>
<td>U. Southern California</td>
<td>UC Davis</td>
<td>Tech-X Corp.</td>
<td>$1.8 M/Yr</td>
</tr>
<tr>
<td>SciDAC Center of Supernova Research</td>
<td>UC Santa Cruz</td>
<td>U. Arizona</td>
<td>LANL</td>
<td>LLNL</td>
<td>$0.52M/Yr</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Particle Physics Data Grid Collaborative Pilot</td>
<td>Univ. of Wisconsin</td>
<td>UCSD</td>
<td>Caltech</td>
<td>BNL</td>
<td>FNAL</td>
<td>ANL</td>
<td>JLAB</td>
<td>LBNL</td>
<td>SLAC</td>
<td>$1.44M/Yr (Joint with ASCR)</td>
<td></td>
</tr>
</tbody>
</table>
High Energy and Nuclear Physics Projects

SciDAC-Integrating Science and Computing

National Computational Infrastructure For Lattice Gauge Theory (Sugar)

Representing ~60 theorists in the US. Funding to 3 labs and 6 universities

Analysis of calculation on Lattice

Quark propagator
Simulation of intense beam exhibits ultra-low density beam halo

Simulation provides high resolution needed to predict peak power loss in complex electromagnetic cavity
High Energy and Nuclear Physics Projects

SciDAC-Integrating Science and Computing

Shedding New Light on Exploding Stars: Terascale Simulations of Neutrino-Driven Supernovae and their Nucleosynthesis (Mezzacappa)

Funding to 1 lab and 8 universities

Problem: How to reveal local features occurring in a larger structure? Build customized RenderMan *data shaders.*

Instrumented 3D hydrodynamics code monitored using CUMULUS
SciDAC-Integrating Science and Computing

SciDAC Center for Supernova Research (Woosley)

Funding to 2 labs and 2 universities

Impact of a supernova on an adjacent star

Protoneutron star convection
High Energy and Nuclear Physics Projects

SciDAC-Integrating Science and Computing

The Particle Physics Datagrid Collaboratory Pilot
(Mount, Newman, Livny)

Jointly funded with ASCR

Funding to 6 labs and 3 universities
Involves major HEP and NP experiments
Babar, DO, CMS, ATLAS, STAR
Climate Modeling Awards

SciDAC-Integrating Science and Computing

Two large projects to develop a community climate model and the climate model of the future; 13 smaller projects to develop advanced climate model elements - $7.9M

<table>
<thead>
<tr>
<th>Project Title</th>
<th>Institution(s)</th>
<th>Principal Investigator(s)</th>
<th>FY 2004 Award Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative Design and Development of the Community Climate System Model for Terascale Computers</td>
<td>ORNL, LBNL, PNNL, ANL, LLNL, NCAR (Robert Malone, LANL)</td>
<td></td>
<td>$4.0M/Yr</td>
</tr>
<tr>
<td>A Geodesic Climate Model with Quasi-Lagrangian Vertical Coordinates</td>
<td>LANL, Clarkson Univ. (David Randall, Colorado State Univ.)</td>
<td></td>
<td>$802K/Yr</td>
</tr>
<tr>
<td>Predictive Understanding of the Oceans' Wind-Driven Circulation on Interdecadal Time Scales</td>
<td>Indiana Univ. (Michael Ghil, UCLA)</td>
<td></td>
<td>$192K/Yr</td>
</tr>
<tr>
<td>Decadal Variability in the Coupled Ocean-Atmosphere Systems</td>
<td>(Paola Cessi, UCSD)</td>
<td></td>
<td>$155K/Yr</td>
</tr>
<tr>
<td>Improving the Processes of Land-Atmosphere Interactions in CCSM 2.0 at High Resolution</td>
<td>University of Arizona (Robert Dickinson, Georgia Tech)</td>
<td></td>
<td>$245K/Yr</td>
</tr>
<tr>
<td>Continuous Dynamic Grid Adaptation in a Global Atmospheric Model</td>
<td>(William Gutowski, Iowa State Univ.)</td>
<td></td>
<td>$148K/Yr</td>
</tr>
</tbody>
</table>
## Climate Modeling Awards (Continued)

<table>
<thead>
<tr>
<th>Project Description</th>
<th>Investigator(s)</th>
<th>Award Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decadal Regional Climate Studies and Applications with Variable-Resolution GCMs Using Advanced Numerical Techniques</td>
<td>Michael Fox-Rabinovitz, Univ. of Maryland</td>
<td>$228K/Yr</td>
</tr>
<tr>
<td>Multi-Resolution Climate Modeling</td>
<td>Ferdinand Baer, Univ. of Maryland</td>
<td>$180K/Yr</td>
</tr>
<tr>
<td>Development of an Atmospheric Climate Model with Self-Adapting Grid and Physics</td>
<td>Joyce Penner, Univ. of Michigan</td>
<td>$270K/Yr</td>
</tr>
<tr>
<td>Testing a New Hybrid Ocean Circulation Model Based on POP</td>
<td>LANL (Kirk Bryan, Princeton Univ.)</td>
<td>$87K/Yr</td>
</tr>
<tr>
<td>Decadal Regional Climate Studies and Applications with Variable-Resolution GCMs Using Advanced Numerical Techniques</td>
<td>Jean Cote, Univ. of Quebec</td>
<td>$69K/Yr</td>
</tr>
<tr>
<td>Modeling Dynamic Vegetation for Decadal to Multi-Century Climate Change Studies</td>
<td>Goddard Institute for Space Sci. (Andrew Friend, Rutgers Univ.)</td>
<td>$72K/Yr</td>
</tr>
<tr>
<td>Description</td>
<td>Principal Investigator</td>
<td>Funding</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>----------</td>
</tr>
</tbody>
</table>
| Towards the Prediction of Decadal to Multi-Century Processes in a High-Throughput Climate System Model | Zhengyu Liu, Univ. of Wisconsin  
(John Kutzbach, Univ. of Wisconsin) | $231K/Yr |
| Modeling and Analysis of Global and Regional Hydrologic Processes and Appropriate Conservation of Moist Entropy | Donald Johnson, Univ. of Wisconsin  | $231K/Yr |
Collaboratory Awards

SciDAC-Integrating Science and Computing

Three Pilot Collaboratories and One Services Infrastructure Pilot to develop and demonstrate software that allows distributed teams to work together effectively and access data and facilities remotely - $8.7M

<table>
<thead>
<tr>
<th>Project Description</th>
<th>Institution(s)</th>
<th>Funding $ (Yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A National Collaboratory to Advance the Science of High Temperature Plasma Physics for Magnetic Fusion</td>
<td>Princeton Univ., Univ. of Utah, LBNL (David Schissel, General Atomics)</td>
<td>$1.8M/Yr</td>
</tr>
<tr>
<td>Particle Physics Data Grid Collaborative Pilot</td>
<td>Univ. of Wisc., UCSD, Cal Tech, SLAC (Ruth Pordes, FNAL)</td>
<td>$3.19M/Yr (Joint with HENP)</td>
</tr>
<tr>
<td>Earth Systems Grid II: Turning Climate Datasets into Community Resources</td>
<td>UCAR, ORNL, LLNL (Don Middleton, UCAR)</td>
<td>$1.82M/Yr</td>
</tr>
<tr>
<td>DOE Science Grid: Enabling and Deploying the SciDAC Collaboratory Software Environment</td>
<td>ANL, PNNL, ORNL (Bill Johnston, LBNL)</td>
<td>$1.88M/Yr</td>
</tr>
</tbody>
</table>
Large-scale science and engineering is typically done through the interaction of
- People,
- Heterogeneous computing resources,
- Multiple information systems, and
- Instruments
All of which are geographically and organizationally dispersed.
The overall motivation for “Grids” is to enable the routine interactions of these resources to facilitate this type of large-scale science and engineering.

ESNet provides critical infrastructure
- Advanced network infrastructure and specialized network services such as adaptive monitoring supporting high performance distributed applications, and Quality-of-Service and bandwidth reservation for building large-scale aggregated systems
- Collaboration services such as support for IP multicast based video teleconferencing
- Persistent support for directory services that provide naming and location of resources across the Grid and services that enable secure access to these resources
Six Middleware and Network Research Projects that contribute to the development of the Collaboratory Software Environment - $3.4M

<table>
<thead>
<tr>
<th>Project Description</th>
<th>University 1</th>
<th>University 2</th>
<th>Amount Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middleware Technology to Support Science Portals</td>
<td>Indiana Univ.</td>
<td></td>
<td>$182k</td>
</tr>
<tr>
<td>(Dennis Gannon)</td>
<td></td>
<td></td>
<td>$182k/Yr</td>
</tr>
<tr>
<td>A High Performance Data Grid Toolkit: Enabling Tech. For Wide Area Data Intensive Applications</td>
<td>Univ. of Wisconsin</td>
<td>USC</td>
<td>$748k</td>
</tr>
<tr>
<td>(Ian Foster, ANL)</td>
<td></td>
<td></td>
<td>$748k/Yr</td>
</tr>
<tr>
<td>Security and Policy for Group Collaboration</td>
<td>Univ. of Wisconsin</td>
<td>USC</td>
<td>$750k</td>
</tr>
<tr>
<td>(Steve Tuecke, ANL)</td>
<td></td>
<td></td>
<td>$750k/Yr</td>
</tr>
<tr>
<td>Optimizing Performance and Enhancing Functionality of Distributed Applications Using Logistical Networking</td>
<td>Univ. of Tennessee</td>
<td></td>
<td>$500k</td>
</tr>
<tr>
<td>(Micah Beck)</td>
<td></td>
<td></td>
<td>$500k/Yr</td>
</tr>
<tr>
<td>Bandwidth Estimation: Measurement Methodologies and Applications</td>
<td>Wellesley College</td>
<td>Univ. of Delaware</td>
<td>$433k</td>
</tr>
<tr>
<td>(K. Claffy, UCSD)</td>
<td></td>
<td></td>
<td>$433k/Yr</td>
</tr>
<tr>
<td>INCITE: Edge-Based Traffic Processing and Service Inference for High Performance Networks</td>
<td>LANL SLAC</td>
<td></td>
<td>$800k</td>
</tr>
<tr>
<td>(Richard Baraniuk, Rice Univ.)</td>
<td></td>
<td></td>
<td>$800k/Yr</td>
</tr>
</tbody>
</table>
Three projects that will take up the challenge of providing scalable numerical libraries - $8.6M

<table>
<thead>
<tr>
<th>Project</th>
<th>Institutions</th>
<th>Budget/Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terascale Simulation Tools &amp; Technologies Center (TSTT)</td>
<td>ANL, LLNL, ORNL, PNNL, SNL, RPI, SUNY Stony Brook (Jim Glimm, BNL)</td>
<td>$2.6M/Yr</td>
</tr>
</tbody>
</table>
Computer Science Integrated Software Infrastructure Centers

Four activities focused on a comprehensive, portable, and fully integrated suite of systems software and tools for effective utilization of terascale computers - $10.7M

<table>
<thead>
<tr>
<th>Activity</th>
<th>Participants</th>
<th>Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalable Tools for Large Clusters; Resource Management; System Interfaces; System Management Tool Framework (SSS)</td>
<td>ANL      LBNL       Ames       PNNL       SNL       LANL       National Center for Supercomputing App (AI Geist, ORNL)</td>
<td>$2.2M/Yr</td>
</tr>
<tr>
<td>High-End Computer Systems Performance: Science &amp; Engineering (PERC)</td>
<td>ANL      LLNL       ORNL       U. of Illinois U. of Maryland UCSD       U. of Tennessee U. of Maryland (David Bailey, LBNL)</td>
<td>$2.4M/Yr</td>
</tr>
<tr>
<td>Center for Component Technology for Terascale Simulation Software (CCA)</td>
<td>ANL      LANL       LLNL       ORNL       PNNL       U. of Utah Indiana U. U. of Maryland (Rob Armstrong, SNL)</td>
<td>$3.1M/Yr</td>
</tr>
<tr>
<td>Scientific Data Management Enabling Technology Center (SDM)</td>
<td>ANL      LLNL       ORNL       Georgia Tech. UCSD       Northwestern U. North Carolina State (Ari Shoshani, LBNL)</td>
<td>$3M/Yr</td>
</tr>
</tbody>
</table>
Scalable Systems Software for TeraScale Computers

www.scidac.org/ScalableSystems

Problem
- Computer centers use incompatible, ad hoc set of systems tools
- Present tools are not designed to scale to multi-Teraflop systems

Solution
- Collectively (with industry) define standard interfaces between systems components for interoperability
- Create scalable, standardized management tools for efficiently running our large computing centers

Impact
- Revolutionize the way system software is designed and used.
Developing a science for understanding performance of scientific applications on high-end computer systems, and engineering strategies for improving performance on these systems.

**GOALS**

Optimize and Simplify:
- Profiling of real applications
- Measurement of machine capabilities (emphasis on memory hierarchy)
- Performance prediction
- Performance monitoring
- Informed tuning

Understand the key factors in applications that affect performance.
Understand the key factors in computer systems that affect performance.
Develop models that accurately predict performance of applications on systems.
Develop an enabling infrastructure of tools for performance monitoring, modeling and optimization.
Validate these ideas and infrastructure via close collaboration with DOE SC and other application owners.
Transfer the technology to end-users.

**ENABLING TECHNOLOGIES**

- Application Signatures
- Machine Signatures
- Bound Models
- Phase Models
- ROSE
- SvPablo
- PAPI
- Sigma++ (IBM)
- DynInst

**DOE LABS**
- ANL
- LBNL
- LLNL
- ORNL

**UNIVERSITIES**
- UCSD
- UI-UC
- UTenn
- UMD
Component Technology for Terascale Simulation Software: Bringing Components to High Performance Computing

Currently DOE Apps Are:
- One-off, stove-pipe, few participants
  
But Also:
- Scalable, and low latency
- Large & rich legacy investment

Challenges
- Practical parallel HPC component model
- Integrate legacy software investment
- Connection to Grid components
- Multi-language

Goals
- Enable plug-and-play parallel simulations
- Establish component "marketplace" with applications partners
- Extend commodity component technology for HPC

Today:
Single-purpose, monolithic, tightly-coupled parallel codes

HPC Component Framework

SCIENTIFIC COMPONENTS

“MxN” Parallel Data Redistribution

APPLICATIONS INTEGRATION

Port

Industry: Discretization Engine

DOE Lab: Implicit Solve

University: Visualization

Link

Component-Based Scientific Application

Participants:
- ANL
- Indiana Univ.
- LANL
- LLNL
- ORNL
- PNNL
- SNL
- Univ. of Utah

Partners:
- Climate Modelling
- Quantum Chemistry
- Combustion Modelling

Common Component Architecture

CCA

http://www.cca-forum.org/ccttss
Scientific Simulations & experiments
- Climate Modeling
- Astrophysics
- Genomics and Proteomics
- High Energy Physics

Data Manipulation:
- Getting files from Tape archive
- Extracting subset of data from files
- Reformatting data
- Getting data from heterogeneous, distributed systems
- Moving data over the network

Scientific Analysis & Discovery

SDM-ISIC Technology
- Optimizing shared access from mass storage systems
- Metadata and knowledge-based federations
- API for Grid I/O
- High-dimensional cluster analysis
- High-dimensional indexing
- Adaptive file caching
- Agents...

Goals
Optimize and simplify:
- Access to very large datasets
- Access to distributed data
- Access of heterogeneous data
- Data mining of very large datasets

Current

Goal
### Summary of Interactions

#### SciDAC - Integrating Science and Computing

<table>
<thead>
<tr>
<th>Office</th>
<th>Project Title</th>
<th>SA $</th>
<th>CCA</th>
<th>SDM</th>
<th>PERC</th>
<th>SSS</th>
<th>TSTT</th>
<th>TOPS</th>
<th>APDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FES</td>
<td>Magnetic Reconnection: Applications to Sawtooth Oscillations, Error Field Induced Islands and the Dynamo Effect</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FES</td>
<td>Numerical Computation of Wave-Plasma Interactions in Multi-dimensional Systems</td>
<td>225</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FES</td>
<td>Center for Extended Magnetohydrodynamic Modeling</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FES</td>
<td>The Plasma Microturbulence Project</td>
<td>35</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HENP</td>
<td>Advanced Computing for 21st Century Accelerator Science and Technology</td>
<td>655</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HENP</td>
<td>National Computational Infrastructure for Lattice Gauge Theory</td>
<td>134</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HENP</td>
<td>Shedding New Light on Exploding Stars: Terascale Simulations of Neutrino-Driven SuperNovae and their Nucleosynthesis</td>
<td>416</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HENP</td>
<td>SciDAC Center for Supernova Research</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BER</td>
<td>Collaborative Design and Development of the Community Climate System Model for Terascale Computers</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>A Computational Facility for Reacting Flow Science</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>Terascale High-Fidelity Simulations of Turbulent Combustion with Detailed Chemistry</td>
<td>138</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>Advanced Methods for Electronic Structure</td>
<td>250</td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>Advancing Multi-Reference Methods in Electronic Structure Theory</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>Accurate Properties for Open-Shell States of Large Molecules</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BES</td>
<td>Advanced Software for the Calculation of Thermochemistry, Kinetics and Dynamics</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>DOE Science Grid: Enabling and Deploying the SciDAC Collaboratory Software Environment</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>A National Collaboratory to Advance the Science of High-Temperature Plasma Physics for Magnetic Fusion</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Particle Physics Data Grid Collaborative Pilot</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Earth System Grid II: Turning Climate Datasets into Community Resources</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>A High-Performance Data Grid Toolkit: Enabling Technology for Wide Area Data-Intensive Applications</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Scalable Systems Software Center (SSS)</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>High-End Computer Systems Performance: Science and Engineering (PERC)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Center for Component Technology for Terascale Simulation Software (CCA)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Scientific Data Management Center (SDM)</td>
<td></td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Terascale Optimal PDE Simulations (TOPS)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Terascale Simulation Tools &amp; Technologies Center (TSTT)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCR</td>
<td>Algorithmic and Software Framework for Applied Partial Differential Equations (APDE)</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>