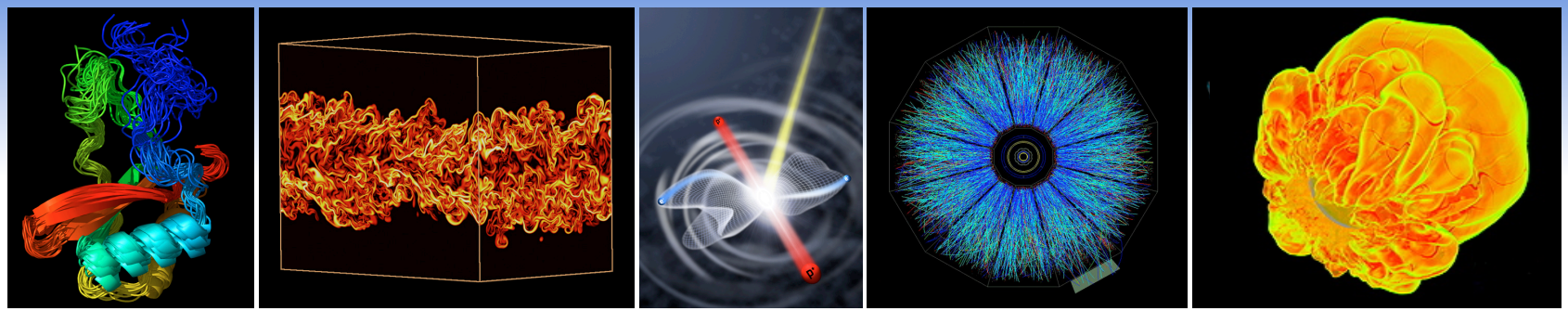


BER/NERSC/ASCR Requirements Workshop
Large Scale Computing and Storage Requirements for Biological and Environmental Research

NERSC
Lawrence Berkeley National Laboratory

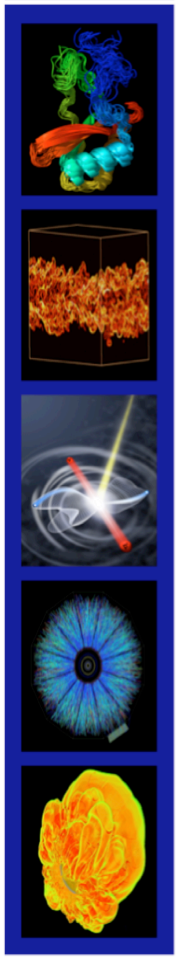
Workshop Held May 7-8, 2009, Rockville, MD

U.S. DEPARTMENT OF ENERGY | Office of Science

NERSC

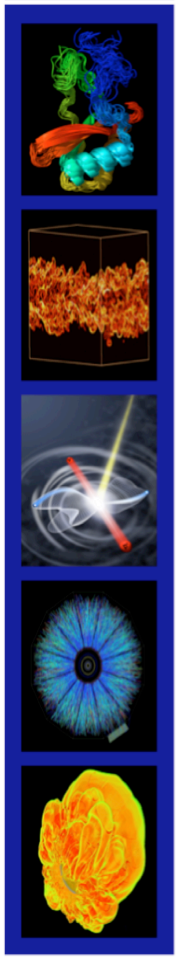
BERKELEY LAB

Yukiko's Charge to Workshop Attendees



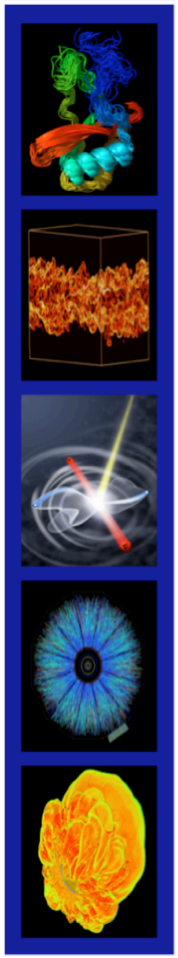
- The goal of this workshop is to accurately characterize the High Performance Computing (HPC) requirements of current and future work funded by the Office of Biological and Environmental Research (BER).
- These requirements will serve as input to the NERSC architecture and planning processes, and will help ensure that NERSC continues to provide world-class support for scientific discovery for DOE scientists and their collaborators.
- The tangible outcome of the workshop will be a document that includes both HPC requirements and a supporting narrative.

Case Studies



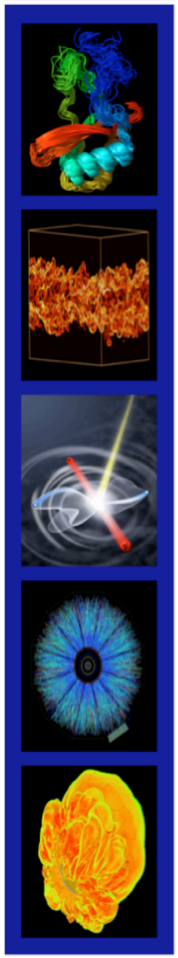
- Workshop discussions and final report based on “case studies”
- Narratives describing science & NERSC requirements
- Based on format used by ESNet
- Tried so select the minimum set of case studies needed to characterize research of relevance to the BER mission as it relates to NERSC

Case Study Titles



NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

Participants



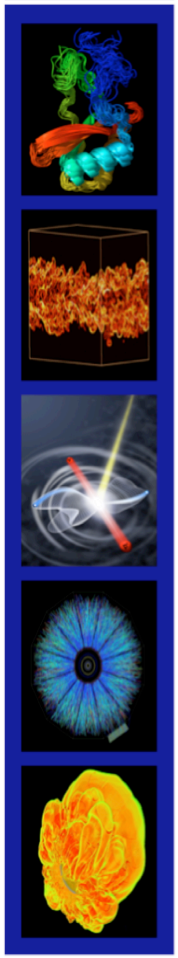
- R. Todd Anderson ,DOE / BER Program Manager
- Ioan Andricioaei , UC Irvine ,Biological Science
- David Beck, U. of Washington, Biological Science
- Anjuli Bamzai ,DOE / BER, BER Program Manager
- Lawrence Buja ,NCAR ,Climate Science
- Dylan Chivian ,LBNL ,Biological Science
- John Dennis ,NCAR ,Climate Science
- John Drake ,ORNL ,Climate Science
- Richard Gerber ,NERSC, Facilitator
- Susan Gregurick ,DOE / BER Program Manager
- Teresa Head-Gordon ,LBNL ,Biological Science
- Christopher Kerr ,GFDL ,Climate Science
- Costas Maranas ,Penn State U., Biological Science
- Victor Markowitz ,LBNL ,JGI
- Lee Ann McCue ,PNNL ,Biological Science
- Lucy Nowell, DOE/ASCR Program Manager
- Bruce Palmer ,PNNL ,Environmental Science
- Edward Patton, NCAR , Climate / Environ Science
- David Randall ,Colorado State U. ,Climate Science
- Timothy Scheibe ,PNNL ,Environ Science
- Yukiko Sekine, DOE / ASCR, NERSC Program Manager
- John Shalf ,NERSC, System Architecture
- Jeremy Smith ,ORNL ,Biological Science
- David Thomassen, DOE/BER, Chief Scientist
- Harvey Wasserman, NERSC , Facilitator
- Michael Wehner , LBNL ,Climate Science
- Christopher L. Wolfe Scripps ,Climate Science
- Kathy Yelick ,NERSC ,NERSC Director

U.S. DEPARTMENT OF ENERGY | Office of Science

The Office Of ASCR Advanced Scientific Computing Research

BERKELEY LAB

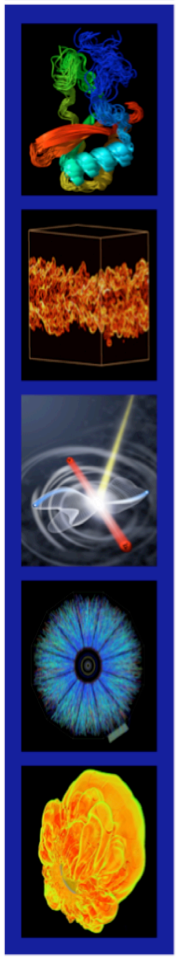
Findings



1. Without a significant increase in available resources, scientific progress will be acutely limited.
2. Many projects have mission-critical time constraints that demand guaranteed throughput on powerful resources.
3. To exploit many-core (10K – 1M) computers, BER scientists request access to test-bed machines and assistance from NERSC in choosing effective programming models.

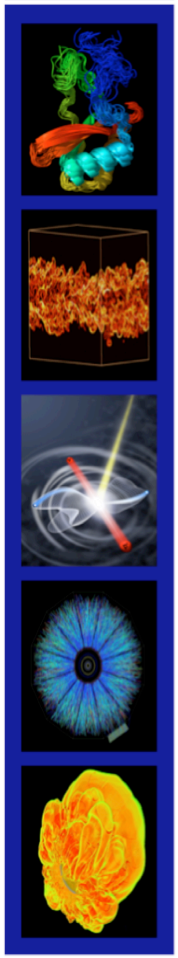
NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

Findings



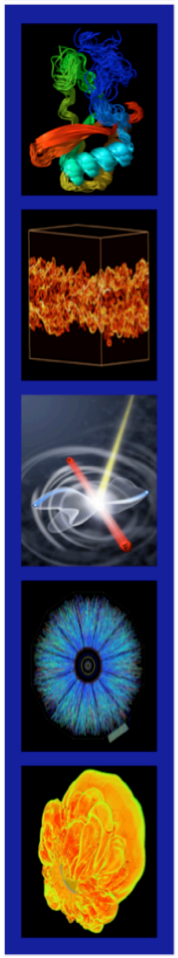
4. There is a need for high-throughput development and analysis services to facilitate uncertainty quantification.
5. Ensemble runs are important methods for scientific discovery and simulation validation.
6. Researchers need access to resources that are stable, available, and reliable. Scientific productivity rivals machine “speeds and feeds” in importance.

Findings



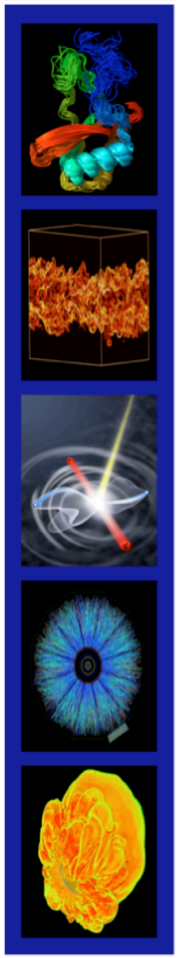
7. Manipulation and analysis of data is becoming a problem that can be addressed only by large HPC systems.
8. Several BER projects require efficient parallel I/O via MPI-IO and high-level libraries.
9. Highly efficient global reduction operations are needed.

Findings



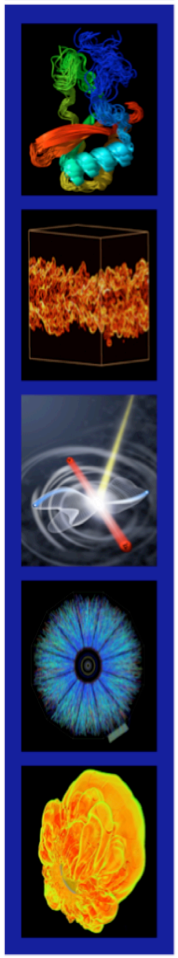
10. Memory requirements are generally 2 GB per core or less. Scientists want to use more memory per core, but realize that architectural trends do not favor this and they may need help reducing their memory footprint.
11. Users underscored a critical need for highly available and reliable HPC resources to support scientific discovery and productivity.

1. Request for Resources



- Users have a need for more HPC resources
  - ✔ Workshop participants need about 10X increase over current allocation.
- Strong request for support services
  - ✔ Help selecting new programming languages and paradigms
  - ✔ Training, guidance
  - ✔ Interfaces: common scheduler, portals support
  - ✔ Agile, timely consulting support; help with workflows
  - ✔ Help optimizing I/O

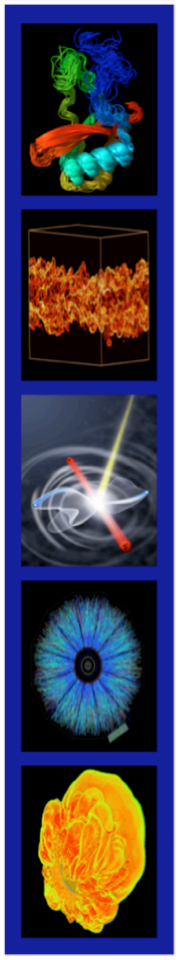
2. Predictable Throughput



Need for predictable throughput

- Microbial Genomics (JGI) and climate modeling for IPCC reports must meet demanding schedules
- Can't have predictable throughput if a job has to restart at the back of queue when it fails due to system errors
- Need quick turn around or reserved access for developing and debugging at scale

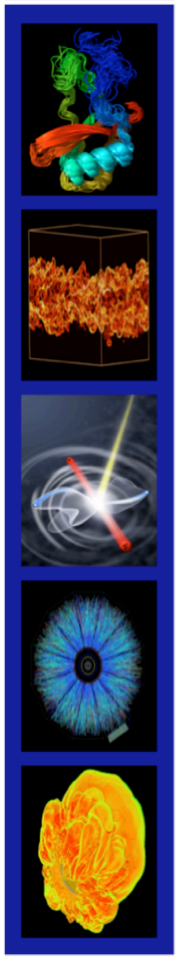
3. Many-Core/Accelerator Programming Models



Users need guidance / help programming for many-core systems

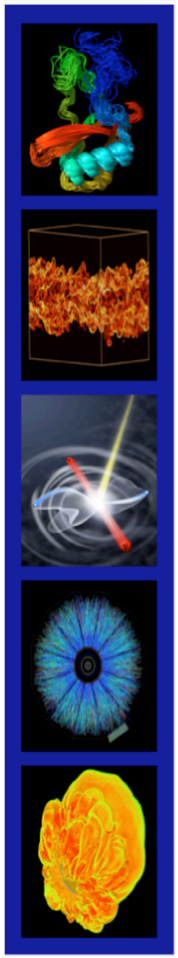
- ✔ Need training, advice, consulting support, deployment of new programming software on production machines
- ✔ Request for NERSC to provide a training test bed for heterogeneous and many-core systems

5. Ensemble / Long Runs



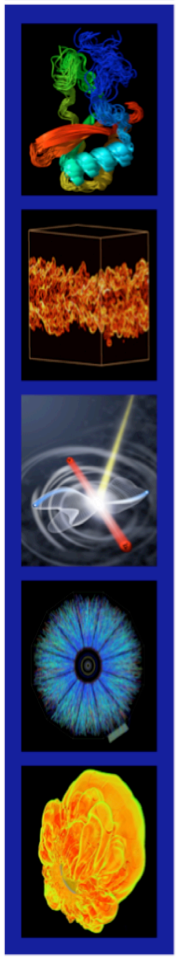
- Ensemble runs are an important method for scientific discovery
  - ✔ Simultaneous instances of single code
  - ✔ Different initial or boundary conditions
  - ✔ Important for Uncertainty Quantification studies
- Some jobs need a long run time at lower concurrency
  - ✔ May be difficult to checkpoint; or a community application may not have the ability
  - ✔ Checkpoint I/O times may be excessive
  - ✔ Eases workflow complications

6. Stability, Availability, Reliability



- Scientists want stable systems
  - ✔ For reliable throughput
  - ✔ Don't want to implement their own fault tolerance
  - ✔ Don't want the hassle of dealing with failed jobs
- Very much want two big systems at NERSC
  - ✔ One stable production system
  - ✔ One newer system for moving forward; OK if it is still undergoing "new-system" shakeout instabilities
- Scientific Productivity is the goal
  - ✔ It's not all about "speeds and feeds"
  - ✔ Make scientific productivity a reportable metric

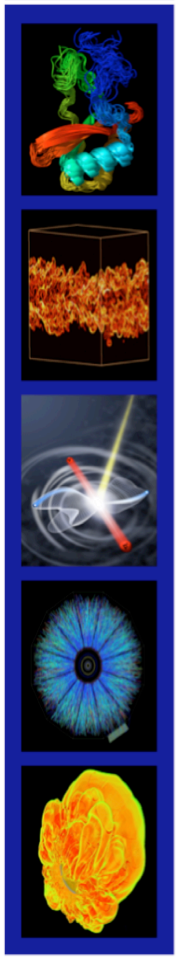
7. Data



- Data is a problem
  - ✔ Data storage needs are increasing exponentially.
  - ✔ Output is too large to move to local sites.
  - ✔ Data analysis requires HPC: *in situ* or *post-hoc*
  - ✔ Users want data management advice from NERSC.
  - ✔ NERSC should provide workflow, data portal, & database technologies.
  - ✔ A significant increase in real-time-accessible data storage capacity is needed.

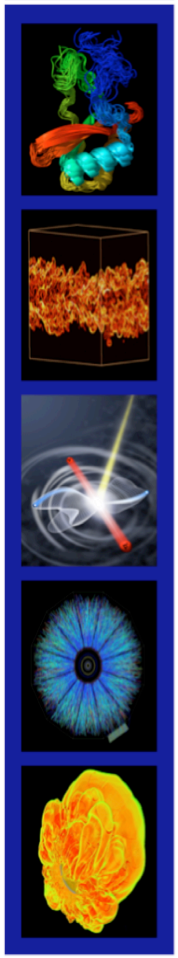
NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

8. I/O



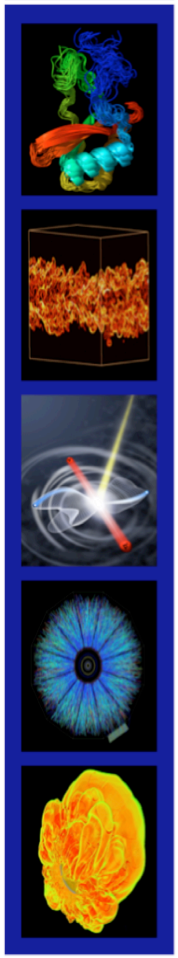
- Users need efficient parallel I/O libraries
  - ✔ HDF5
  - ✔ netCDF
  - ✔ MPI-IO
  - ✔ Need DOE investment to maintain software infrastructure
- I/O rates need improvement
  - ✔ Don't want I/O to consume >10% of run time
  - ✔ Some users feel checkpointing "not worth it" because I/O takes too long
- Guidance wanted to help achieve efficient I/O

9. MPI & 10. Memory



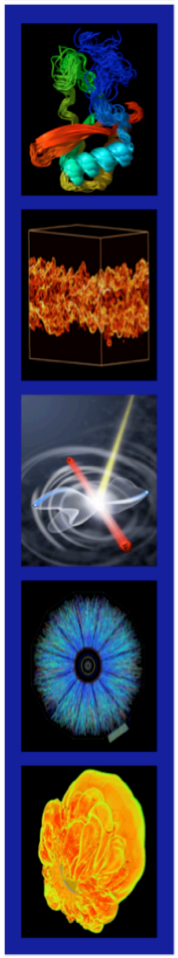
- Communications Needs
  - ✔ Fast global reductions
  - ✔ Low-latency MPI
  - ✔ Several key algorithms use CG solver
- Memory Requirements
  - ✔ Many codes can fit into 2 GB/task
  - ✔ Some are more efficient with access to more memory
  - ✔ Scientists want more memory/core, but realize that is not the architectural trend; want programming help

11. Discovery & Productivity



- Support effective workflows
  - ✔ Kepler/SDM
  - ✔ Special workflow needs
  - ✔ Computational “beam line” (dedicated, reserved periods of access)
  - ✔ Integrate portals & workflows into NERSC environment
  - ✔ Ability to log in when system down to access data & code
  - ✔ DOE centers shouldn’t schedule downs at the same time
- Reliable resources
  - ✔ System outages and transient errors that cause job failures have a big negative impact on productivity

Questions



- How easy is it for you to translate what you need to meet your scientific objectives into quantitative computing and storage requirements?
- How can we help you do so?