Mathematical, Information and Computational Sciences Program

An Overview

Princeton Plasma Physics Laboratory
August 29, 2002

Walter M. Polansky, Acting Director
Mathematical, Information and
Computational Sciences Division

Staff

Mathematical, Information and Computational Sciences

- Ed Oliver, Associate Director for Advanced Scientific Computing Research
- Dan Hitchcock, Senior Scientific Advisor
- Linda Twenty, Senior Budget & Financial Specialist

- Walt Polansky, Acting Director MICS

- Gary Johnson, ACRTs, Computational Biology
- Fred Johnson, Computer Science
- William (Buff) Miner, NERSC & Scientific Applications
- Thomas Ndousse-Fetter, Network Research
- Chuck Romine, Applied Mathematics
- Mary Anne Scott, Collaboratories
- George Seweryniak, Esnet
- John van Rosendale, Computer Science- Visualization and Data Management

- Vacancies- (2)

- Jane Hiegel
- Susan Kilroy

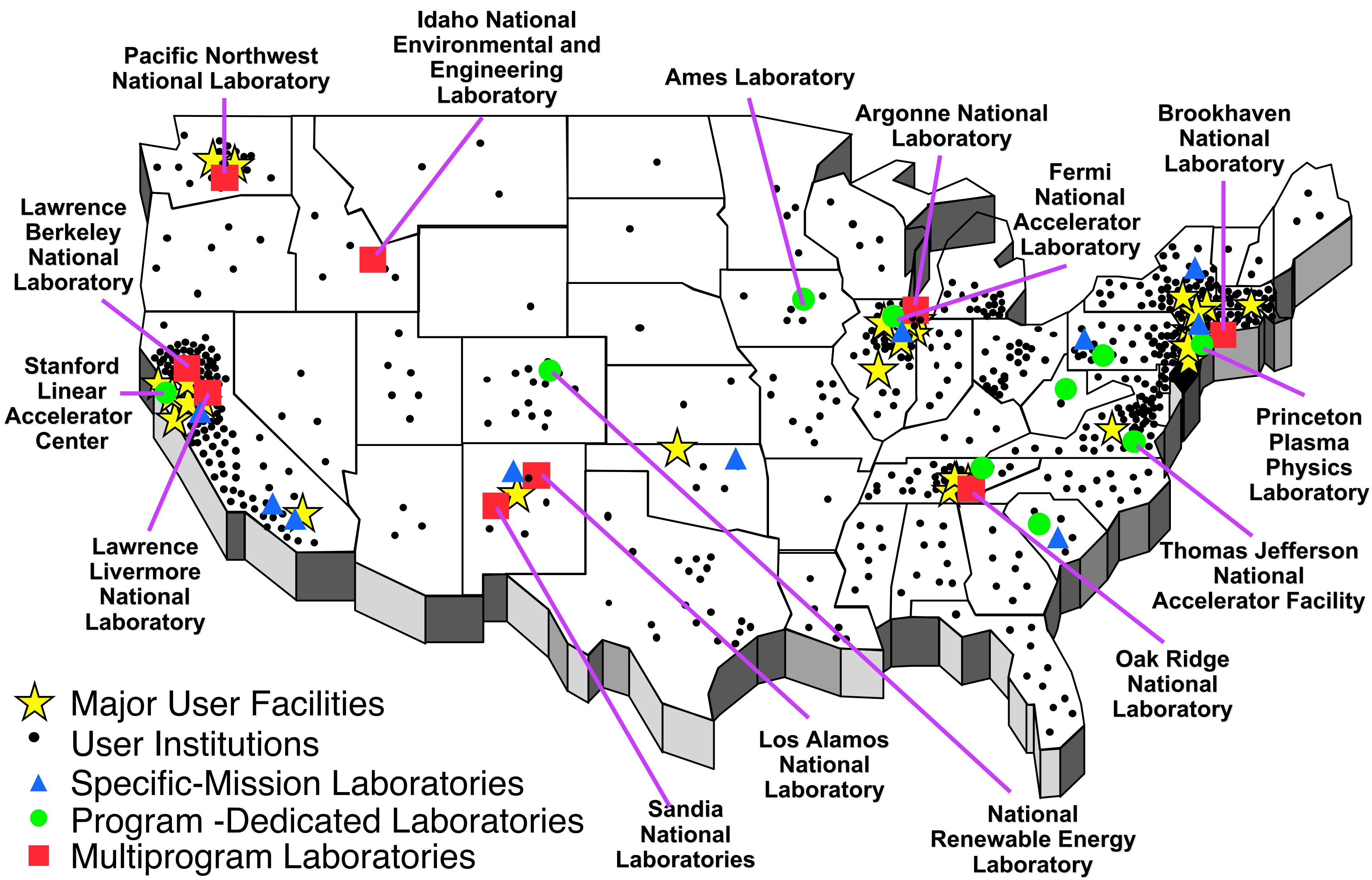
Phone- 301-903-5800
Fax- 301-903-7774
http://www.sc.doe.gov/production/octr/mics/index.html

Mission and Priorities for the Department

- Guaranteeing the safety and reliability of the nuclear stockpile;
- Ensuring the R&D and production plans support the Administration's nuclear strategy;
- Resolving the threat of weapons of mass destruction;
- Provide safe, efficient and effective nuclear power for the Navy;
- Ensuring energy security through infrastructure protection;
- Implementing the President's National Energy Plan;
- Directing R&D budgets to innovative new ideas while ensuring applications of mature technologies;
- Exploring new energy sources with dramatic environmental benefits;
- Supporting Homeland Defense through a focus on the threats of weapons of mass destruction posed by terrorist or nation states

Strong Relevance to Mathematical, Information and Computational Sciences

High Performance Computing, Collaboration and Networks- Critical for DOE Science



Mission

Discover, develop, and deploy the computational and networking advances that enable researchers in the scientific disciplines to analyze, model, simulate, and predict complex physical, chemical, and biological phenomena important to the Department of Energy (DOE).

support a broad research portfolio in advanced scientific computing – applied mathematics, computer science, networking and collaboratory software

operate supercomputers, a high performance network, and related facilities.

Program Strategy

Mathematical, Information and Computational Sciences

Research to enable…

Basic Research

…simulation of complex systems

…distributed teams, remote access to facilities

BES, BER, FES, HEP, NP

- Materials
- Chemical
- Combustion
- Accelerator
- HEP
- Nuclear
- Fusion
- Climate
- Astrophysics

Computational Biology

- Integrated Software Infrastructure Centers

Teams- mathematicians, computer scientists, application scientists, and software engineers

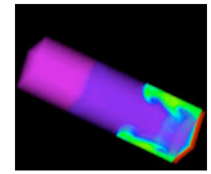
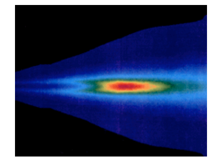
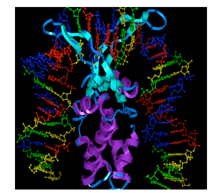
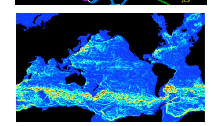
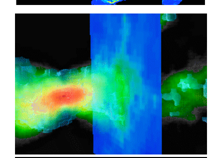
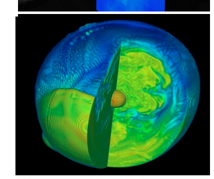
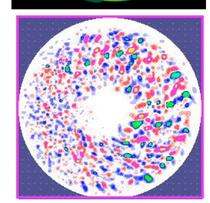
- Grid enabling research
- Nanoscience
- Topical Computing

SciDAC!

- Applied Mathematics
- Computer Science
- Collaboratory Tools
- Networking
- Scientific Application Pilots
- Collaboratory Pilots

High Performance Computing and Network Facilities for Science

- National Energy Research Scientific Computing Center (NERSC)
- Advanced Computing Research Facilities
- Energy Sciences Network (ESnet)

Principles and Attributes

Mathematical, Information and Computational Sciences

- MICS will advance DOE science and missions through
  - world-class applied mathematics and computer science research,
  - innovative software technologies
  - cost-effective high-performance computational and networking
- Strategies-
  - review MICS management processes and MICS portfolio
  - expand applied math and computer science base research
  - develop research partnerships with all Office of Science programs
  - acquire/deploy high-end computational and network resources in a systematic manner
  - enhance coordination with other agencies
- Values-
  - excellence and innovation in everything we do
  - consistent high standards to all existing and prospective research performers
  - the best research and PIs, regardless of affiliation

FY2001 MICS Research Budget by Institution
$ in thousands (# of projects)

Mathematical, Information and Computational Sciences

	Base Research					
	AMS	CS	NC-ACST	NR	SAPP	SciDAC
Univ. (& Others)	3,903	4,380	1,141	885	1,703	17,548
Laboratories	13,345	8,461	11,284	2,444	1,120	19,895
Totals	17,248 (43)	12,841 (25)	12,425 (20)	3,329 (13)	2,823 (11)	37,443 (34)

Legend
AMS- Applied Mathematical Sciences
CS- Computer Sciences
NC-ACST- National Collaboratories- Advanced Computing Software Tools
NR- Networking Research
SAPP- Scientific Application Pilot Projects- Computational Biology
SciDAC- Scientific Discovery through Advanced Computing

Award Size	Mean	Median
Base	$ 435	$ 198
SciDAC	$ 1,100	$ 811

FY2001 MICS Budget
Laboratories ($ in thousands)

Mathematical, Information and Computational Sciences

	Base Research	SciDAC	Facilities
Ames	1,698	66	0
ANL	7,656	2,157	2,037
BNL	155	182	0
LBNL	8,512	1,851	53,424
LLNL	3,625	1,273	0
LANL	4,270	207	1,250
ORNL	4,283	1,117	15,657
PNNL	2,321	688	22
Sandia	4,374	422	0
SLAC	234	81	0

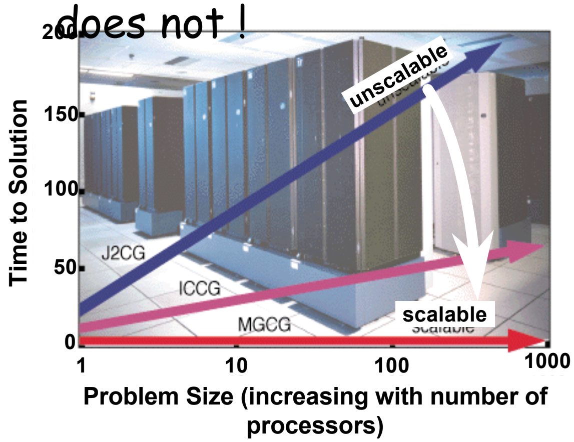
Legend
Facilities- NERSC, ACRTs, ESnet
SciDAC- Scientific Discovery through Advanced Computing

Applied Mathematical Sciences

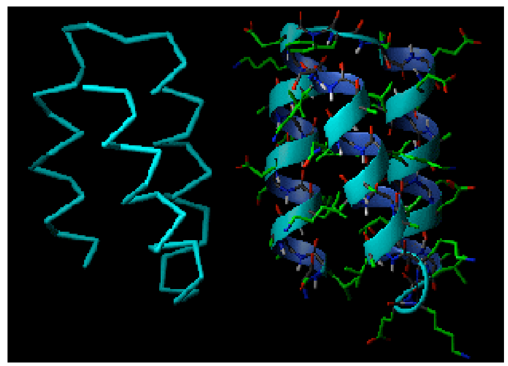
From the "simple"… …to the complex!

Ax=b	→	Ax=□Bx	→	F(u,x,y,z)=0	→	F(u,u',u",…,x,y,z,t)=0
Linear Solvers		Eigensolvers		Nonlinear Solvers		Partial Differential Equation Solvers

Algorithms must be *scalable. Ideally*, as the problem size grows and the number of processors grows, the solution time does not !



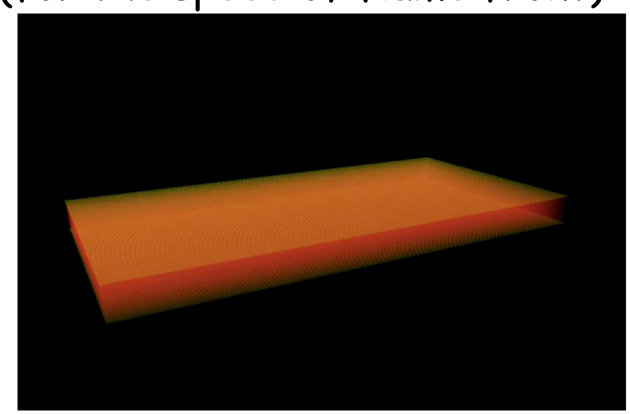
Protein Folding



Current simulations use 44 amino acids.
Actual protein ~300 amino acids. Run times using current techniques? Greater than life of the universe!

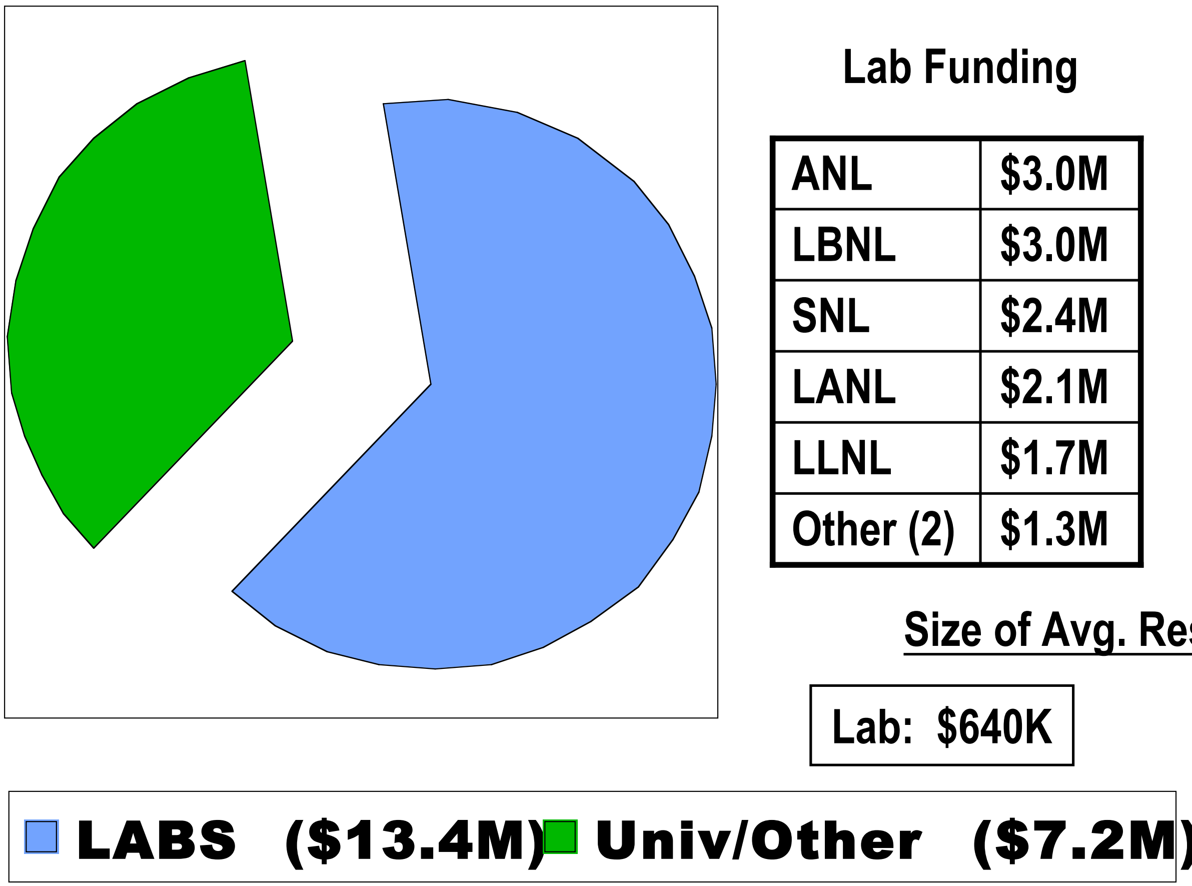
Combustion

~60 *coupled, nonsymmetric, nonlinear time-dependent* PDEs on 10M mesh points. Time steps range from 10^{-12} (for chemical reaction rates) to 10^{-2} (for the speed of flame front)



FY2001 Budget
Applied Mathematical Sciences

Mathematical, Information and Computational Sciences



Lab Funding

ANL	$3.0M
LBNL	$3.0M
SNL	$2.4M
LANL	$2.1M
LLNL	$1.7M
Other (2)	$1.3M

Univ/Other Funding

CSGF	$3.00M
NYU	$1.20M
SUNYSB	$0.50M
MSRI	$0.25M
Brown U.	$0.22M
Other (19)	$2.00M

Size of Avg. Research Project

Lab: $640K

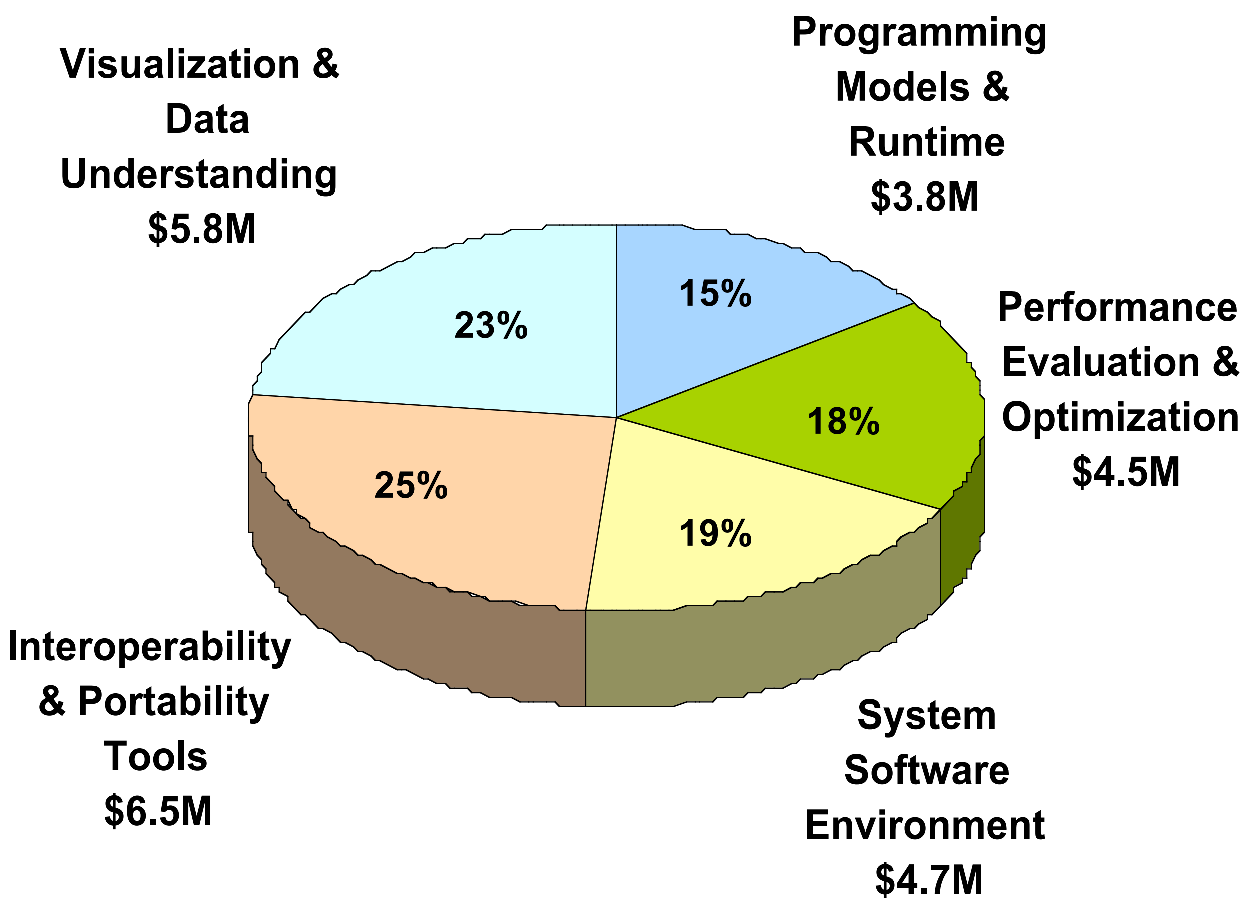
Univ: $154K

Computer Science

- Base Program
  - Evolutionary and revolutionary software methodologies for future generations of HPC architectures
- SciDAC Integrated Software Infrastructure Centers
  - Enable effective application of current terascale architectures to SciDAC applications through focused research and partnerships

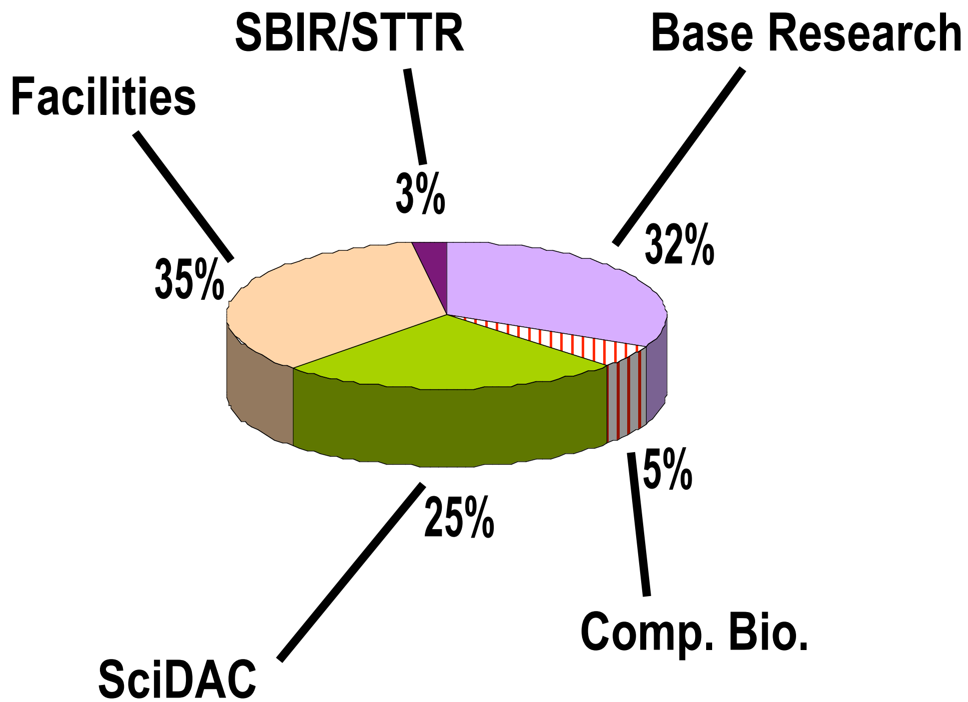
Computer Science Technical Elements

Mathematical, Information and Computational Sciences



Budget Request

FY2003- $166,625,000



Enhancements over FY2002

- **Computational Biology** **+$5.6M**
- **SciDAC** **+$5.3M**
- **Facilities** **+$1.3M**

Program Evolution

- **FY 2001**
  - Initiated software infrastructure portion of SciDAC
  - Initiated computational biology research efforts
  - Upgraded NERSC to 5 teraflops
  - Acquired IBM Power 4 Hardware for evaluation/scaling studies

- **FY 2002**
  - Convened first SciDAC Principal Investigator Meeting
  - Issued research call for Early Career Principal Investigators
  - Approved LBNL proposal for management and operation of NERSC
  - Conducted several Genomes to Life workshops with BER
  - Planned workshops with BES (computational nanoscience) and FES

- **FY 2003 Plans**
  - Launch computational component of Genomes to Life, in partnership with BER
  - Initiate computational nanoscience partnership with BES as part of SciDAC
  - Provide topical high performance computing resources to support SciDAC research
  - Lead development of SC strategy for high-performance networks

Background

- The Earth Simulator, Japan's New Supercomputer, presents both a formidable challenge and an opportunity for DOE computational science.

- Domestically, commercial market forces (e-business) determine the computational choices for science.

- Bandwidth and latency- (processor to memory and processor to processor) are major "bottlenecks" for science.

- Opportunity to make a case for providing new building blocks for scientific computing.

Oops !

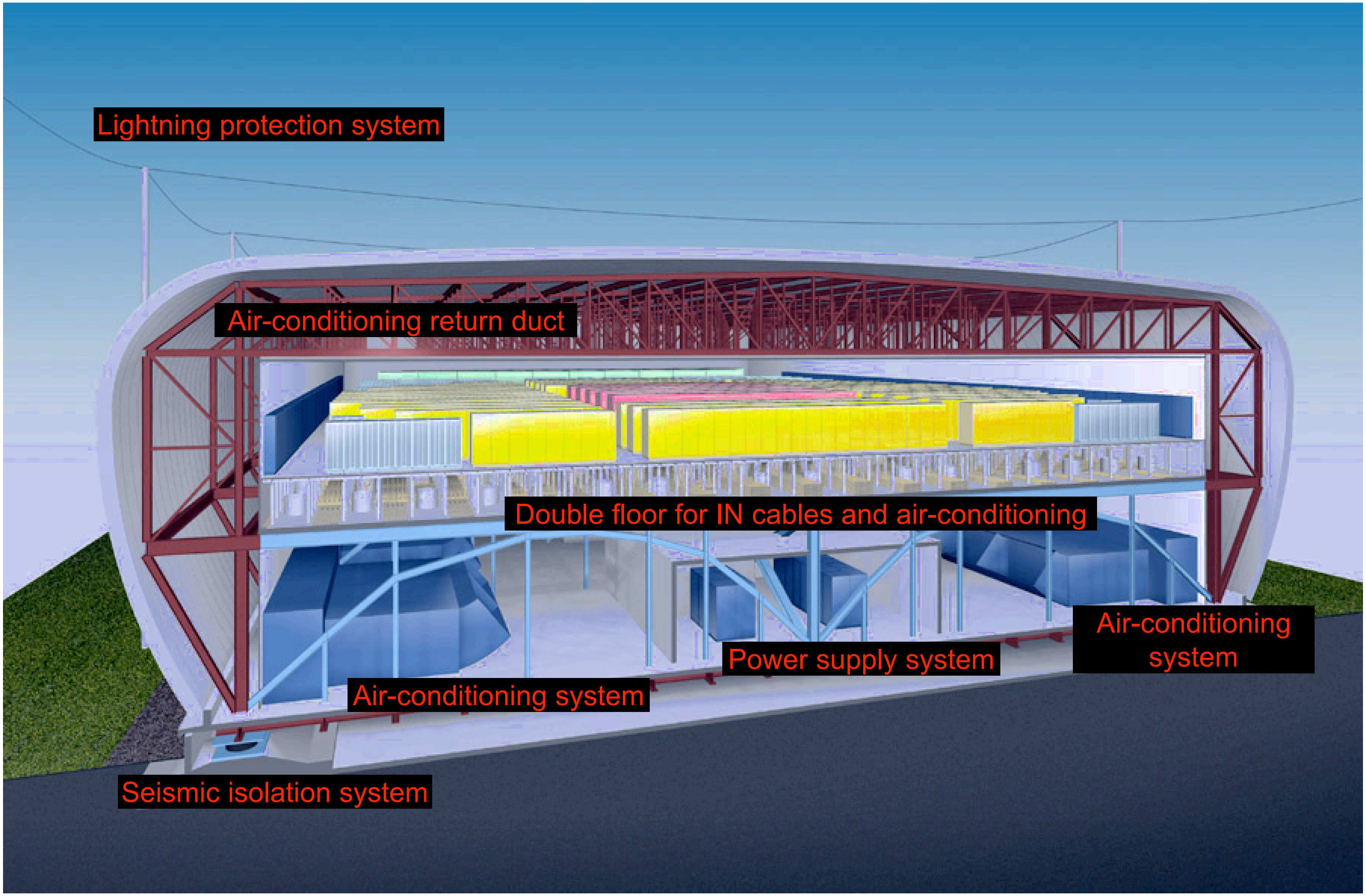
6/02

Rank	Manufacturer	Computer	Rmax	Installation Site	Country	Year	Area of Installation	# Proc	Rpeak	Nmax	N1/2
1	NEC	Earth-Simulator	**35860**	Earth Simulator Center Kanazawa	Japan	2002	Research	5120	40960	1075200	266240
2	IBM	ASCI White, SP Power3 375 MHz	**7226**	Lawrence Livermore National Laboratory Livermore	USA	2000	Research Energy	8192	12288	518096	179000
3	Hewlett-Packard	AlphaServer SC ES45/1 GHz	**4463**	Pittsburgh Supercomputing Center Pittsburgh	USA	2001	Academic	3016	6032	280000	85000
4	Hewlett-Packard	AlphaServer SC ES45/1 GHz	**3980**	Commissariat a l'Energie Atomique (CEA) Bruyeres-le-Chatel	France	2001	Research	2560	5120	360000	85000
5	IBM	SP Power3 375 MHz 16 way	**3052**	NERSC/LBNL Berkeley	USA	2001	Research	3328	4992	371712	102400
6	Hewlett-Packard	AlphaServer SC ES45/1 GHz	**2916**	Los Alamos National Laboratory Los Alamos	USA	2002	Research	2048	4096	272000	.
7	Intel	ASCI Red	**2379**	Sandia National Laboratories Albuquerque	USA	1999	Research	9632	3207	362880	75400
8	IBM	pSeries 690 Turbo 1.3GHz	**2310**	Oak Ridge National Laboratory Oak Ridge	USA	2002	Research	864	4493	275000	62000
9	IBM	ASCI Blue-Pacific SST, IBM SP 604e	**2144**	Lawrence Livermore National Laboratory Livermore	USA	1999	Research Energy	5808	3868	431344	.
10	IBM	pSeries 690 Turbo 1.3GHz	**2002**	IBM/US Army Research Laboratory (ARL) Poughkeepsie	USA	2002	Vendor	768	3994	252000	.
		SP Power3 375 MHz		Atomic Weapons							

Earth Simulator

- Based on the NEC SX architecture, 640 nodes, each node with 8 vector processors (8 Gflop/s peak per processor), 2 ns cycle time, 16GB shared memory.
  - Total of 5104 total processors, 40 TFlop/s peak, and 10 TB memory.
- It has a single stage crossbar (1800 miles of cable) 83,000 copper cables, 16 GB/s bandwidth, into and out of each node.
- 700 TB disk space
- 1.6 PB mass store
- Area of computer = 4 tennis courts, 3 floors

Cross-sectional View of the Earth Simulator Building



Global Atmospheric Simulation with the Spectral Transform Method on the Earth Simulator

- 26.6 Tflops
- 640 nodes
- 65% of peak
- High Performance Fortran
- AFES - spectral atmospheric general circulation model using Fourier transform and Legendre transforms
- Grid resolution – T1279L96 (grid interval of about 10km around the equator)

Three-dimensional Fluid Simulation for Fusion Science with HPF on the Earth Simulator

- 12.5 TFLOPS
- 512 nodes
- 39% of peak
- High Performance Fortran
- IMPACT-3D three-dimensional compressible and inviscid Eulerian fluid computation with explicit 5-point stencil scheme for spatial differentiation
- Mesh size – 2048x2048x4096

Software Environment- Earth Simulator

- Operating System
  - UNIX-based (Enhanced version of NEC SUPER-UX)
  - Parallel file system (MPI-IO), high-performance Fortran)
- Programming Environment
  - Parallel programming environment ({Fortran90, C}+MPI2, HPF2)
  - Program Analyzer + Tuning tools
  - Parallel debugger
- Flexible job scheduler
  - Earth Simulator specifications
  - Job assignment to processing nodes with file loading to appropriate system disks

Software has good scalability to 640 nodes

Operating system, compiler and libraries- NEC

Debuggers and performance tools- 3rd party vendors.

Calendar of Events

April 20, 2002	New York Times	"Japanese Computer Is World's Fastest, as U.S. Falls Back"
May 15-16,2002	Earth Simulator Rapid Response Meeting	ES performance a credible threat to US computational science leadership
June 12, 2002	IBM/ORNL/NCAR meeting	
June 14, 2002	OSTP Meeting (Marburger)	Possible need for interagency response
June 20, 2002	Cray/ORNL/NCAR meeting	
June 19-20, 2002	Visit to NASA Ames	
June 21, 2002	Visit to Silicon Graphics, Inc.	
July 8, 2002	SIAM Mini-Symposium	Presentation of ES challenge
July 17, 2002	SAC Meeting	Overviews of ES and challenge to SC
July 22, 2002	DOE visit to the Earth Simulator	Yokohama, Japan
August 5-30, 2002	Town Meetings with Fusion, Chemistry, Astrophysics, Accelerator Designers, Network communities	

The Science must drive…

…the response to the challenge

For fusion:

- What is the status of the science in this field ?
- How can this science be advanced through simulations at the ultrascale ?
- Why are these advances important to the field ? …to the Office of Science and the DOE ?
- What breakthrough simulations need to be performed ? What knowledge will result ? What would be the benefit to the Office of Science and the DOE ?
- What computational and networking resources would be needed to perform breakthrough simulations ? When would you be ready to utilize those resources ?
- What challenge does the Earth Simulator pose to your field of science ?

Expected Outcomes from Town Meetings

- Further discussions throughout the Office of Science computational community about opportunities presented by ultrascale computing.
- Contribution(s) to "Building the Science Case for Ultra Scale Simulation", http://www.krellinst.org/esinfo/ (See "Reasserting U.S. Leadership…" for a suggested format.)
- Further dialog with us.