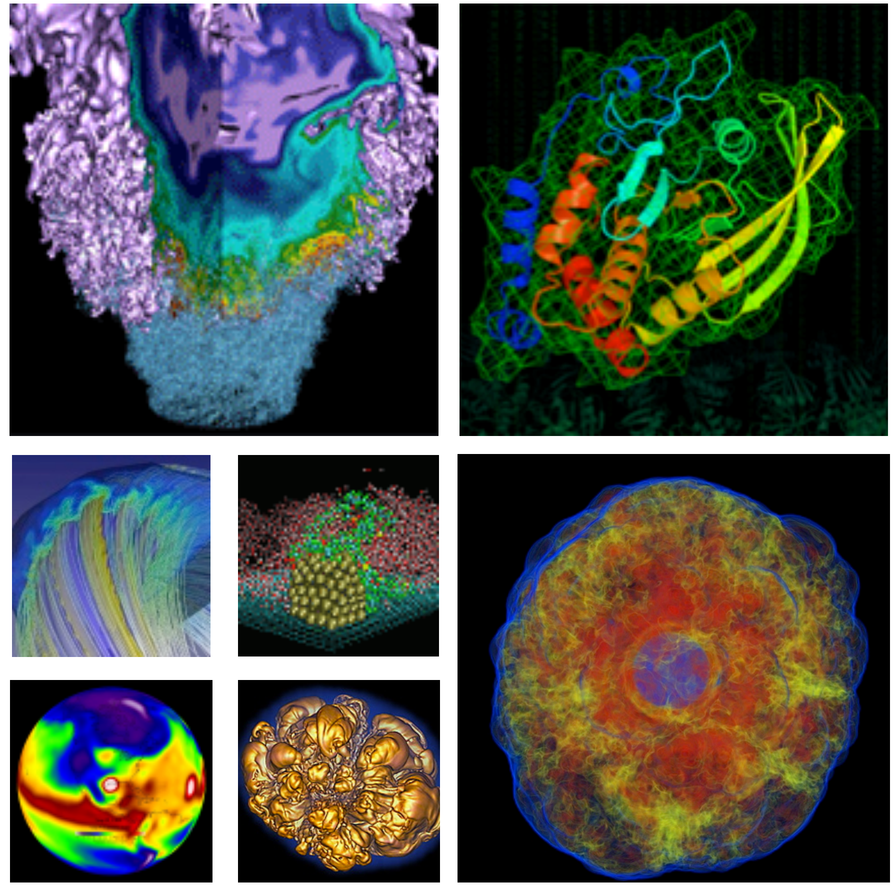
Introduction to High Performance Computing (HPC)



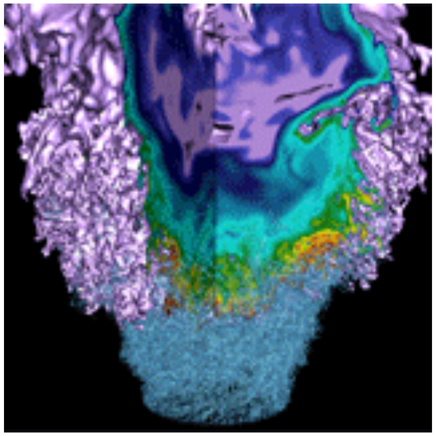
Richard Gerber
Senior Science Advisor
Deputy Group Lead, NERSC User Services

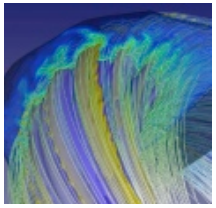
June 11, 2013

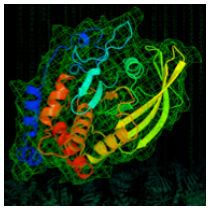
Outline

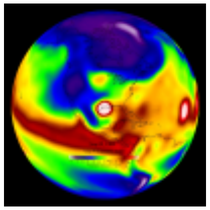
- **What is HPC and who uses it?**
- **What is a supercomputer?**
- **Who uses HPC?**
- **What is NERSC?**
- **What do scientists do at NERSC?**
- **How does HPC computing work?**
- **Challenges in HPC?**

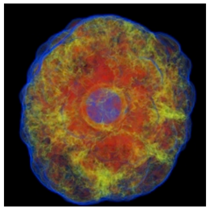
What is HPC and Who Uses It?

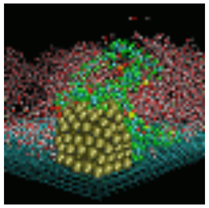






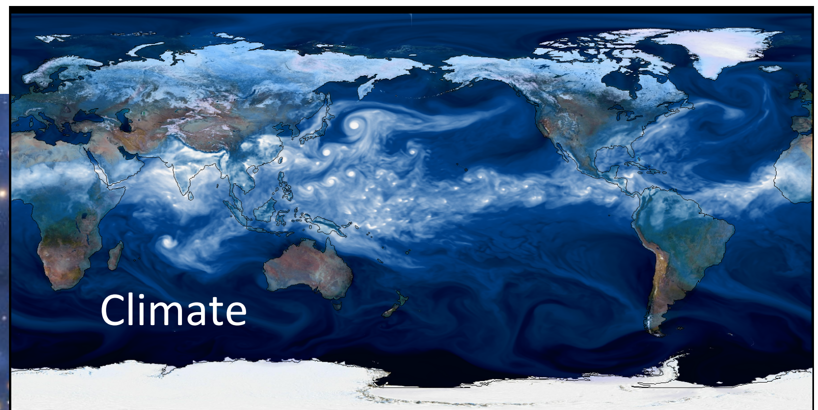




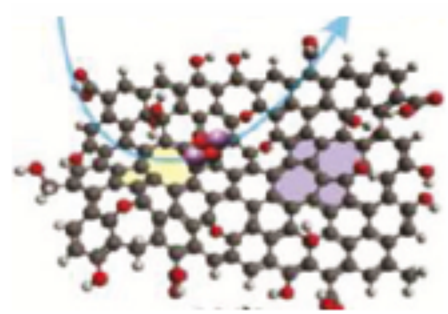


High Performance Computing is …

… the application of "supercomputers" to scientific computational problems that are either too large for standard computers or would take them too long.



Understanding How Proteins Work



Designing Better Batteries

Why Use HPC?

Length (m)	Phenomena
10^{-18}-10^{-15}	quarks, strings
10^{-15}-10^{-12}	proton, neutron
10^{-12}-10^{-9}	gamma rays, X rays, hydrogen atom
10^{-9}-10^{-6}	DNA, virus, optical light
10^{-6}-10^{-3}	bacteria, fog, human hair
10^{-3}-10^{0}	mosquito, golf ball, football
10^{0}-10^{3}	people, football field, Eiffel tower
10^{3}-10^{6}	Mt. Everest, Panama Canal, asteroid
10^{6}-10^{9}	Moon, Earth, light-second
10^{9}-10^{12}	Sun, light-minute, Earth's orbit
10^{12}-10^{15}	Solar System
10^{15}-10^{18}	light-year, nearest star
10^{18}-10^{21}	galactic arm
10^{21}-10^{24}	Milky Way, distance to Andromeda galaxy
10^{24}-10^{26}	visible universe

Direct Human Experience (10^{-6} to 10^{6})

U.S. DEPARTMENT OF ENERGY | Office of Science

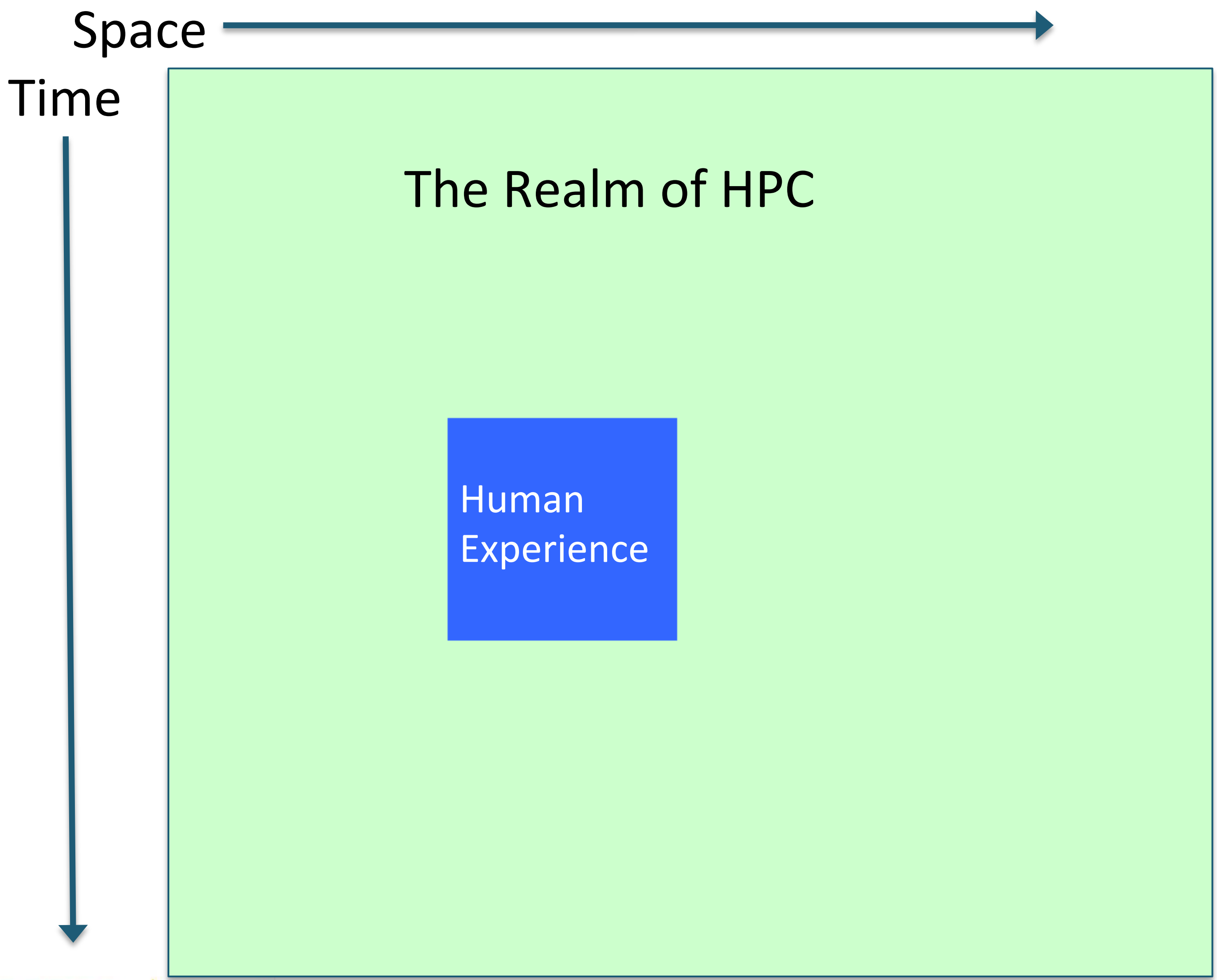
BERKELEY LAB

Why Use HPC

Time Scale (s)	Phenomena
10^{-44}	Planck time
10^{-24}	light crosses nucleus
10^{-15}	atomic vibration, visible light
10^{-12}	IBM SiGe transistor
10^{-9}	1 Gz CPU
10^{-6}	protein folding, lightning bolt
10^{-3}	hard disk seek time, blink of an eye
10^{0}	earthquakes
10^{2}	tornadoes
10^{5}	hurricanes
10^{7}	year
10^{9}	human life span
10^{10}	deep ocean mixing time
10^{12}	first homo sapiens
10^{15}	Milky Way rotation period
10^{17}	age of universe

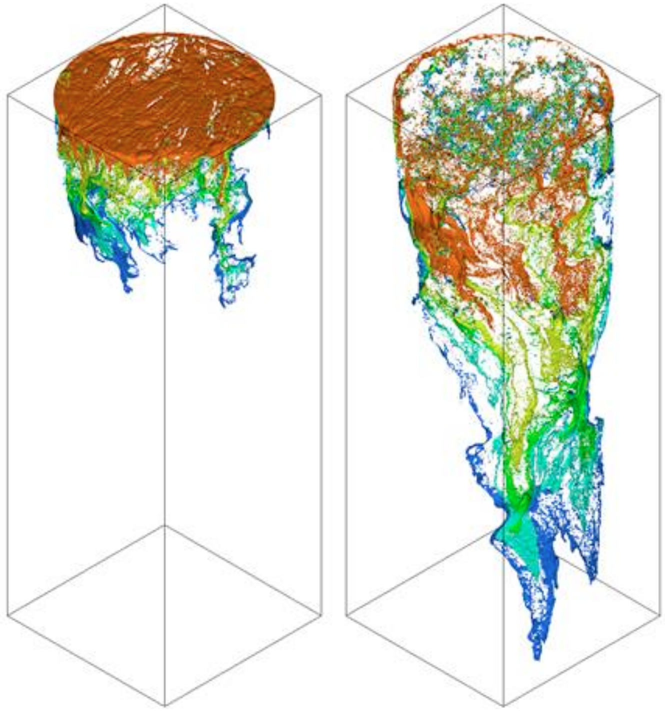
Direct Human Experience (10^{-3} to 10^{9})

HPC: Access to the Universe Past, Present, & Future

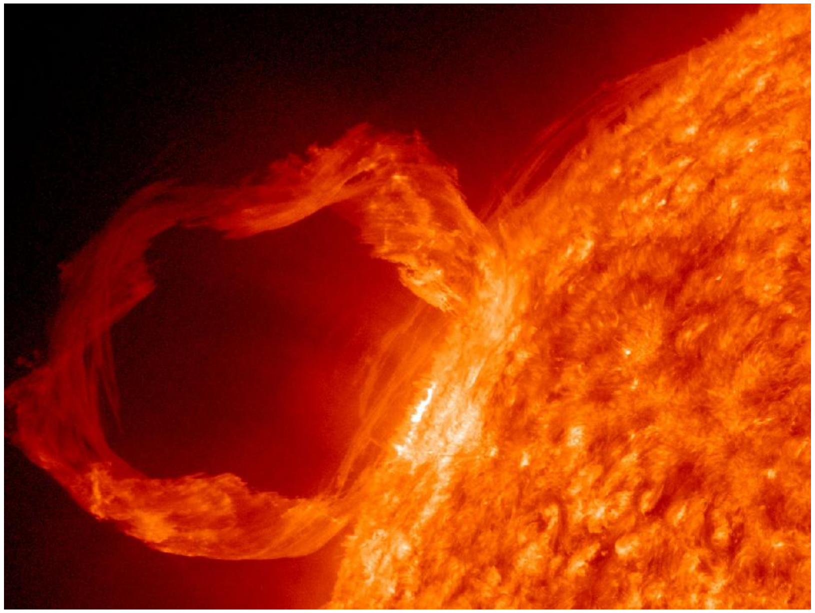


Why Use HPC?

- **Explore dangerous or inaccessible domains**



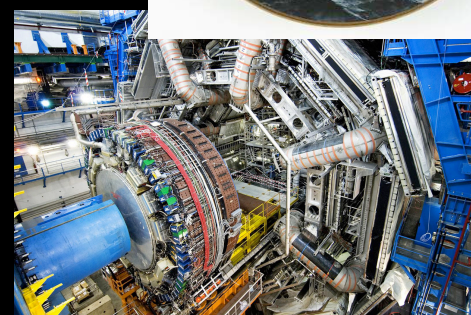
Biological and radioactive flow of materials through porous rock



Magnetic storms on the Sun

HPC is a Tool for Discovery

NERSC



HPC are used by …

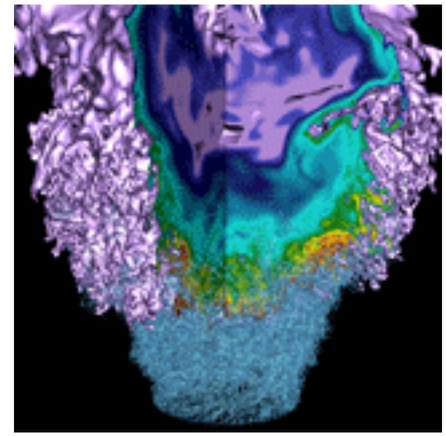
Scientists and Engineers

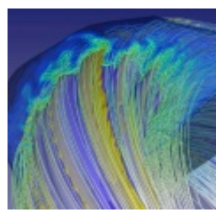
- Environment
- Biosciences
- Chemistry
- Materials & Nanotech
- Geology
- Physics
- Mechanical Engineering
- Electrical Engineering
- Alternative Energy
- Military

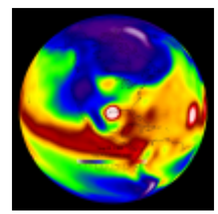
Industry and Commercial

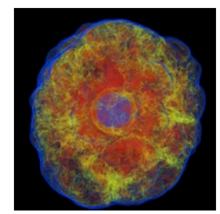
- Energy Exploration
- Automotive Design
- Aerospace
- Medial Imaging
- Data Mining
- Web Search & Services
- Financial Modeling
- Advanced Graphics (think Pixar)

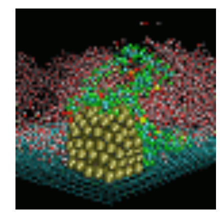
What is a Supercomputer?











A Supercomputer is …

… not so different from a super high-end desktop computer.

Or rather, a lot of super high-end desktop computers.

Hopper, show above, has 6,384 “nodes” (~one desktop), each with 24 compute cores for a total of

153,216 compute cores

But There's More …

The nodes are all connected to each other with a high speed, low latency network.

This is what allows the nodes to "talk" to each other and work together to solve problems you could never solve on your laptop or even 150,000 laptops.

Typical point-to-point bandwidth

Supercomputer: 10 GBytes/sec
Your home: 0.02* GBytes/sec

Latency

Supercomputer: 1 µs
Your home computer: 20,000* µs

* If you're really lucky

Cloud systems have slower networks

And Even More …

PBs of fast storage for files and data

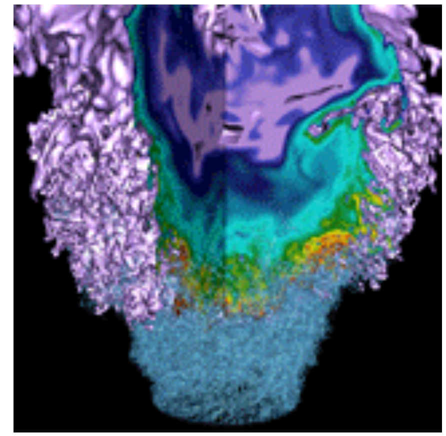
Hopper: 2 PB
Your laptop: 0.0005 PB
Your iPhone: 0.00005 PB

Write data to permanent storage

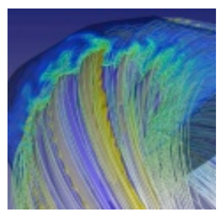
Hopper: 70 GB/sec
My iMac: 0.01 GB/sec

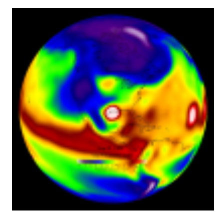
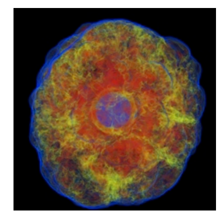
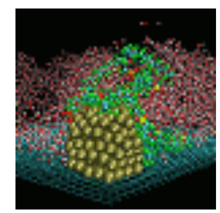
Cloud systems have slower I/O and less permanent storage

What is NERSC?



NERSC



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

National Energy Research Scientific Computing Center (NERSC)

- **NERSC is a national supercomputer center funded by the U.S. Department of Energy Office of Science (SC)**
  - Supports SC research mission
  - Located at Berkeley Lab
- **SC supports a broad spectrum of energy-related research**
- **If you have SC funding, you can use NERSC**
  - Other researchers can apply if research is in SC mission
- **4,500 users, 700 projects**
- **From 48 states; 65% from universities**
- **Hundreds of users each day**

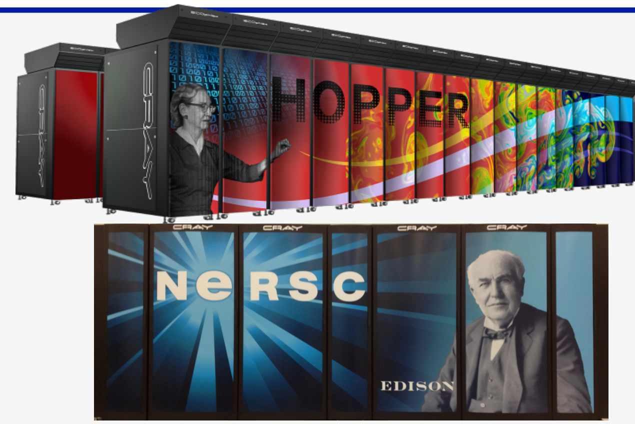
NERSC's focus is on enabling scientific productivity

13 Journal Covers in 2012

- 1,500 refereed publications per year
- ~10 major journal covers per year
- Key contributor to 2 Nobel Prizes (2007 & 2011)
  - Accelerating Universe
  - Climate Change
- Data services contributed to 2 of Science Magazine's Top 10 breakthroughs of 2012
  - Higgs Boson
  - Weak Mixing Angle

Current NERSC Systems

World-Class Supercomputers



Hopper: Cray XE6
- 6,384 compute nodes, 153,216 cores
- 144 Tflop/s on applications; 1.3 Pflop/s peak

Edison: Cray XC30 (Cascade)
- Phase I (10K processors), Phase II in 2013 (~120K)
- Over 200 Tflop/s on applications, 2 Pflop/s peak

Midrange

140 Tflops total

Carver
- IBM iDataplex cluster
- 9884 cores; 106TF

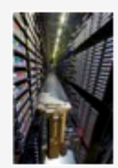
PDSF (HEP/NP)
- ~2K core cluster

GenePool (JGI Genomics)
- ~5K core cluster
- 2.1 PB Isilon File System

NERSC Global Filesystem (NGF)

Uses IBM's GPFS
- 8.5 PB capacity
- 15 GB/s of bandwidth

HPSS Archival Storage



- 240 PB capacity
- 5 Tape libraries
- 200 TB disk cache

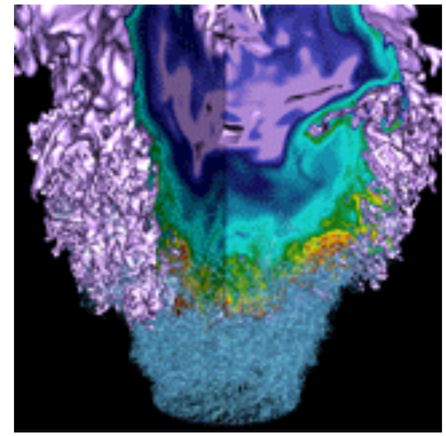
Analytics & Testbeds

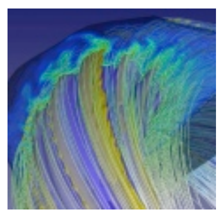
Dirac 48 Fermi GPU nodes

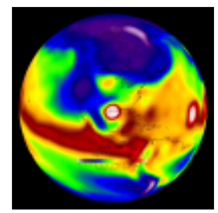
U.S. DEPARTMENT OF ENERGY | Office of Science

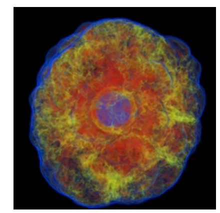
BERKELEY LAB

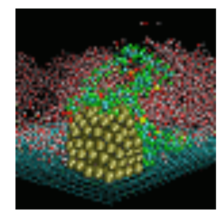
What Do Scientists Do at NERSC?







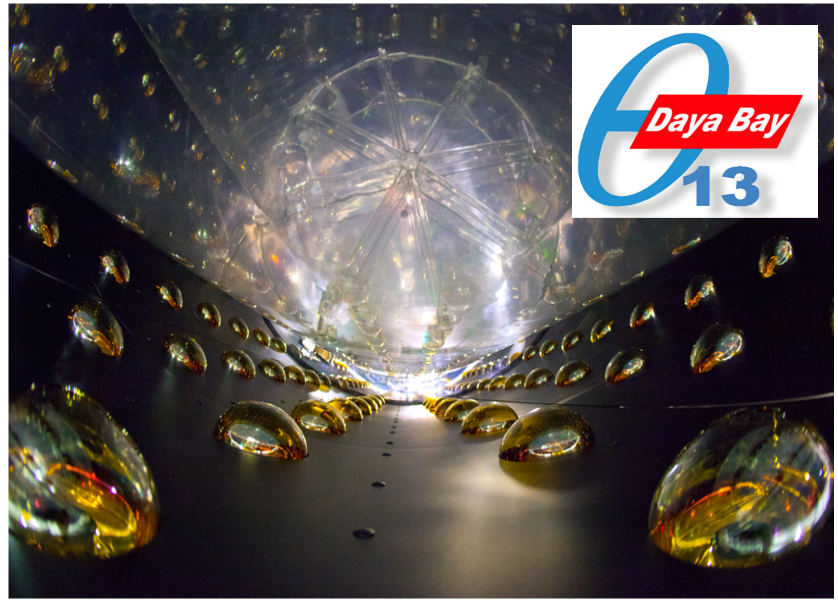




One of *Science* Magazine's Top 10 Breakthroughs of 2012

Discovery of θ_{13} weak mixing angle

- The last and most elusive piece of a longstanding puzzle: How can neutrinos appear to vanish as they travel?
- The answer – a new, large type of neutrino oscillation
  - Affords new understanding of fundamental physics
  - May help solve the riddle of matter-antimatter asymmetry in the universe.



Detectors count antineutrinos near the Daya Bay nuclear reactor in China. By calculating how many would be seen if there were no oscillation and comparing to measurements, a 6.0% rate deficit provides clear evidence of the new transformation.

Experiment Could Not Have Been Done Without NERSC and ESNet

- PDSF for simulation and analysis
- HPSS for archiving and ingesting data
- ESNet for data transfer into NERSC
- NERSC Global File System & Science Gateways for distributing results
- NERSC is the *only* US site where all raw, simulated, and derived data are analyzed and archived

PI: Kam-Biu Luk (LBNL)

NERSC Played Key Role in Nobel Prize-Winning Discovery


Physics

Accelerating Expansion of the Universe Subject of 2011 Prize

Type Ia supernovae are used as "standard candles" to measure the distance to remote galaxies.

Simulations run at NERSC modeled how Type Ia supernovae should appear from Earth.

This provided the crucial calibration needed to enable the Nobel Prize-winning discovery.

When NERSC moved to Berkeley 1996, this project's work was one of the first funded in a new computational science program created to encourage collaborations between physical and computer scientists.

Berkeley Lab's Saul Perlmutter was awarded the 2011 Nobel Prize in Physics along with two others for their discovery.

It implies the existence of so-called dark energy, a mysterious force that acts to oppose gravity.

The nature of dark energy is unknown and has been termed the most important problem facing 21st century physics.

NERSC Award for High Impact Scientific Achievement 2013

NERSC

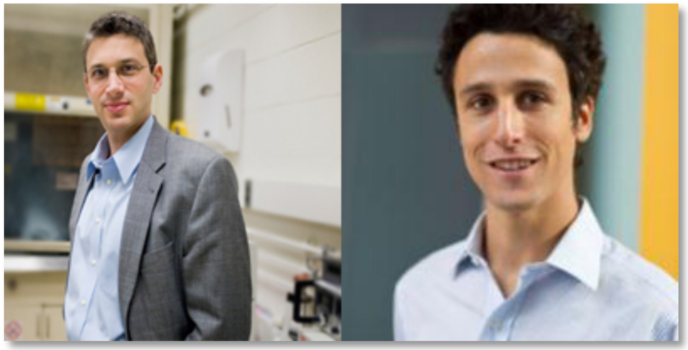
A New Approach to Water Desalination

Jeff Grossman and David Cohen-Tanugi, MIT

New material's water permeability is several orders of magnitude greater than conventional membranes.

Using supercomputers at NERSC, Grossman and Cohen-Tanugi came up with a new approach for desalinating seawater using sheets of graphene, a one-atom-thick form of the element carbon. This method holds the promise of being far more efficient and less expensive than existing desalination systems.

Grossman's project "Quantum Simulations of Nanoscale Energy Conversion" has used 5.6 Million hours at NERSC since 2010.



Jeff Grossman (left) and David Cohen-Tanugi are the recipients of the inaugural NERSC award.

One of Smithsonian Magazine's Top Ten "Surprising Scientific Milestones of 2012"

Early Career Award: Fundamental Properties of Novel Superconductors

Tanmoy Das, Las Alamos National Laboratory

Das completed groundbreaking computational work to understand fundamental properties of novel superconductors and spin-orbit ordering effects in two-dimensional electron gases. A postdoctoral researcher at LANL, Das was the first author on three 2012 articles published in the highly regarded journal Physical Review Letters.

NERSC Award for Innovative Use of High Performance Computing

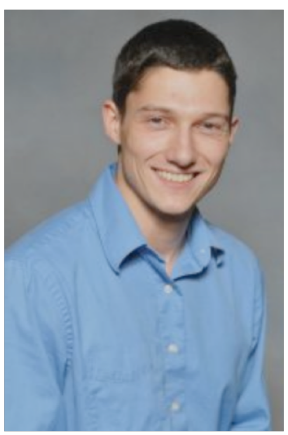
Data Pipeline Transfers, Analyzes, Stores, & Disseminates Astronomical Observations

Peter Nugent and the PTF Team, Berkeley Lab & UC Berkeley, California Institute of Technology

Innovative workflow enables earliest supernova discovery, first direct observations of progenitor systems.

Every night observations from the Palomar Observatory in Southern California are sent to NERSC where computers running machine learning algorithms scour the data for transients. Once an interesting event is discovered, an automated system sends its coordinates to ground-based telescopes around the world for follow-up observations. NERSC also archives this data and allows collaborators to access it over the Internet through a web-based science gateway.

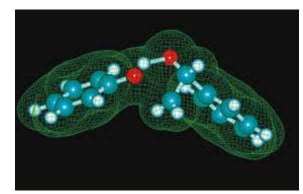
Peter Nugent, the PTF's Realtime Transient Detection Lead, is interviewed on the PBS News Hour following the discovery of supernova PTF 11kly within hours of its appearance.



Early Career Award

Edgar Solomonik, UC Berkeley

$$\Lambda^{\alpha}{}_{\beta} M^{\beta}{}_{\gamma} = N^{\alpha}{}_{\gamma}$$



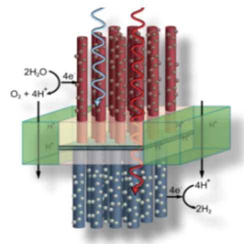
Solomonik, a graduate student at UC Berkeley, has developed novel algorithms for massively parallel tensor contractions and applied them to quantum chemistry problems. Solomonik's algorithmic developments are instantiated in the Cyclops Tensor Framework (CTF), which has been used on some of the largest supercomputers in the world, including the NERSC Hopper system, and the IBM Blue Gene/Q systems at the Lawrence Livermore National Laboratory and Argonne Leadership Computing Facility.

Basic Energy Sciences

Batteries

Breakthroughs in battery technologies may extend the range of electric cars.

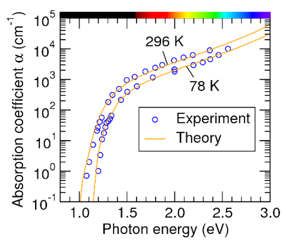
(D. Mei, PNNL, L-W. Wang, LBNL)



Artificial Photosynthesis

Simulation is playing a key role in highly visible quest to develop artificial photosynthesis.

(L. Wang, LBNL)



Solar Energy

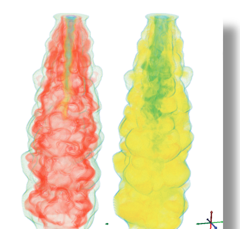
A NERSC “NISE” award and software by a NERSC consultant yield an important new method for characterizing solar energy materials.

(E. Kioupakis, U. Michigan)

Coal Gasification

NERSC resources were used to model a real coal gasifier with a Large Eddy Simulation code.

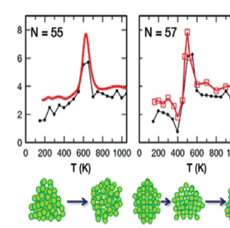
(P. Smith, U. Utah)



Nanotechnology

Computation explains the size-sensitive melting behavior of metal nanoclusters.

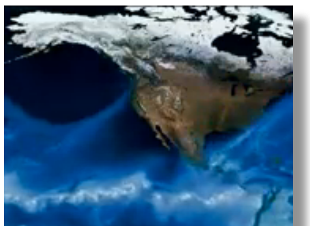
(S. Wei, NREL)



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

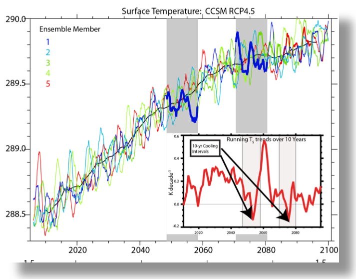
Biological and Environmental Research



Extreme Climate

New techniques help detect extreme events buried in immense data sets.

(Prabhat, M. Wehner, LBNL)



Oceanic Heat Reserviors

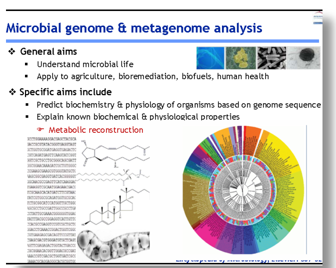
Key finding that deep oceans can mask global warming for decade-long periods

(G. Meehl, A. Hu, NCAR)

Genomes

Genomes pipeline at NERSC can process 100 million genes in few days, a task that used to require weeks at the JGI.

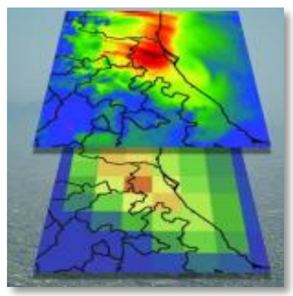
(V. Markowitz, LBNL)



Aerosol Effects

Atmospheric scientists have shown how small-scale effects of aerosols contribute to errors in climate models.

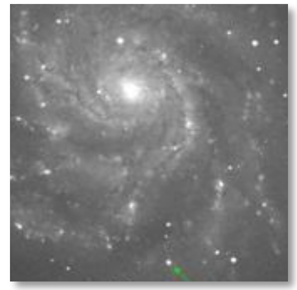
(W. Gustafson, PNNL)



High Energy Physics

Acceleration of the Universe

NERSC played a key role in the discovery that led to the 2011 Nobel Prize in Physics.
(S. Perlmutter, UC Berkeley/LBNL)

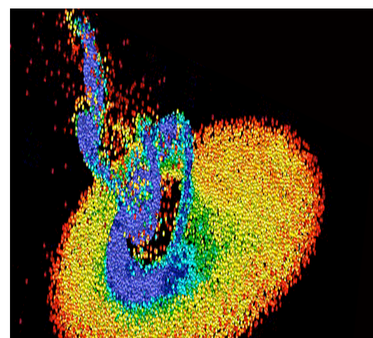


Supernova

The earliest-ever detection of a supernova was made possible by NERSC and Esnet.
(P. Nugent, LBNL)

Neutrino Decay

An important piece of the neutrino puzzle has fallen into place, thanks to data transfer, storage, analysis, archive, and gateway capabilities at NERSC and ESnet.
(K.-B. Luk, LBNL)



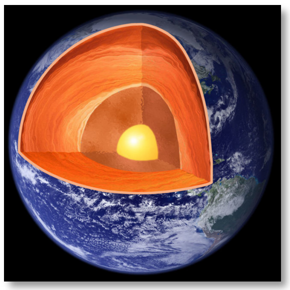
Dark Matter

Simulations done at NERSC helped validate a key new method that reveals a dark companion to the Milky Way.
(S. Chakrabarti, UC Berkeley)

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

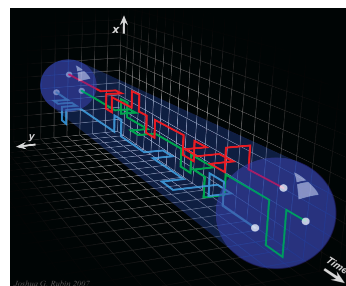
Nuclear Physics



Nuclear Decay Heating

The KamLAND neutrino experiment showed that radioactivity cannot be Earth's only heat source; it accounts for only ½ of it.

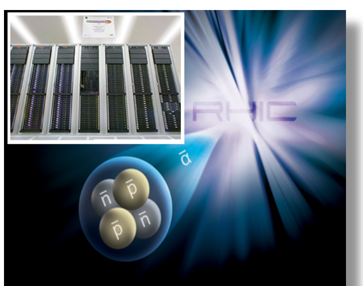
(S. Freedman, LBNL)



6-Quark Nucleons

Computations done at NERSC suggest the possible existence of a so-called H-dibaryon bound state, an exotic nucleus first envisaged in 1977.

(M. Savage, U. Washington)



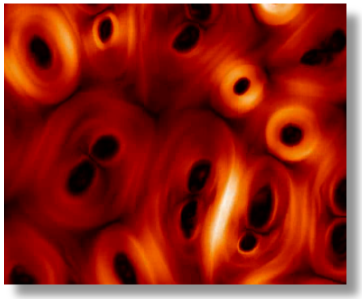
Antimatter

The heaviest antimatter particle has been discovered with NERSC help. Antihelium-4 is likely to hold the title for decades.

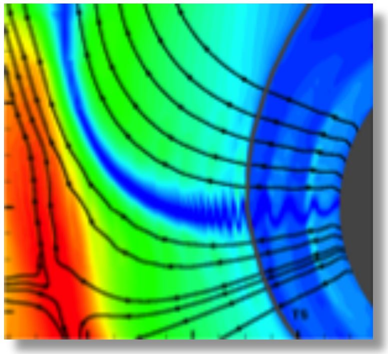
(STAR Collaboration at RHIC, BNL/ LBNL)

Fusion Energy Sciences

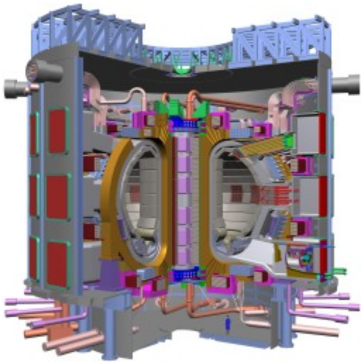
Magnetic Reconnection
Magnetic reconnection simulations done at NERSC along with NASA Voyager probe data help shake up prevailing views of the solar system's outer reaches.
(J. Drake, U. Maryland)



Voyager Surprise Explained
Study by award-winning researcher explained magnetic reconnection phenomenon observed by *Voyager* spacecrafts.
(J. Drake, U. Maryland)



ITER Design
Study supported by NERSC NISE award suggests ITER might require an alternate mitigation strategy for "runaway electron" current.
(V. Izzo, General Atomics)



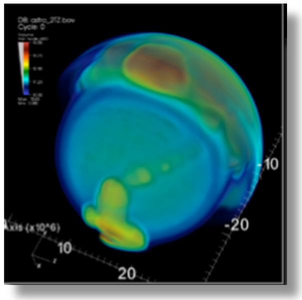
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

Advanced Scientific Computing Research

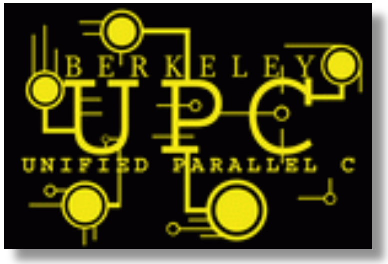
Visualization Technology

Demonstrated that visualization R&D has produced technology that can ingest and process tomorrow's "datasets" today.
(W. Bethel, LBNL)



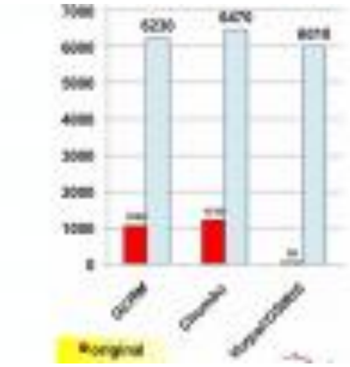
Programming Languages

Unified Parallel C (**UPC**) is an extension of the C programming language designed for high performance computing on large-scale parallel machines.
(K. Yelick, UC Berkeley/ LBNL)

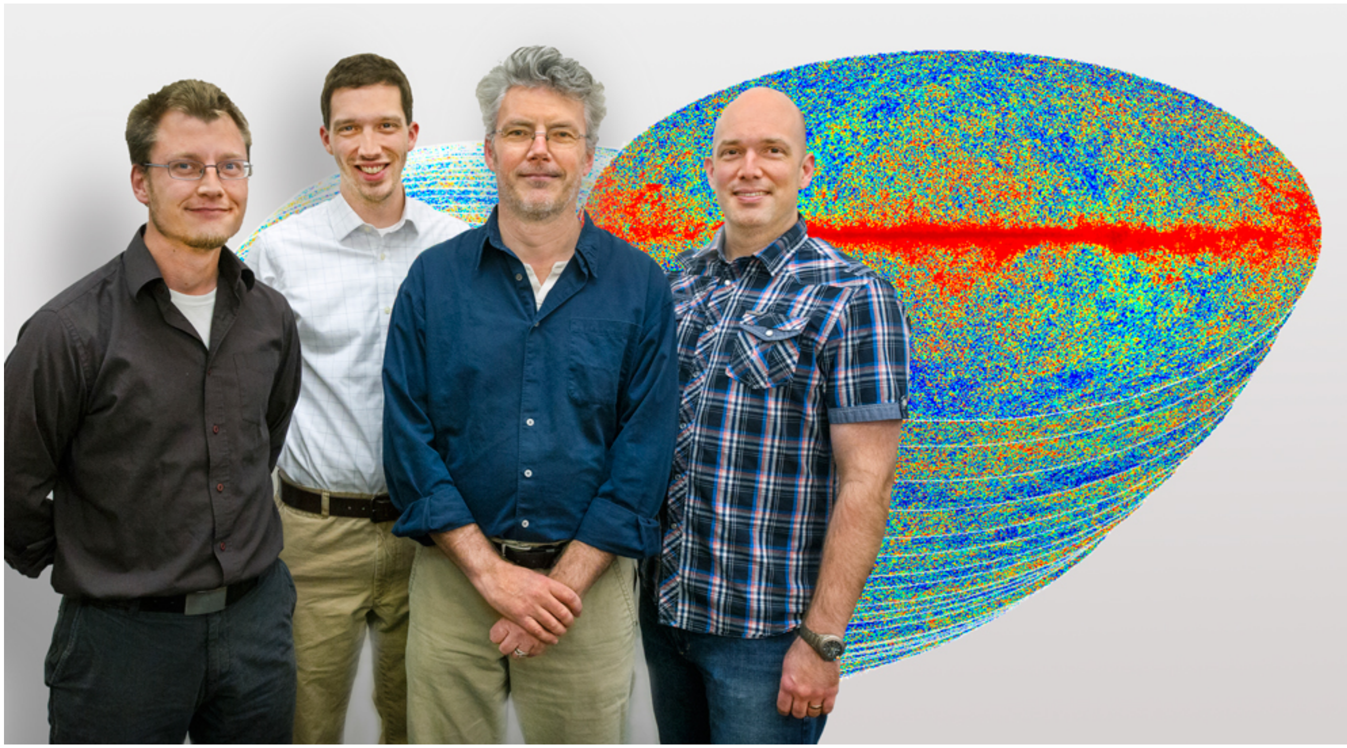


I/O Optimization

NERSC staff profiled & optimized HDF5 for Lustre and helped Cray optimize their MPI-IO, achieving a 10X improvement.



Building the Massive Simulation Sets Essential to Planck Results



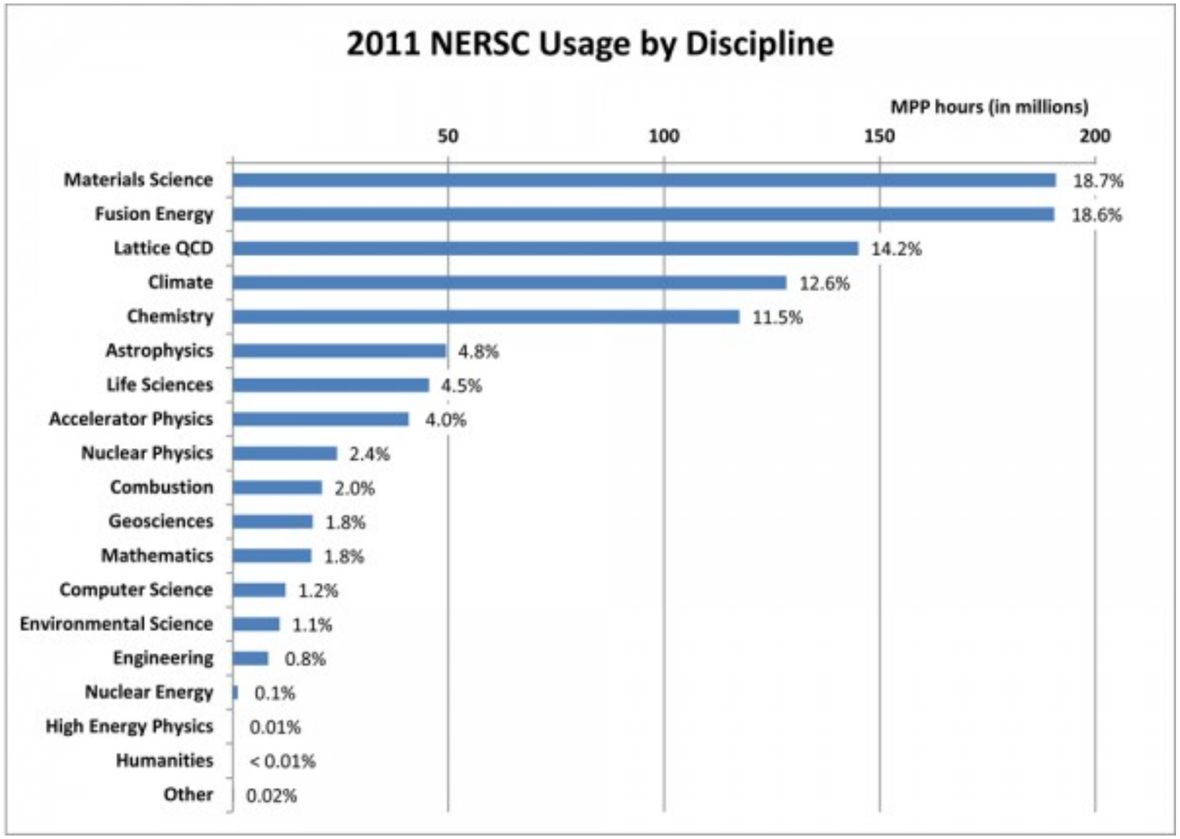
First Planck results: the Universe is still weird and interesting

Using NERSC supercomputers, Berkeley Lab scientists generate thousands of simulations to analyze the flood of data from the Planck mission.

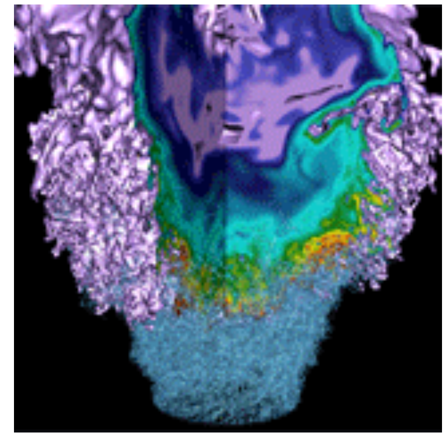
Generated 250,000 simulated maps of the Planck sky, all needed to minimize statistical uncertainty.

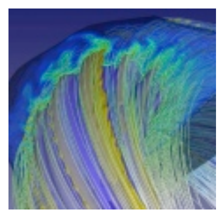
NERSC Supports DOE Open Science

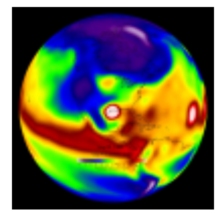
- Scientists from all Office of Science offices rely on NERSC
- Universities (54%), DOE Labs (39%)
- U.S. and International
- Individuals, teams of all sizes

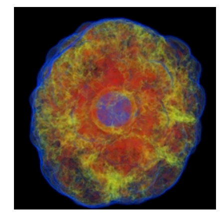


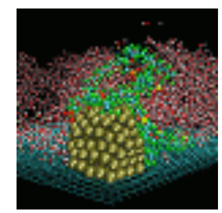
How Does HPC Computing Work?





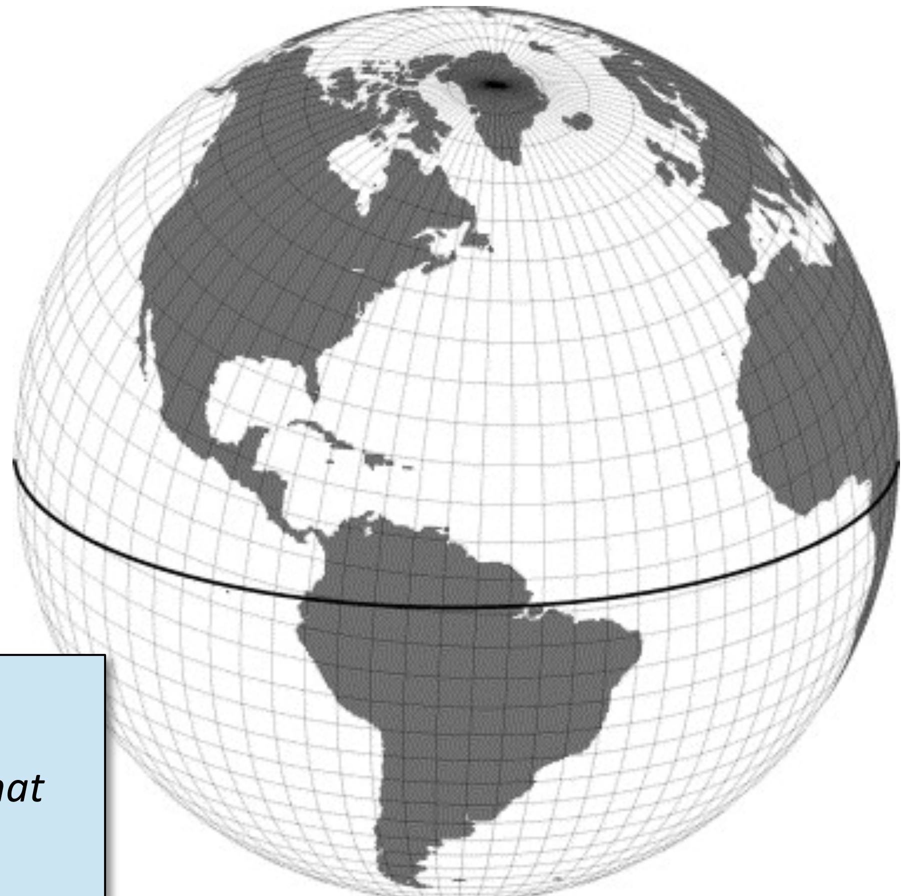






HPC Computing …

- implies parallel computing
- In parallel computing, scientists divide a big task into smaller ones
- “Divide and conquer”



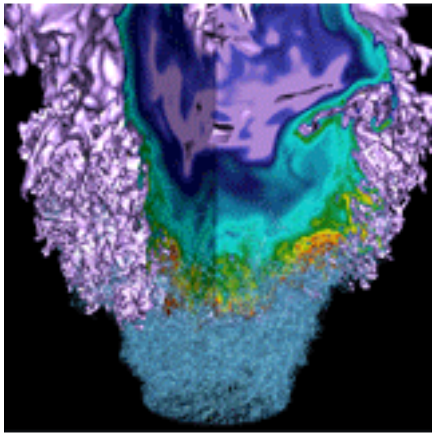
For example, to simulate the behavior of Earth’s atmosphere, you can divide it into zones and let each processor calculate what happens in each.

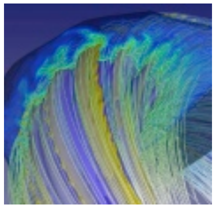
From time to time each processor has to send the results of its calculation to its neighbors.

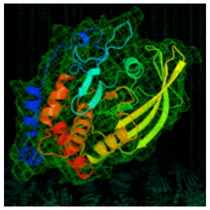
Distributed Memory Systems

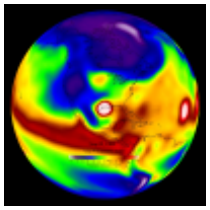
- **This maps well to HPC “distributed memory” systems**
  - Many nodes, each with its own local memory and distinct memory space
  - A node typically has multiple processors, each with multiple compute cores (Hopper has 24 cores per node)
  - Nodes communicate over a specialized high-speed, low-latency network
  - SPMD (Single Program Multiple Data) is the most common model
    - Multiple copies of a single program (tasks) execute on different processors, but compute with different data
    - Explicit programming methods (MPI) are used to move data among different tasks

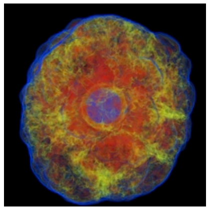
Challenges in HPC

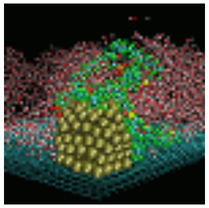










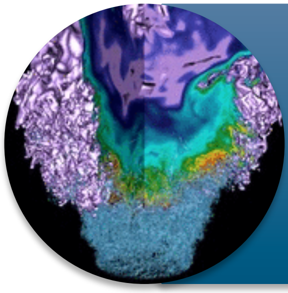


U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

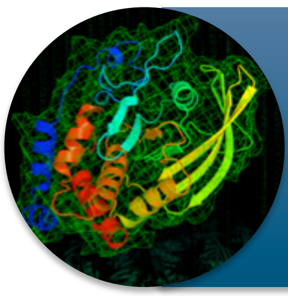
Thrusts in High Performance Computing

NERSC



Science at Scale

Petaflops to Exaflops



Science through Volume

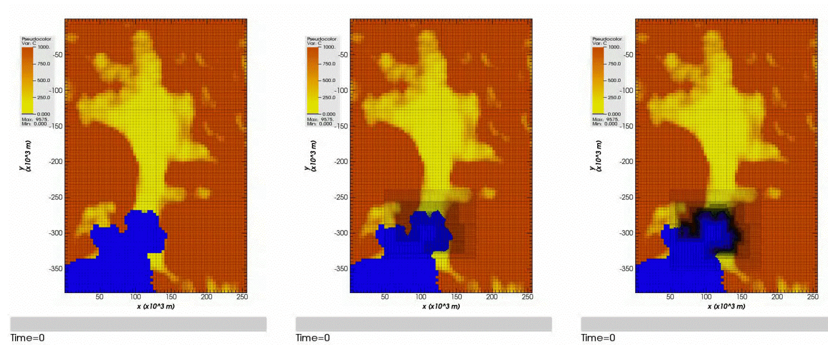
Thousands to Millions of Simulations

Science in Data

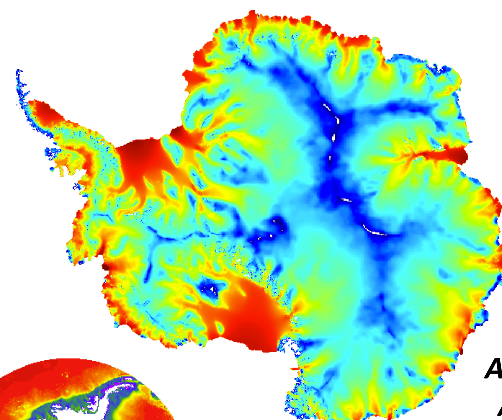
Petabytes to Exabytes of Data

Science at Scale: Simulations Aid in Understanding Climate Impacts

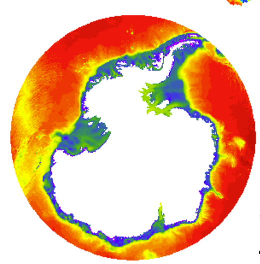
- **Warming ocean and Antarctic ice sheet key to sea level rise**
- **BISICLES ice sheet model uses AMR for ice-ocean interface.**
  - Dynamics very fine resolution (AMR)
  - Antarctica still very large (scalability)
  - Multi-institution (LANL, LBNL)
- **Ongoing collaboration to couple ice sheet and ocean models**
  - 19M ALCC Hours at NERSC



BISICLES Pine Island Glacier simulation – mesh resolution crucial for grounding line behavior.



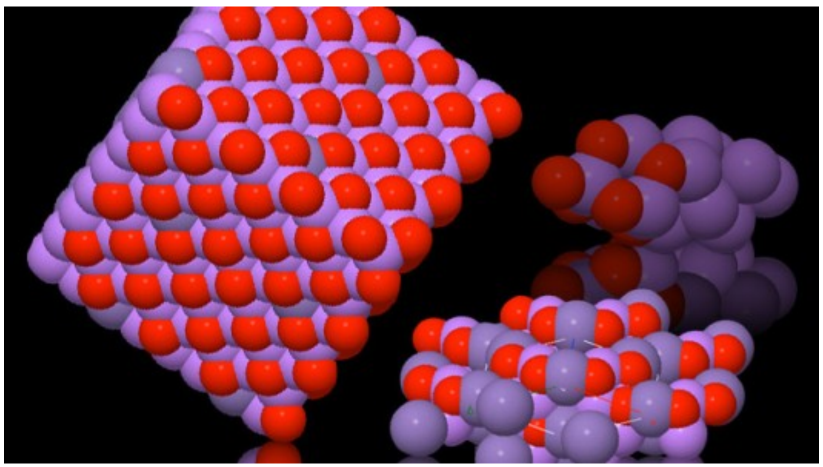
Antarctic ice speed (left): AMR enables sub-1 km resolution (black, above)
(Using NERSC's Hopper)

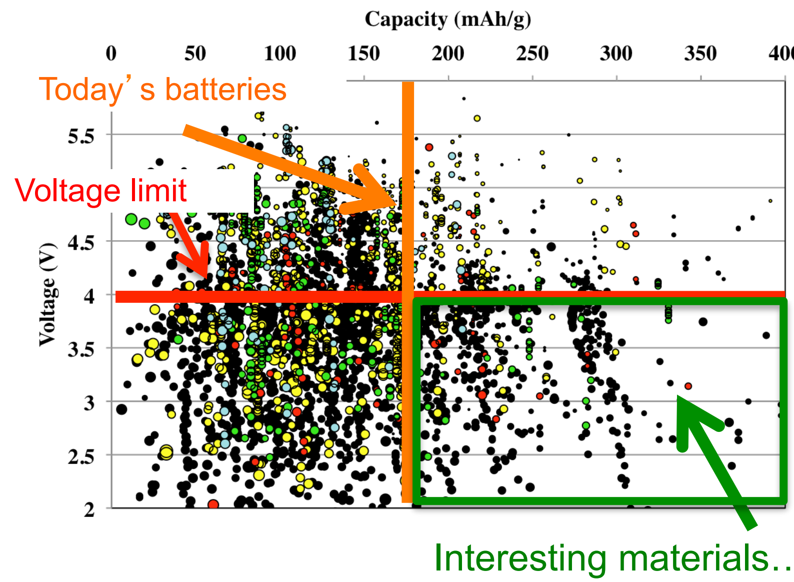


Enhanced POP ocean model solution for coupling to ice

Science through Volume: Large Numbers of Simulations for Materials

- Tens of thousands of simulations are used to screen related materials for use in battery design and other domains
- Goal: cut in half the 18 year from design to manufacturing

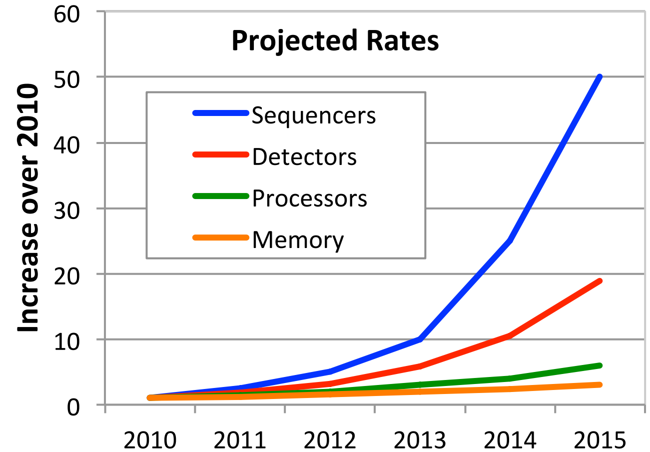




<u>Materials Project, Gerd Ceder PI (MIT):</u> *website has a database materials from simulations, e.g., over 20,000 potential battery materials.*

PIs: Gerd Ceder, MIT and Kristin Persson, LBNL

Science in Data: Automated Image Analysis in Astronomy

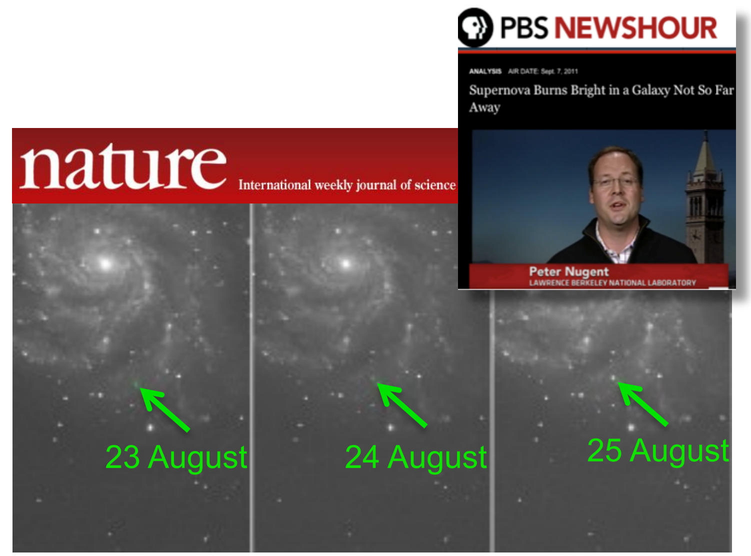


Data from scientific instruments is growing exponentially

- NERSC in 2 Nobel prizes, and 3 Science "best of decade" (CMB and Genomics)
- Far outpacing processor and memory performance growth

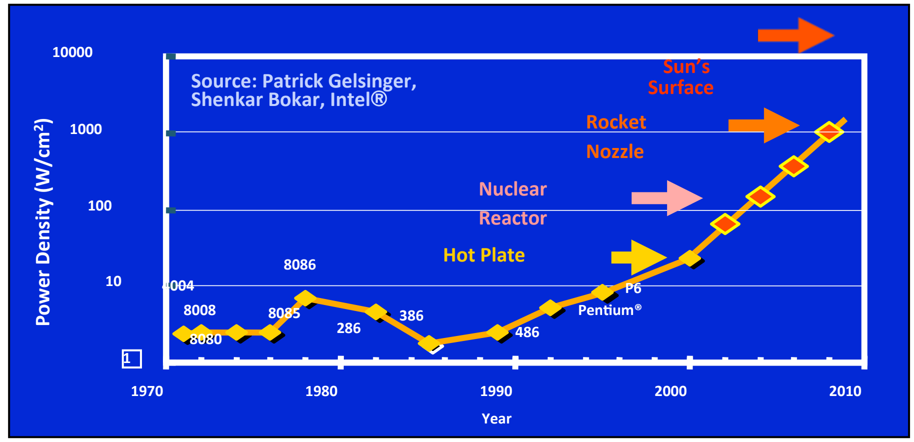
Astrophysicists discover early nearby supernova

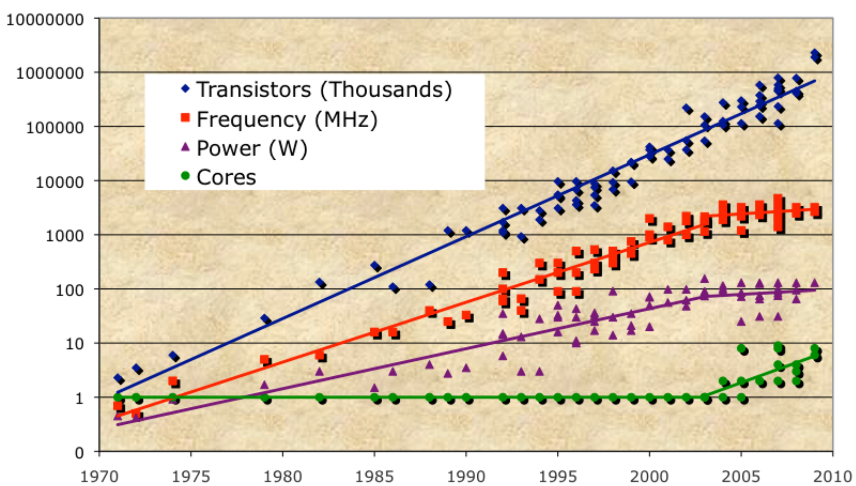
- Palamor Transient Factory runs machine learning algorithms on **~300GB/night** delivered by ESnet "science network"
- Rare glimpse of a supernova within 11 hours of explosion, 20M light years away
- Telescopes world-wide redirected



Power: the Biggest Architectural Challenge

If we just kept making computer chips faster and more dense, they'd melt and we couldn't afford or deliver the power.



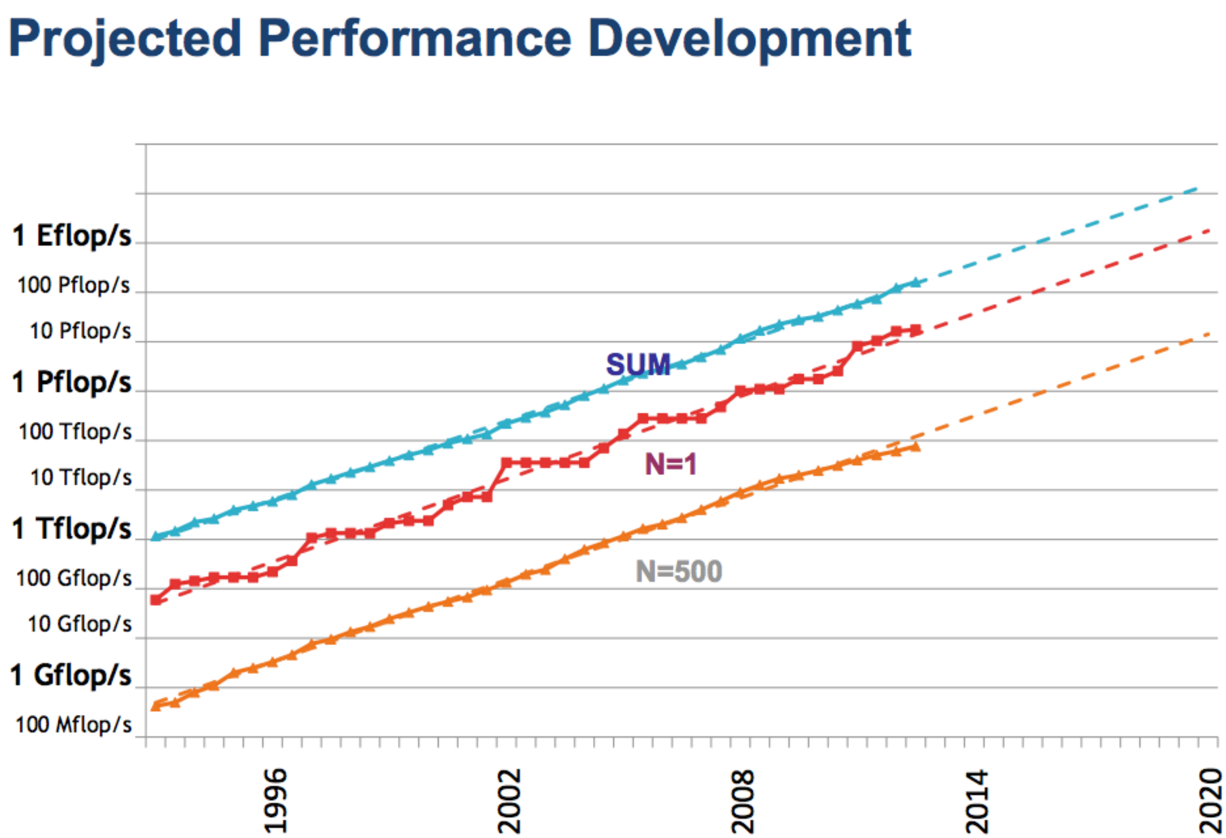


Now compute cores are getting slower and simpler, but we're getting lots more to maintain the performance curve.

GPUs and Intel Phi have hundreds of "light-weight cores"

Revolution in Energy Efficiency Needed

Even though energy efficiency is increasing, today's top supercomputer (N=1) uses ~9 MW or roughly $9M/year to operate. Even if we could build a working exaflop computer today, it would use about 450 MW and cost $450M/year to pay for power.



U.S. DEPARTMENT OF ENERGY | Office of Science

Programming for Many-Core: Biggest Software Challenge

- **To effectively use many-core processors, programs must exploit 100K – 1M way parallelism.**
- **Traditional programming paradigms won't work**
  - Too resource intensive per MPI task
  - Data movement is extremely expensive
- **Current programming methods for accelerators (GPUs) are difficult**
  - Need one "fat core" (at least) for running the OS
  - Data movement from main memory to GPU memory kills performance
  - Programmability is very poor
  - Most codes will require extensive overhauls

Data: Getting Bigger All the Time

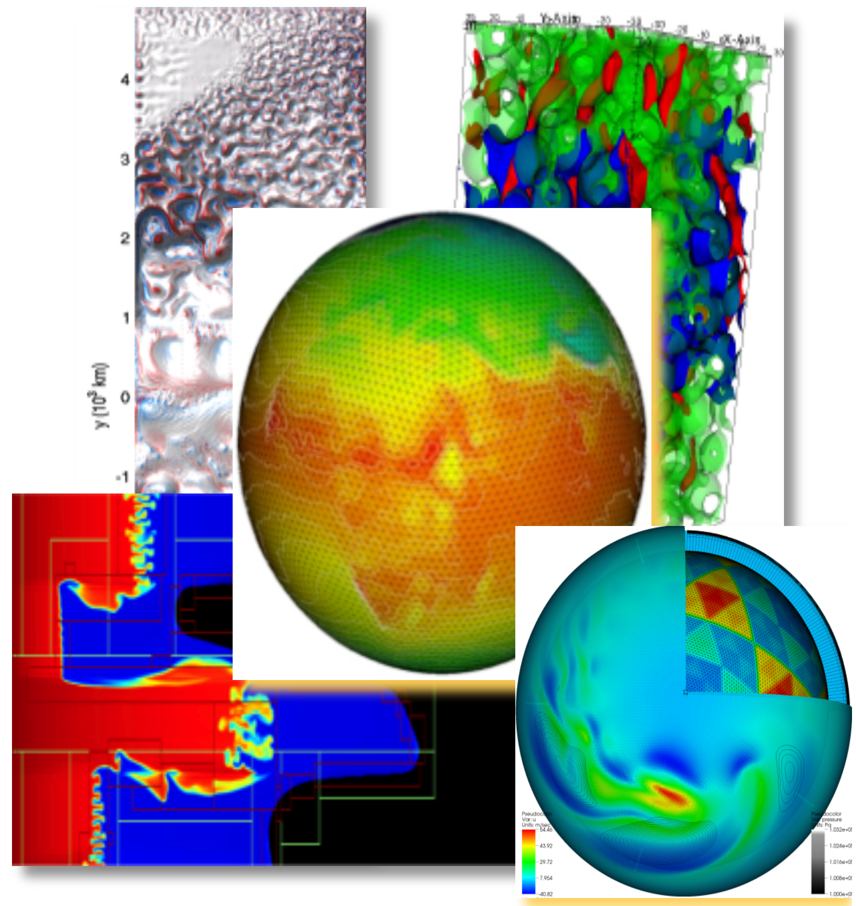
NERSC

Larger simulations are producing ever more data and data subsystems are not keeping up.

Huge experimental facilities like the LHC, beam lines, telescopes, satellites, etc. are producing unprecedented volumes of data at mind-boggling rates.

Reading, writing, and transferring this data is a serious challenge. Making sense of it via data analytics and visualization is too.

Data and job management is another largely unsolved problem. Effective workflows and job schedulers are much needed.



Images from David Randall, Paola Cessi, John Bell, T Scheibe

Your Challenges

- Figure out how to program the next generation of machines
- Find a way to make sense of all the data
- Build faster, more capable hardware that uses less energy
- Create effective data and job management workflows
- Bring new fields of science into HPC
- Tell the world about what you're doing

National Energy Research Scientific Computing Center

Top 500

Top 10 Systems in November 2012

#	Site	Manufacturer	Computer	Country	Cores	Rmax [Pflops]	Power [MW]
1	Oak Ridge National Laboratory	Cray	Titan Cray XK7, Opteron 16C 2.2GHz, Gemini, NVIDIA K20x	USA	560,640	17.6	8.21
2	Lawrence Livermore National Laboratory	IBM	Sequoia BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	1,572,864	16.3	7.89
3	RIKEN Advanced Institute for Computational Science	Fujitsu	K Computer SPARC64 VIIIfx 2.0GHz, Tofu Interconnect	Japan	795,024	10.5	12.66
4	Argonne National Laboratory	IBM	Mira BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	786,432	8.16	3.95
5	Forschungszentrum Juelich (FZJ)	IBM	JuQUEEN BlueGene/Q, Power BQC 16C 1.6GHz, Custom	Germany	393,216	4.14	1.97
6	Leibniz Rechenzentrum	IBM	SuperMUC iDataPlex DX360M4, Xeon E5 8C 2.7GHz, Infiniband FDR	Germany	147,456	2.90	3.52
7	Texas Advanced Computing Center/UT	Dell	Stampede PowerEdge C8220, Xeon E5 8C 2.7GHz, Intel Xeon Phi	USA	204,900	2.66	
8	National SuperComputer Center in Tianjin	NUDT	Tianhe-1A NUDT TH MPP, Xeon 6C, NVidia, FT-1000 8C	China	186,368	2.57	4.04
9	CINECA	IBM	Fermi BlueGene/Q, Power BQC 16C 1.6GHz, Custom	Italy	163,840	1.73	0.82
10	IBM	IBM	DARPA Trial Subset Power 775, Power7 8C 3.84GHz, Custom	USA	63,360	1.52	3.57

Section Title