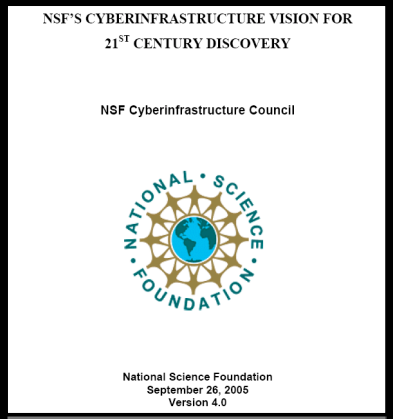
NSF's High-End Computing Activities

Steve Meacham
Sr. Science and Technology Advisor
Office of Cyberinfrastructure
National Science Foundation

SC08
November 2008

Strategic Plan (FY 2006 – 2010)

- Ch. 2: High Performance Computing



5yr Goal:

- To enable petascale science & engineering thro' deployment & support of a world-class HPC environment comprising the most capable combination of HPC assets available to the academic community

Components:

- Acquisition, deployment, operation of science-driven HEC systems
- Development & maintenance of supporting software
  - New design tools, performance modeling tools, system software & fundamental algorithms
- Development and maintenance of portable, scalable applications software

HEC Program Elements

- **Acquisitions**
  - Track 1 - Petascale
  - Track 2 - Mid-range supercomputers
- **Operations**
- **HEC System Software**
  - Compilers, fault-tolerant OS, fault-survivability tools, system status monitoring, file-systems, PSEs, …
- **HEC Petascale Application Development**
  - Scalable math libraries, scalable algorithms, data exploration tools, performance profiling and prediction, large application development
- **Coordinated with other agencies**

Recent High-end HPC Investments

Petascale Software Development (PetaApps, SDCI and STCI) 2006...

2007- 2011- 2016 Track 1 Petascale Acquisition (UIUC/NCSA)

TeraGrid Operations
TACC, NCSA, SDSC, PSC, ORNL, ANL, Indiana, Purdue, NCAR, LSU

TeraGrid Phase III (XD)

Track 2

Track 2A: 2006-2011 TACC

Track 2B: 2007-2012 UTK

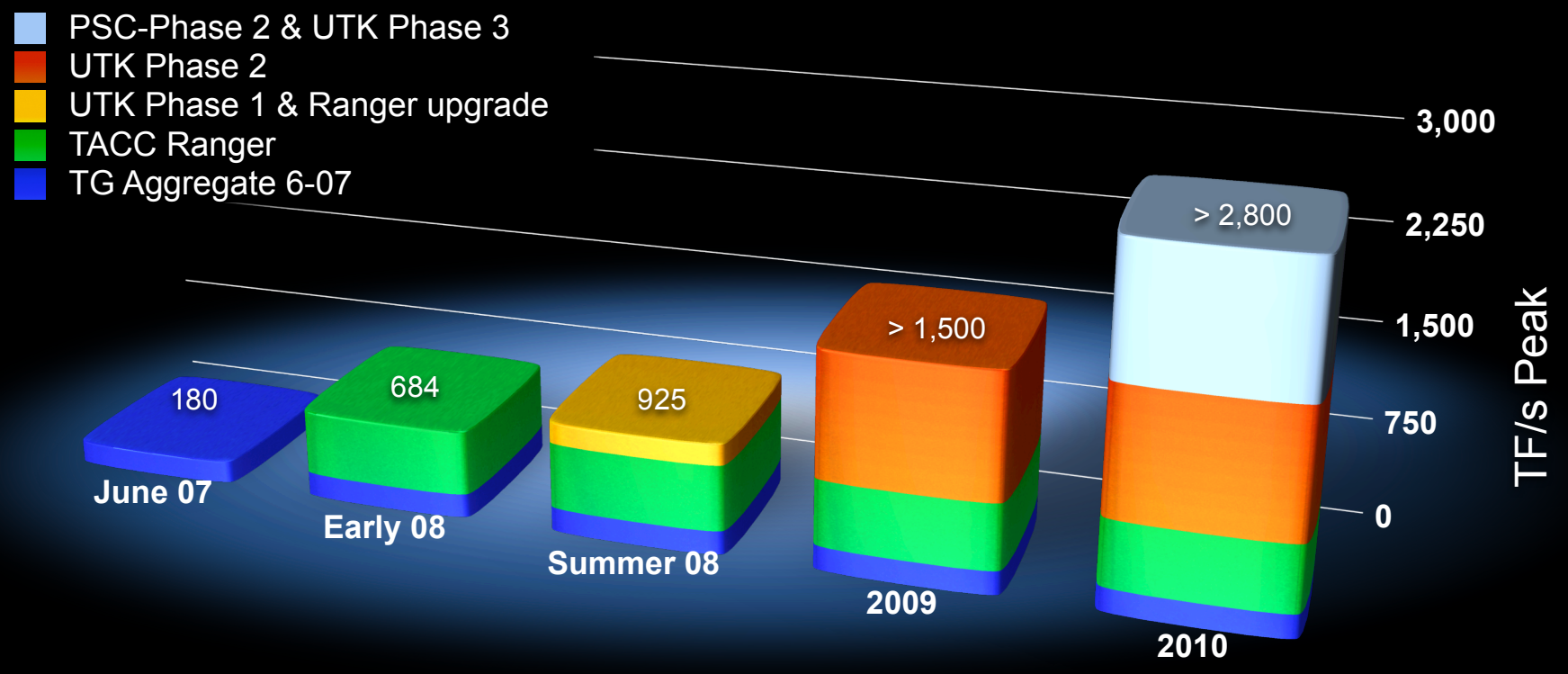
Track 2C: 2008-2013 PSC

Track 2D: 2009-2014 Multiple smaller resources

TACC - Texas Advanced Computing Center, Austin.
UTK - Univ. of Tennessee, Knoxville, Joint Institute of Computational Science.
NCSA - National Center for Supercomputer Applications, Univ. of Illinois, Urbana-Champaign.
SDSC - San Diego Supercomputer Center, Univ. of California, San Diego.
PSC - Pittsburgh Supercomputing Center

Impact

Greatly expanding capacity of the TeraGrid for digital exploration with reduced oversubscription and queue wait times.

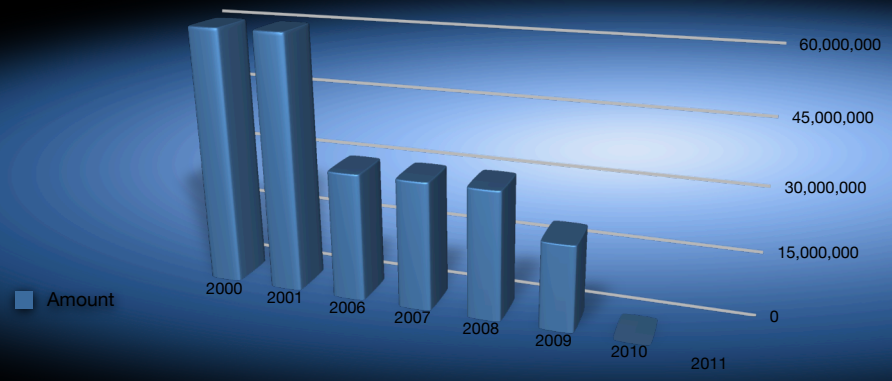


National Science Foundation
Where Discoveries Begin

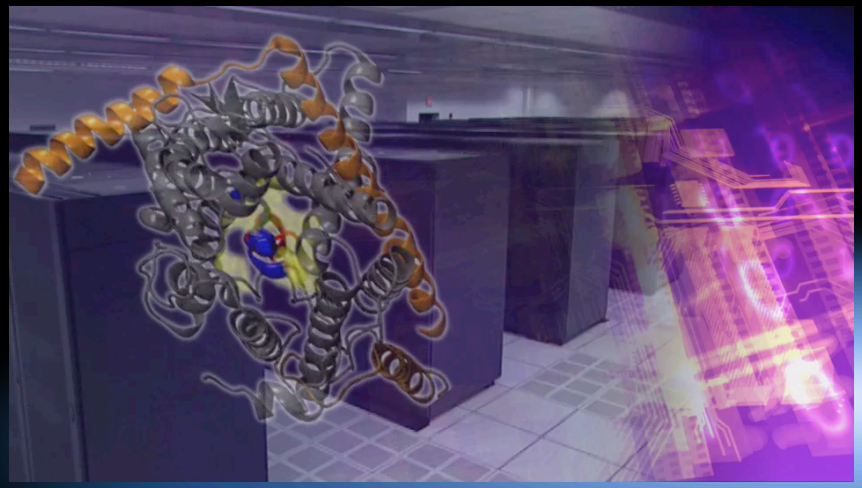
Office of Cyberinfrastructure

Other Trends

- Growing power dissipation
- Complexity of system
- Complexity of programming
- Increase in uptake
- Declining investment



Blue Waters: the next-generation supercomputer



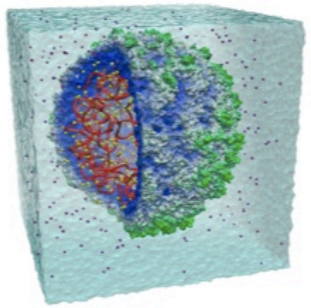
- climate science, particle physics, materials science, stellar evolution
- fluid dynamics, astrophysics and cosmology, condensed matter physics
- cell biology, nano-engineering, chemistry, plasma physics
- the influence of social networks on the spread of contagion

National Science Foundation
Where Discoveries Begin

Office of Cyberinfrastructure

HPC is an increasingly important tool for understanding:

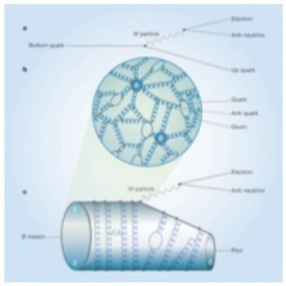
Life



Satellite tobacco mosaic virus, P. Freddolino et al.

Aldehyde dehydrogenase, T. Wymore and S. Brown

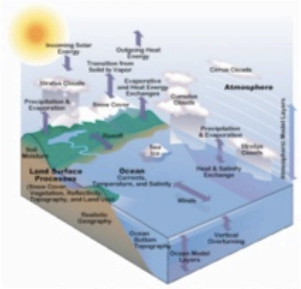
Matter



I. Shipsey

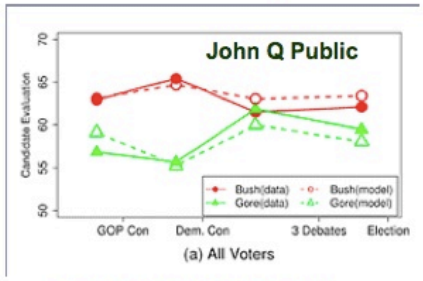
The Environment

K. Droegemeier et al.



Community Climate System Model

Society



S.-Y. Kim, M. Lodge, C. Taber.

D. E. Atkins

Some Science We'll Be Able To Do

Pandemic Network Simulation and Analysis: Models for simulating the spread of infectious diseases using graph-theoretic methods. The influence of social networks. In silico testing of mitigation strategies.

Computational Proteomics: Modeling proteins with ~1000 amino acids to examine protein-protein interactions and trans-membrane signaling. Inverse methods for macromolecular design.

Environmental Decomposition of Complex Molecules: Develop highly scalable, locally correlated, numerical quantum chemistry methods for molecules containing hundreds of heavy atoms. Study decomposition pathways of complex molecules in the environment.

Ocean Modeling: Study ice-ocean interactions using whole-globe coupled ice-ocean models.

Cosmology: Multi-scale cosmological modeling in support of research efforts in precision cosmology, including galaxy surveys, weak lensing, and the analysis of the cosmic microwave background.

Computational Fluid Dynamics: Study the evolution of transient, localized structures in inhomogeneous turbulence using daptive mesh refinement (AMR).

Mining the Internet Archive: Studying the propagation of ideas and interactions.

Analysis of Socio-Economic data: Extraction of patterns and pattern evolution in multi-year socio-economic data collections.

Track 2d

- Track 2 = TeraGrid hardware funding line
- Track 2d = up to $20M in year 1 & constrained O&M
- Looking for innovation
- Up to four awards - two “production” and two “experimental” categories
  - A data-intensive, high-performance computing system
  - An experimental, high-performance computing system of innovative design
  - An experimental, high-performance grid test-bed for grid research (CS & applied)
  - A pool of loosely coupled grid-computing resources
- All “systems” to be made available to national community of users via the TeraGrid

Some examples of existing programs

NSF's investments in HEC R&D are spread over several offices and divisions

- HECURA
- Science and Engineering Beyond Moore's Law - proposed new effort, led by MPS

  Focuses on new paradigms for HW architecture, algorithms, & application software

  E.G. basic research on new types of logic devices: nanophotonics, spintronics, molecular computing, quantum computing
- STCI & SDCI
- Multicore Chip Design and Architecture (NSF and SRC)

  Innovative research on design, fabrication, architecture and programmability of homogeneous and heterogeneous multicore systems. - Includes hardware and software aspects
- PetaApps

  Funding for development of future simulation and analysis tools that can use petascale computing to advance the frontiers of scientific and engineering research.
- Application development in domain science and engineering programs. e.g. climate, nanoscale engineering, biosciences, chemistry, SBE
- National Center for Atmospheric Research

The next five years 2011 - 2016

- Over next year - develop a new five-year plan
- Looking for input on opportunities and priorities
  - From academia and industry
  - From consumers of leading-edge HPC and from researchers and developers of leading-edge HPC
- Anticipate greater emphasis on software development
- Over the current decade, NSF investments in HPC acquisition have been declining; future funding levels are uncertain
- NSF seeks to maintain a balanced cyberinfrastructure portfolio. Other CI areas include: data curation and data management, networking, middleware development, grid/cloud computing
- Exascale and beyond in computing and data are hard challenges
  - hope to expand partnerships with other agencies and foster international collaborations
- No investment similar to Track 2 is planned for 2010

Questions

- What are some of the research areas - hardware & software - which academia & industry would be interested in tackling together?
- What are the science grand challenges that would require the combination of international HPC resources to tackle?
- How does the cloud paradigm fit into the provisioning of high-performance computing resources?

Questions or Comments?