NERSC: Delivering High-End Scientific Computing to the Nation's Research Community

Educause Conference

Katie Antypas
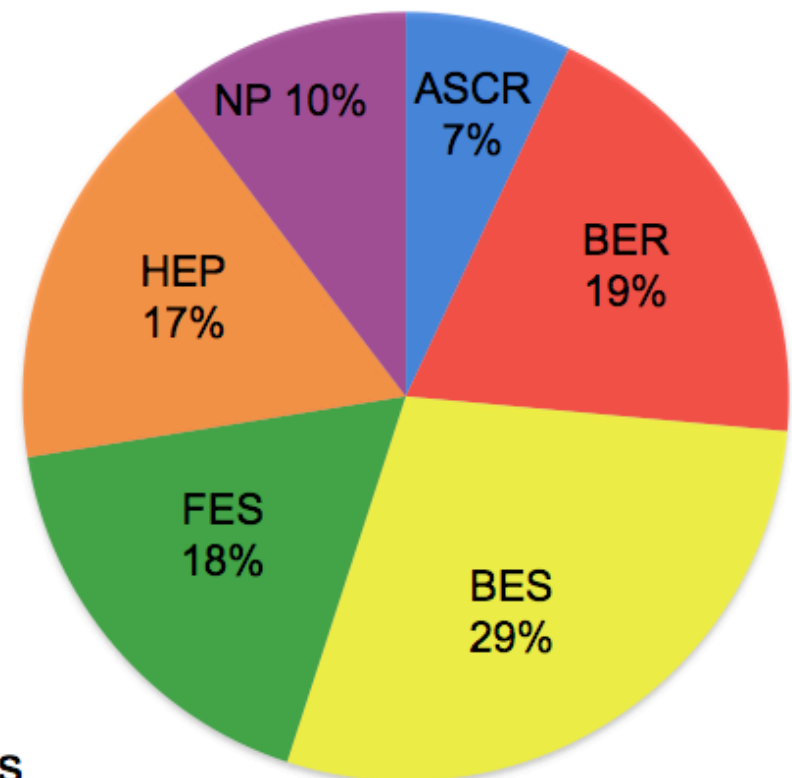**HPC Consultant**

Denver CO Nov 5, 2009

NERSC is the Primary Computing Facility for the Office of Science
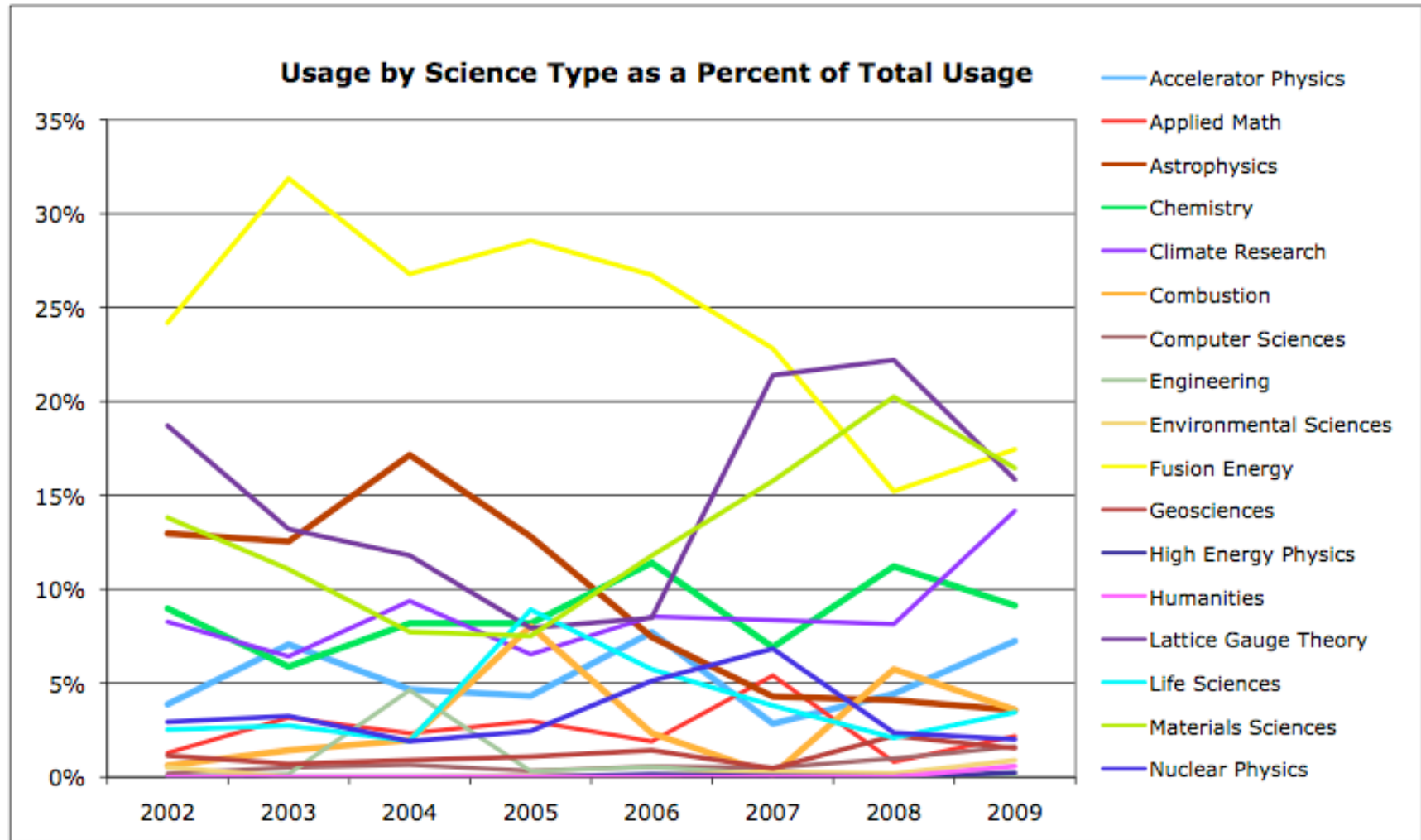
- NERSC serves a large population

  Approximately 3000 users,

  400 projects, 500 code instances

- Focus on "unique" resources

  - High end computing systems
  - High end storage systems
    - File system and tape archive
  - Interface to high speed networking

- Science-driven

  - Science problems used in machine procurements and performance metrics
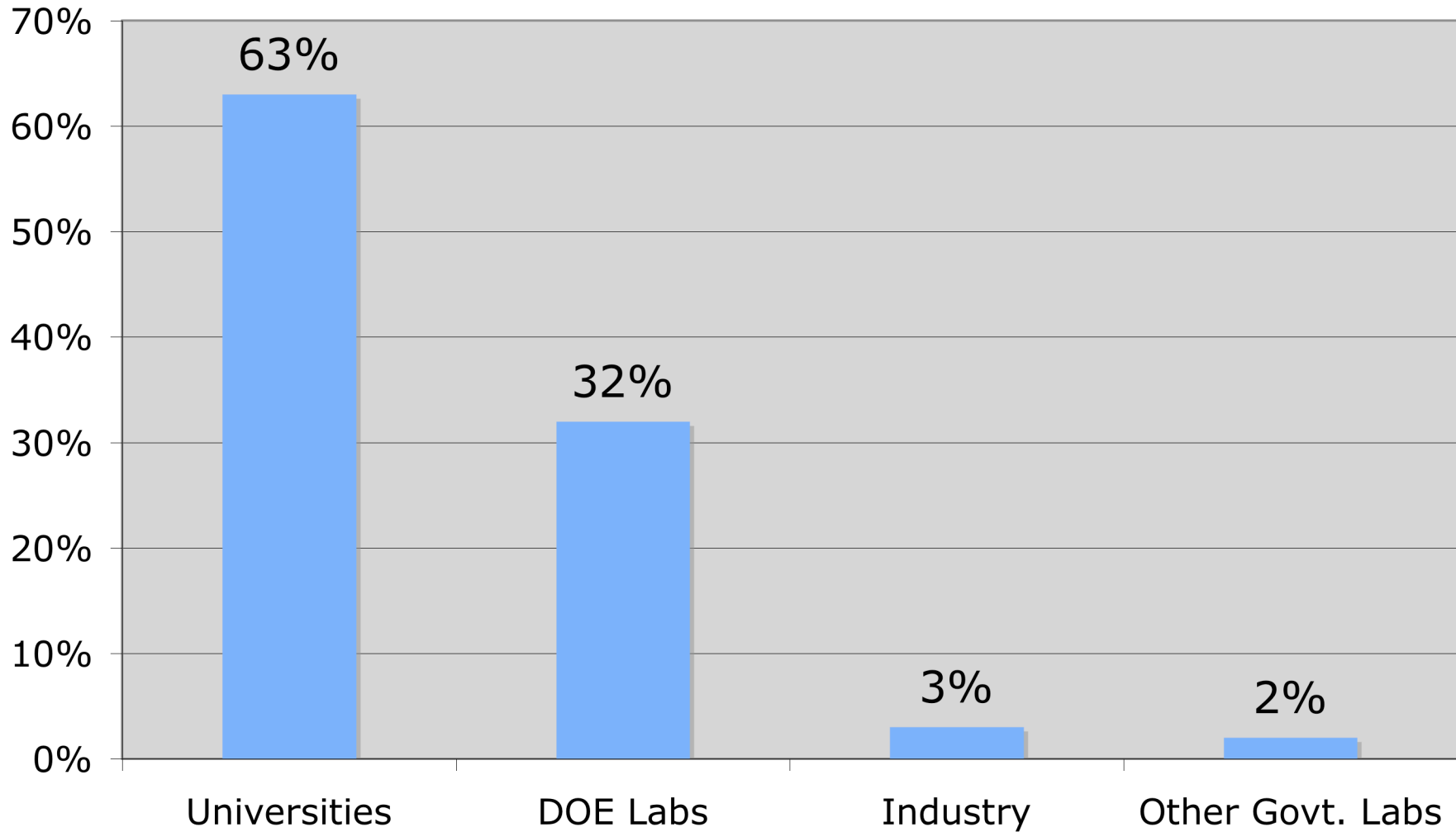 - Science services

2009 Allocations

Pie chart:
- ASCR 7%
- BER 19%
- BES 29%
- FES 18%
- HEP 17%
- NP 10%

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

What's Changed in DOE Priorities for NERSC?

Usage by Science Type as a Percent of Total Usage

Legend:
- Accelerator Physics
- Applied Math
- Astrophysics
- Chemistry
- Climate Research
- Combustion
- Computer Sciences
- Engineering
- Environmental Sciences
- Fusion Energy
- Geosciences
- High Energy Physics
- Humanities
- Lattice Gauge Theory
- Life Sciences
- Materials Sciences
- Nuclear Physics

NERSC User Demographics

NERSC 2009 Configuration

Large-Scale Computing System

Franklin (NERSC-5): Cray XT4

- 9,532 compute nodes; 38,128 cores
- ~25 Tflop/s on applications; 356 Tflop/s peak

Hopper (NERSC-6): Cray XT

- Phase 1: Cray XT5, 668 nodes, 5344 cores
- Phase 2: > 1 Pflop/s peak

Clusters

Jacquard and Bassi

- LNXI and IBM clusters
- Upgrading to Carver (NCS-c)

PDSF (HEP/NP)

- Linux cluster (~1K cores)

NERSC Global Filesystem (NGF)

Uses IBM's GPFS
440 TB; 5.5 GB/s

HPSS Archival Storage
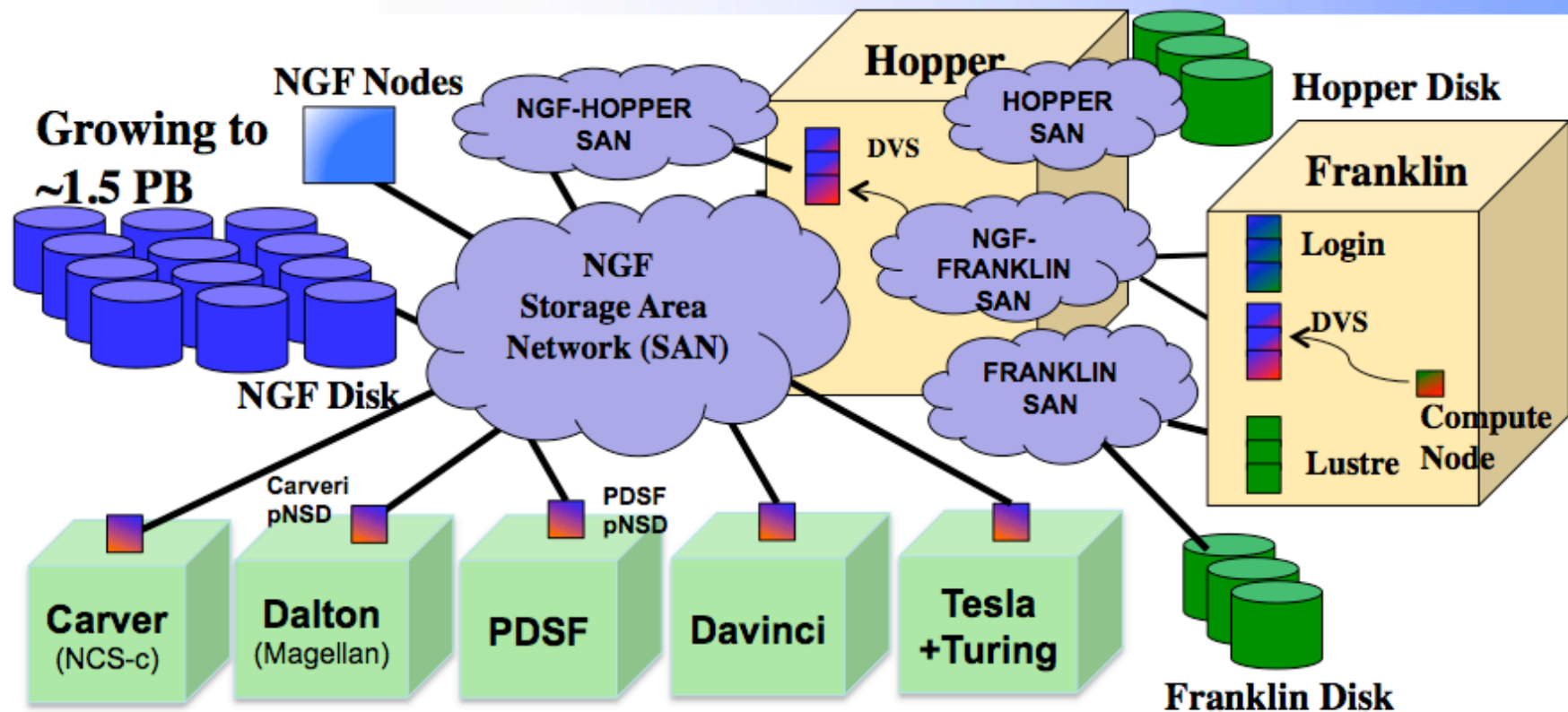
- 59 PB capacity
- 11 Tape libraries
- 140 TB disk cache

Analytics / Visualization

Davinci (SGI Altix)

- Tesla testbed
- Upgrade planned

NERSC Global File system (NGF)

- **Coming soon to NGF**
  - Additional storage, up to ~1.5 PB total
  - Access to NGF from new systems: Carver (replacing Jacquard and Bassi); Dalton (the Magellan testbed); Tesla & Turing (GPU testbed)

Focus on Scientific Productivity

- **Let scientists focus on science and research**

- **Configure and document systems for usability**

- **Wide array of 3 party software application support**

  - **Math libraries - ACML, FFTW, gsl, LibSci, PETSc, SuperLU and more**

  - **I/O - HDF5, nco, netCDF, pNetCDF**

  - **Chemistry/Mat Sci - amber, NAMD, NWChem, abinit, cpmd, lammps, quantum expresso, VASP, and more**

  - **Visualization - IDL, gnuplot, VisIT, ncar**

  - **Debuggers - Allinea's DDT and Totalview**

Efficiencies and Economies of Scale for Computing Centers

- **Large resources with many users allows systems to run more efficiently**

- **Hardware doesn't run by itself**
  - **System administrators**
  - **File system experts**
  - **Network engineers**
  - **Security analysts**
  - **Application specialists**

- **Scientists compute in 'bursty' sessions**
  - **May not compute for months at a time**
  - **Then need to surge to use the entire system**

Communicating with Key Stakeholders

- ## DOE Science Offices

- ## NERSC Users

  - ### Requirements workshops

  - ### Science disciplines have varying computing needs

    - **Bioinfomatics needs surge capability**

    - **Climate needs guaranteed throughput**

    - **Astrophysics needs high I/O bandwidth and disk storage**

Opportunities for Collaboration with Campus Computing

- **Always looking for new science to run at NERSC machines**
- **Training**
  - **Our users are spread across the country**
  - **Localized training**
- **Bug tracking**
- **System configuration settings**
  - **MPI Buffers, Lustre settings, networking, etc.**
- **Software installations**
  - **Optimal compiler settings**
  - **Trouble shooting**

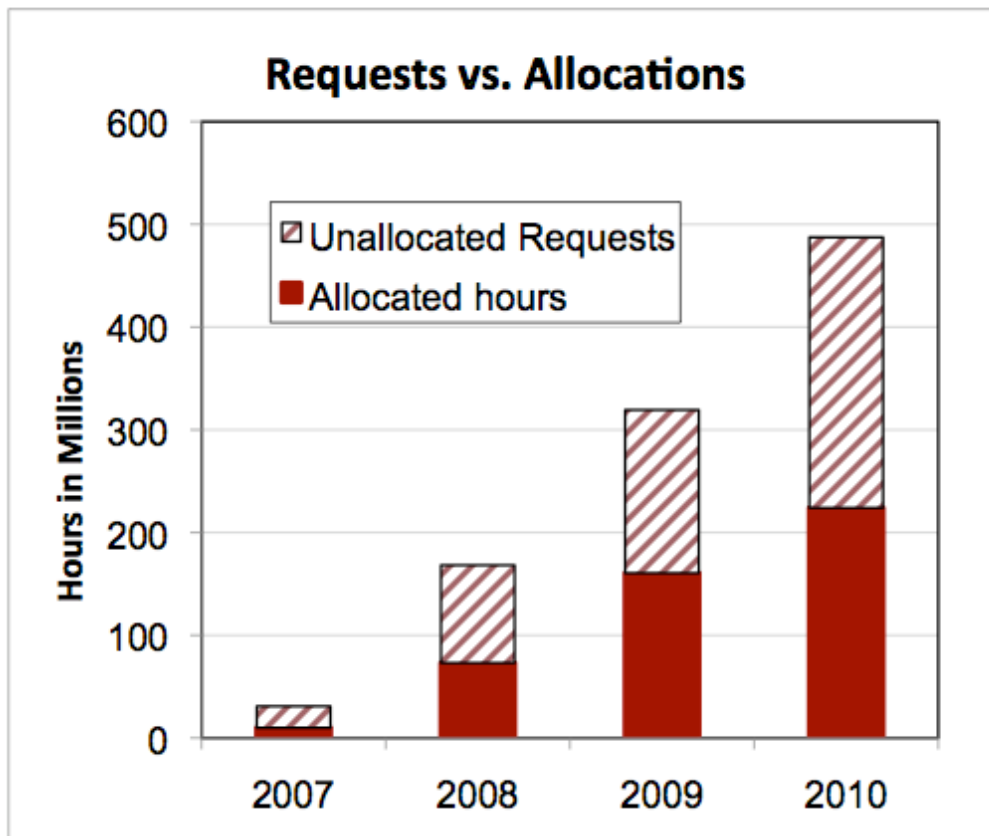# DOE Explores Cloud Computing

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

Demand for More Computing

Compute Hours Requested vs Allocated

Requests vs. Allocations

Chart legend:
- Unallocated Requests
- Allocated hours

Y-axis: Hours in Millions (0, 100, 200, 300, 400, 500, 600)
X-axis: 2007, 2008, 2009, 2010

- **Each year DOE users requests ~2x as many hours as can be allocated**

- **This 2x is artificially constrained by perceived availability**

- **Unfulfilled allocation requests amount to hundreds of millions of compute hours in 2010**

NERSC Allocations

- **Allocations**
  - 80% DOE program managers control
  - 10% ASCR Leadership Computing Challenge
  - 10% NERSC Reserve
- **Science includes all of DOE Science Offices**
- **Start-up allocations available directly from NERSC**
  - 10,000 - 50,000 hours allocations
  - If you have an abstract of your research goals applying will take about 30 min
  - A small allocation is stepping stone toward a large allocation. It helps build a computing relationship with DOE and project reviewers.
- **http://www.nersc.gov/nusers/accounts**

Cover Stories from NERSC Research
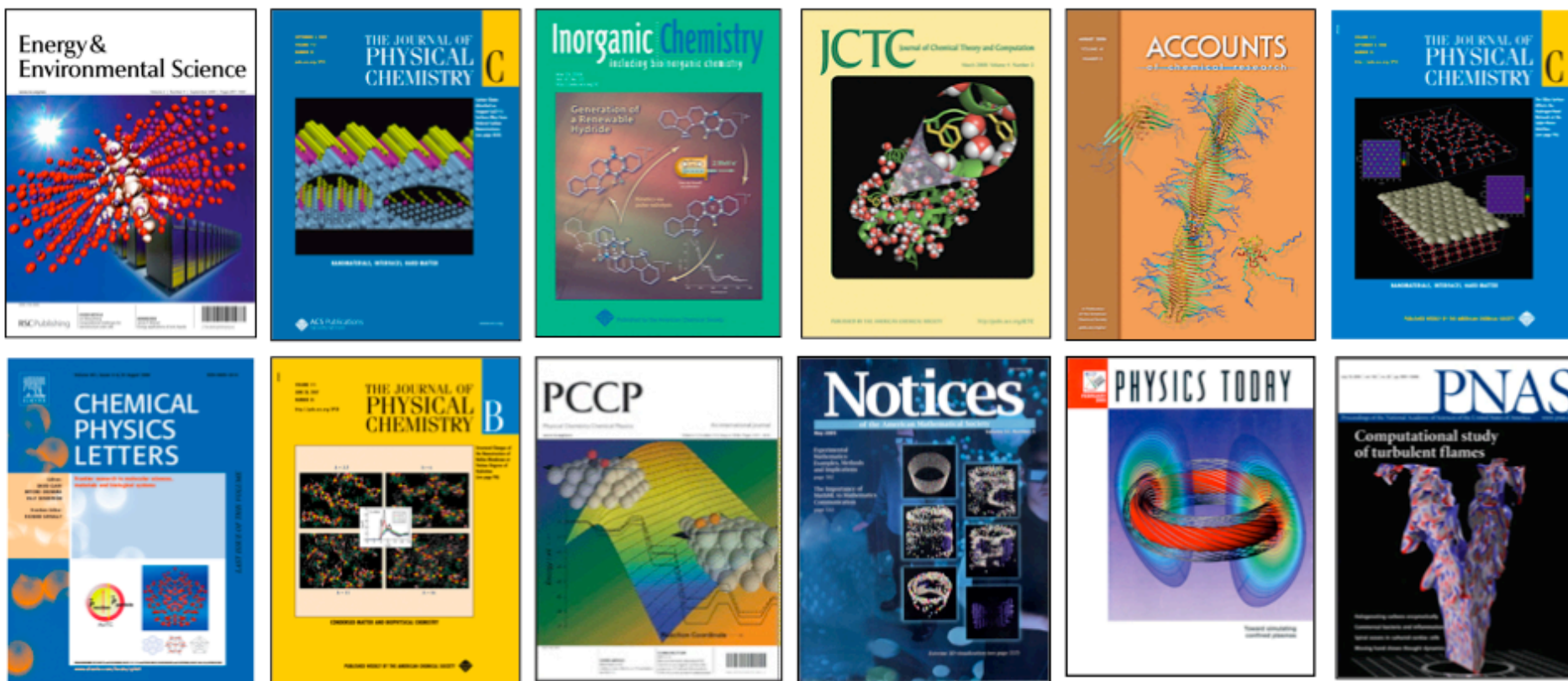
NERSC is enabling new science in all disciplines, with about *1,500 refereed publications* per year

NERSC Mission

The mission of the National Energy Research Scientific Computing Center (NERSC) is to *accelerate the pace of scientific discovery* by providing high performance computing, information, data, and communications services for *all DOE Office of Science (SC) research.*