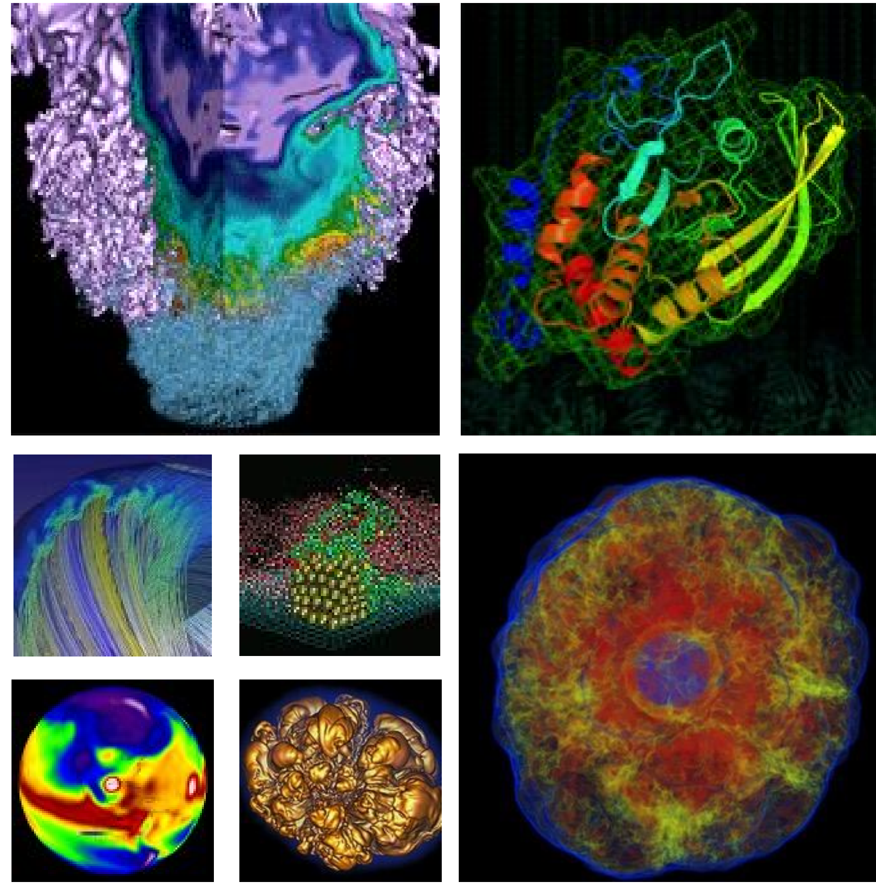
High-Performance Computing and NERSC



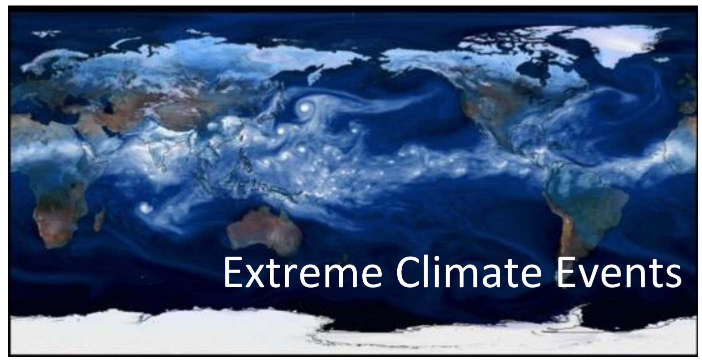
Rebecca Hartman-Baker
Group Lead, User Engagement

Brandon Cook
Application Performance

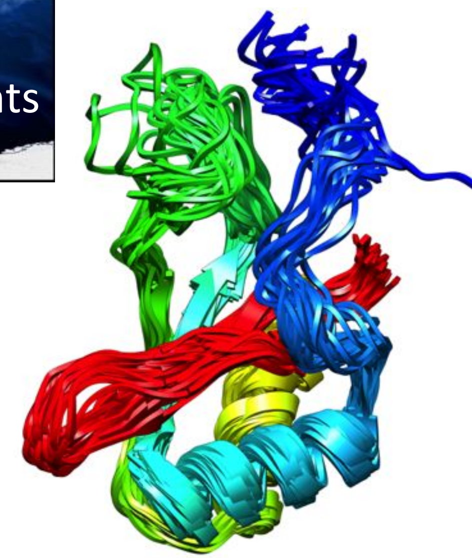
June 8, 2017

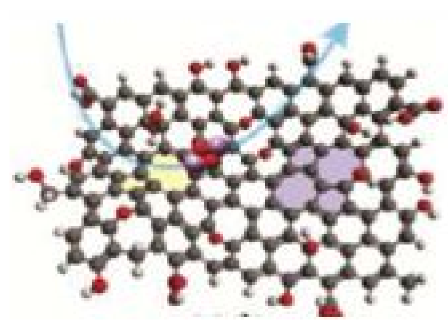
High-Performance Computing is …

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



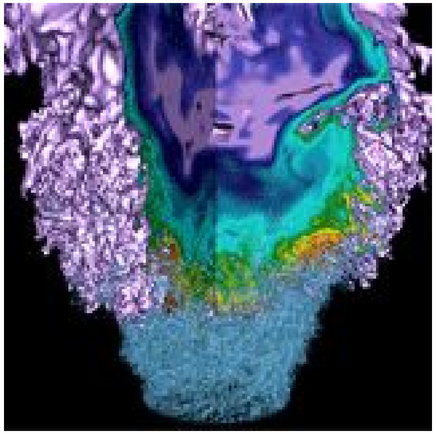
Understanding How Proteins Work



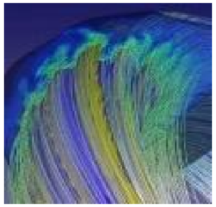
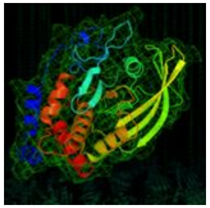
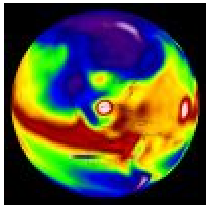
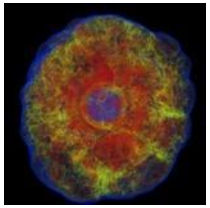
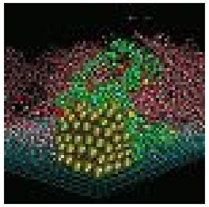


Designing Better Batteries

What is a Supercomputer?



NERSC

What is a supercomputer?

A. A processor (CPU) unimaginably more powerful than the one in my laptop.

B. A quantum computer that takes advantage of the fact that quantum particles can simultaneously exist in a vast number of states.

C. Processors not so different than the one in my laptop, but 100s of thousands of them working together to solve a problem. ✔

A Supercomputer is …

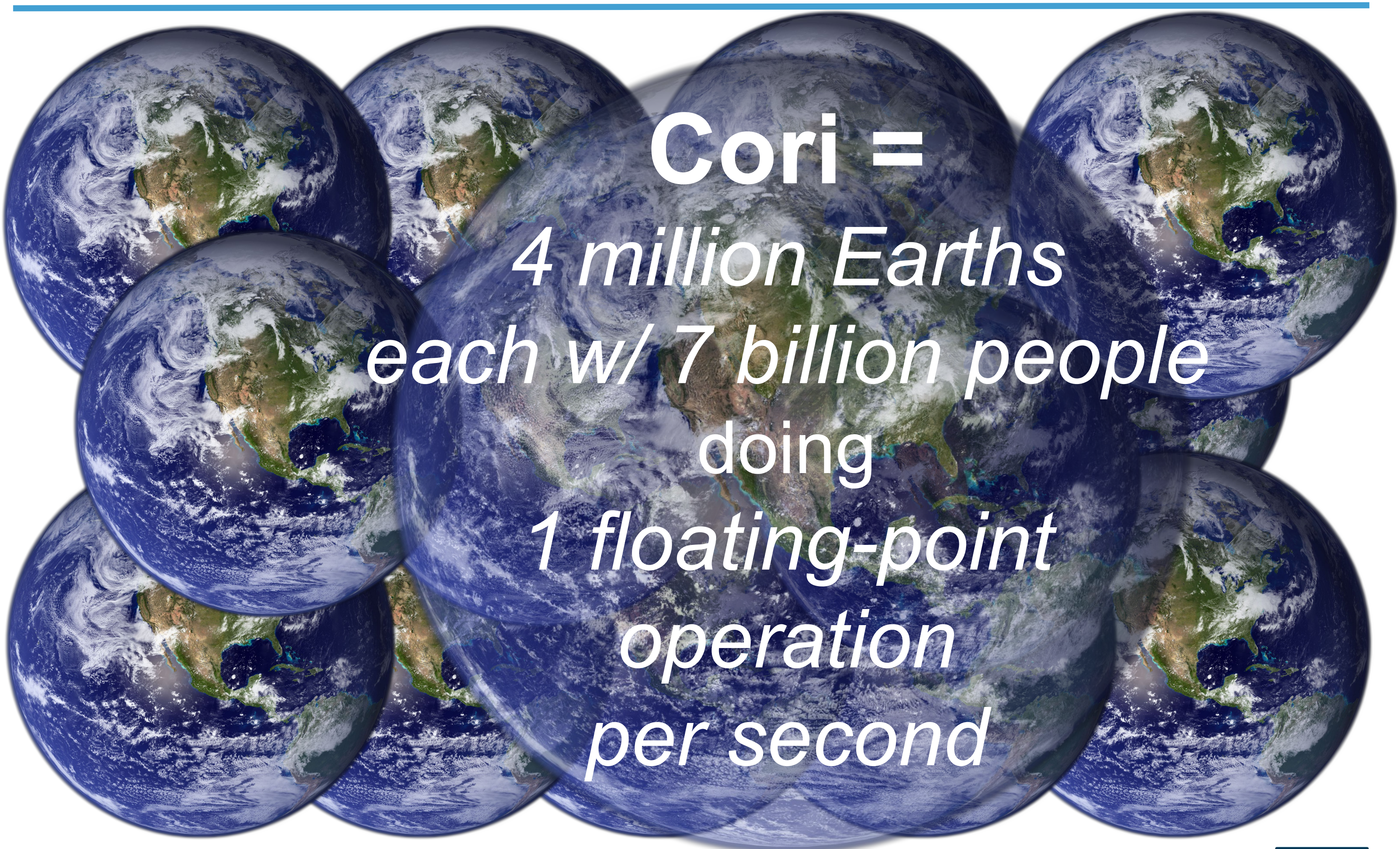
vs.

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

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

Cori (left) has ~11,000 nodes (~ high-end desktop computer)

700,000 compute cores that can perform ~3x10^{16} calculations/second


Cori =
4 million Earths
each w/ 7 billion people
doing
1 floating-point
operation
per second

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

5,000 X

Latency

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

20,000 X

Cloud systems have slower networks

* If you're really lucky

And Even More …

PBs of fast storage for files and data

Cori: 30 PB
Your laptop: 0.0005 PB
Your iPhone: 0.00005 PB

60,000 X

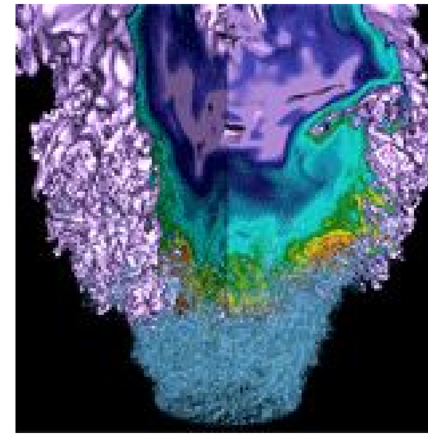
Write data to permanent storage

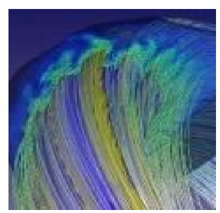
Cori: 700 GB/sec
My iMac: 0.01 GB/sec

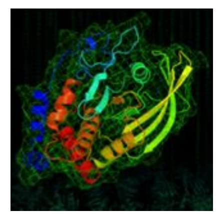
70,000 X

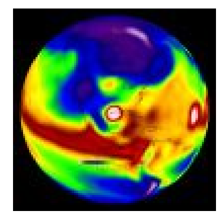
Cloud systems have slower I/O and less permanent storage

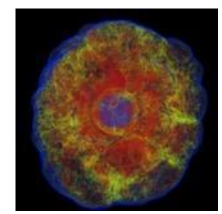
High-Performance Computing

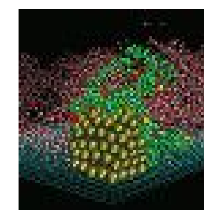






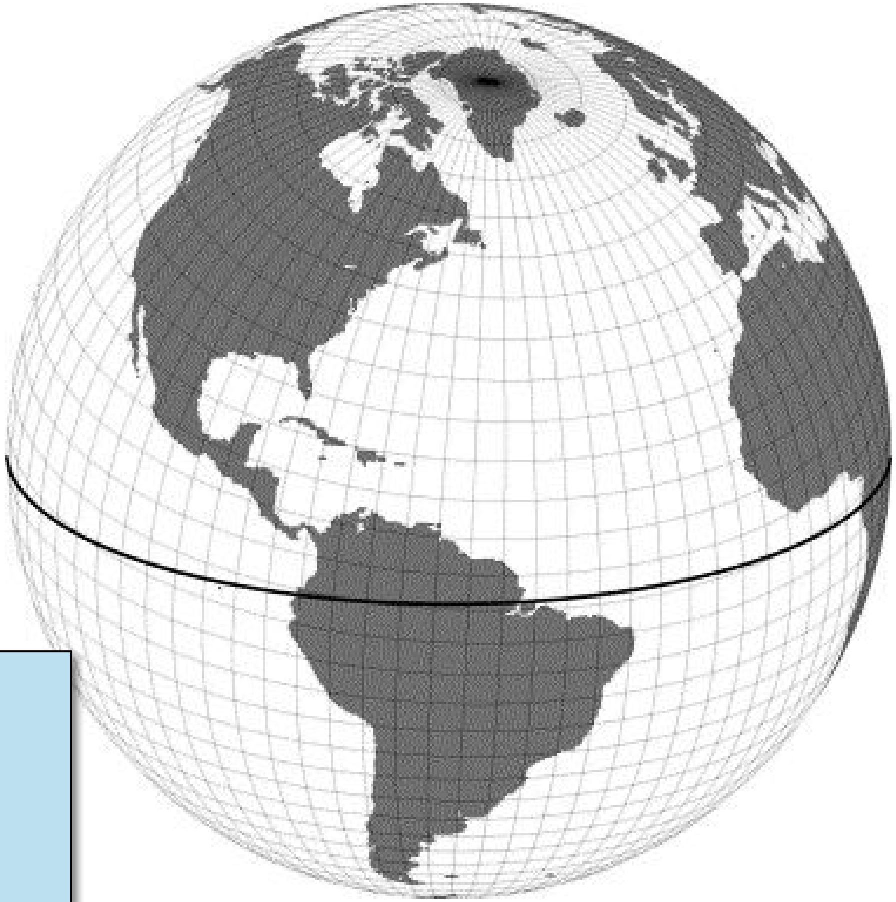






HPC Computing …

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



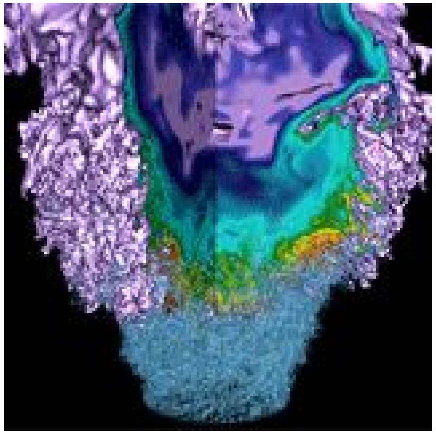
For example, to simulate the behavior of Earth’s atmosphere, you can divide it into zones and let each processor calculate what happens in each.

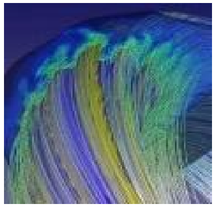
From time to time each processor has to send the results of its calculation to its neighbors.

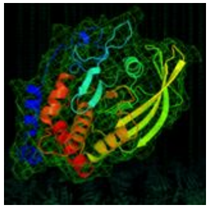
Distributed Memory Systems

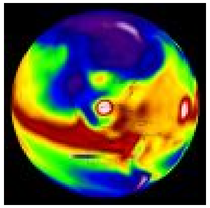
- **This maps well to HPC "distributed memory" systems**
  - Many nodes, each with its own local memory and distinct memory space
  - A node typically has multiple processors, each with multiple compute cores (Cori has 32 or 68 cores per node)
  - Nodes communicate over a specialized high-speed, low-latency 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

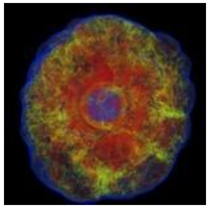
What is NERSC?

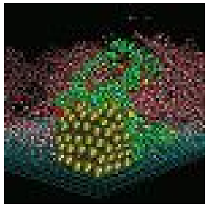












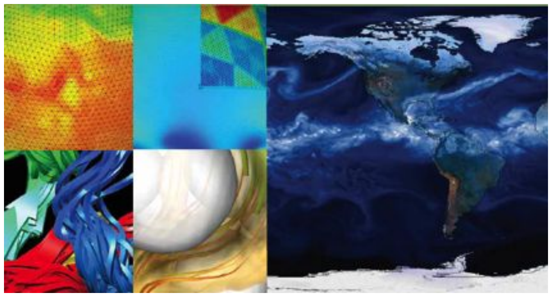
National Energy Research Scientific Computing Center (NERSC)

- **NERSC is a national supercomputer center funded by the U.S. Department of Energy Office of Science (SC)**
  - Supports SC research mission
  - Part of Berkeley Lab
- **If you are a researcher with funding from SC, you can use NERSC**
  - Other researchers can apply if research is in SC mission
- **NERSC supports 7,000 users, 800 projects**
- **From 48 states; 65% from universities**
- **Hundreds of users log on each day**

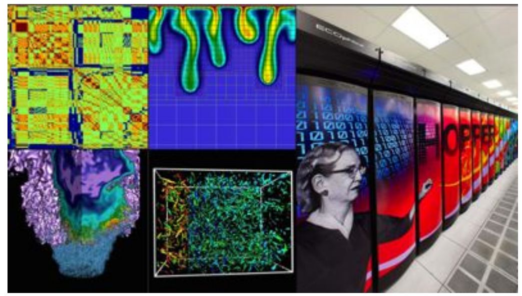
Facility for DOE Office of Science Research

U.S. DEPARTMENT OF ENERGY | Office of Science

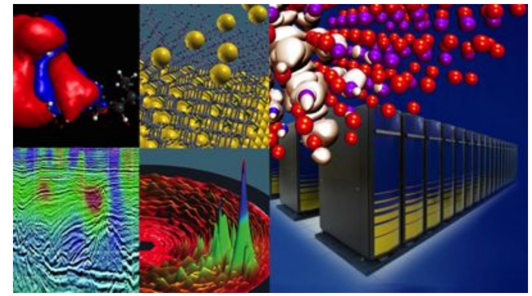
Largest funder of physical science research in U.S.



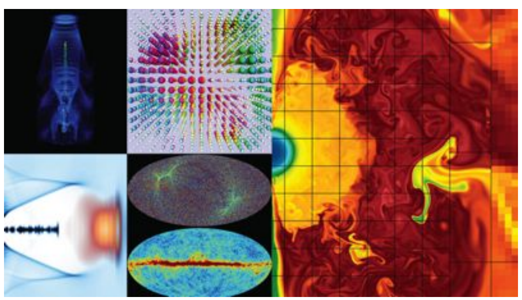
Bio Energy, Environment



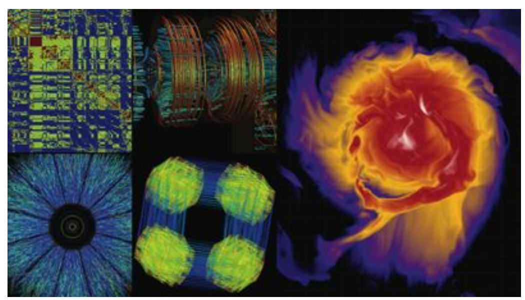
Computing



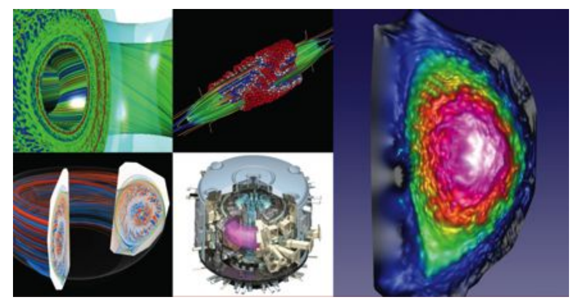
Materials, Chemistry, Geophysics



Particle Physics, Astrophysics



Nuclear Physics



Fusion Energy, Plasma Physics

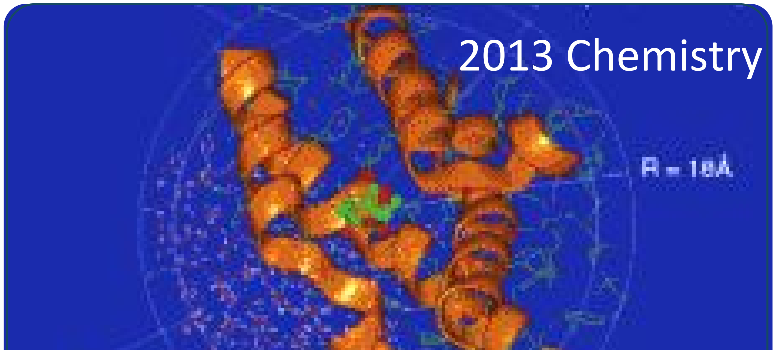
NERSC: Science First

NERSC's mission is to accelerate scientific discovery at the DOE Office of Science through high performance computing and data analysis.

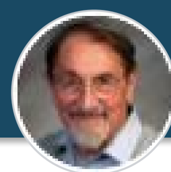
2015 Science Output

2,078 refereed publications

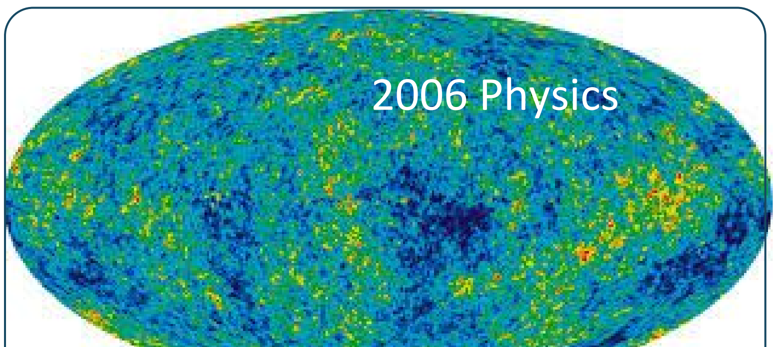
NERSC Nobel Prize Winners



Martin Karplus



Saul Perlmutter



George Smoot

Warren Washington

Nobel Prize in Physics 2015

NERSC

Scientific Achievement

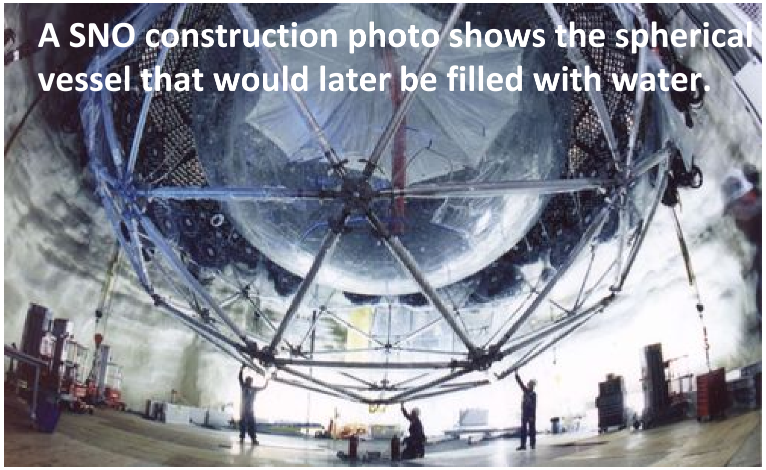
The discovery that neutrinos have mass and oscillate between different types

Significance and Impact

The discrepancy between predicted and observed solar neutrinos was a mystery for decades. This discovery overturned the Standard Model interpretation of neutrinos as massless particles and resolved the "solar neutrino problem"

Research Details

The Sudbury Neutrino Observatory (SNO) detected all three types (flavors) of neutrinos and showed that when all three were considered, the total flux was in line with predictions. This, together with results from the Super Kamiokande experiment, was proof that neutrinos were oscillating between flavors and therefore had mass



Calculations performed on PDSF & data stored on HPSS played a significant role in the SNO analysis. The SNO team presented an autographed copy of the seminal *Physical Review Letters* article to NERSC staff.

Q. R. Ahmad et al. (SNO Collaboration). Phys. Rev. Lett. 87, 071301 (2001)

Nobel Recipients: Arthur B. McDonald, Queen's University (SNO)
Takaaki Kajita, Tokyo University (Super Kamiokande)

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

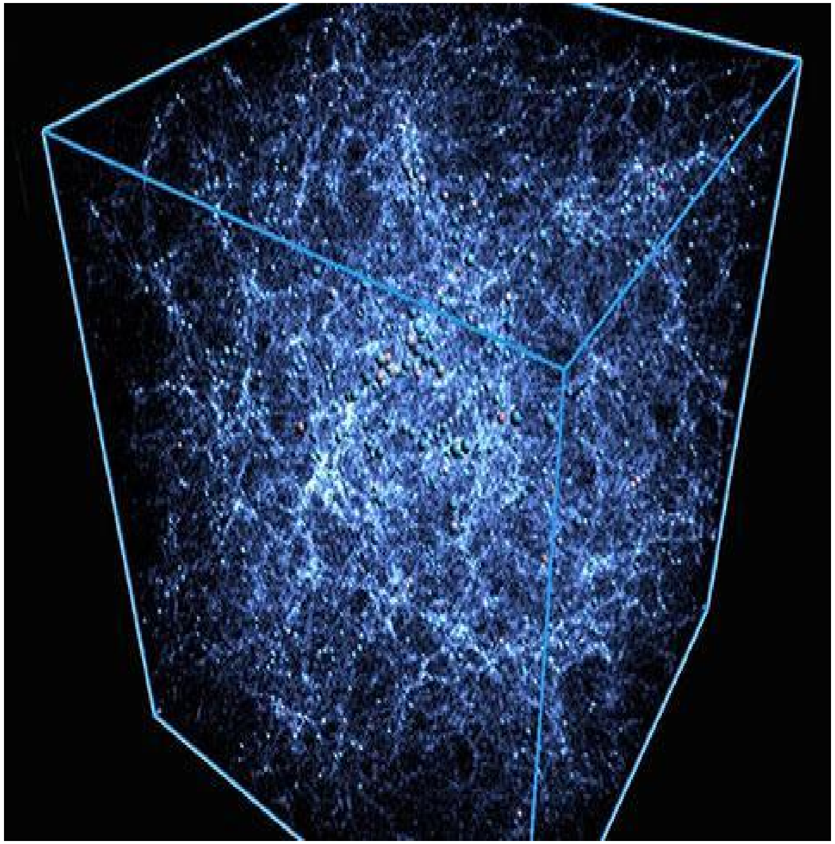
Science on Cori - Hot off the Presses

3-Pt Correlation On 2B Galaxies Recently Completed on Cori

- NESAP For Data Prototype (Galactos)
- First anisotropic, 3-pt correlation computation on 2B Galaxies from Outer Rim Simulation
- Solves an open problem in cosmology for the next decade (LSST will observe 10B galaxies)
- Can address questions about the nature of dark-energy and gravity
- Novel $O(N^2)$ algorithm based on spherical harmonics for 3-pt correlation

Scale:

- 9600+ KNL Nodes (Significant Fraction of Peak)

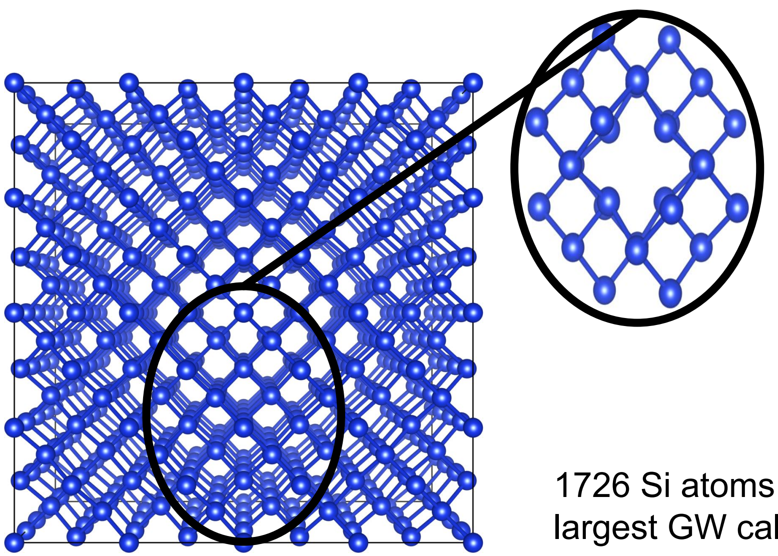


Science on Cori - Hot off the Presses

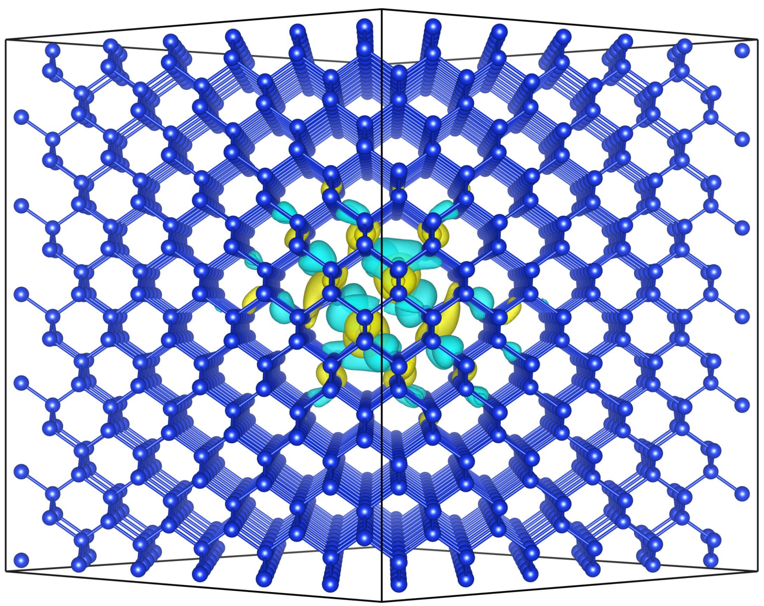
Defect States in Materials:
Important material properties of, for examples, transistors and photovoltaics are often determined by the effects of defects. However, accurately studying defect properties require extremely large calculations to isolate defect states using BerkeleyGW - Featured in **one of 5 BES Material Software Centers**

Scale:
Simulated on Cori with up to 9600 KNL Nodes - Near Perfect Strong and Weak Scaling. Large percentage of peak performance obtained > 10 PFLOPS.

BerkeleyGW



Schematic of di-vacancy defect and localized defect orbital in crystaline Silicon.



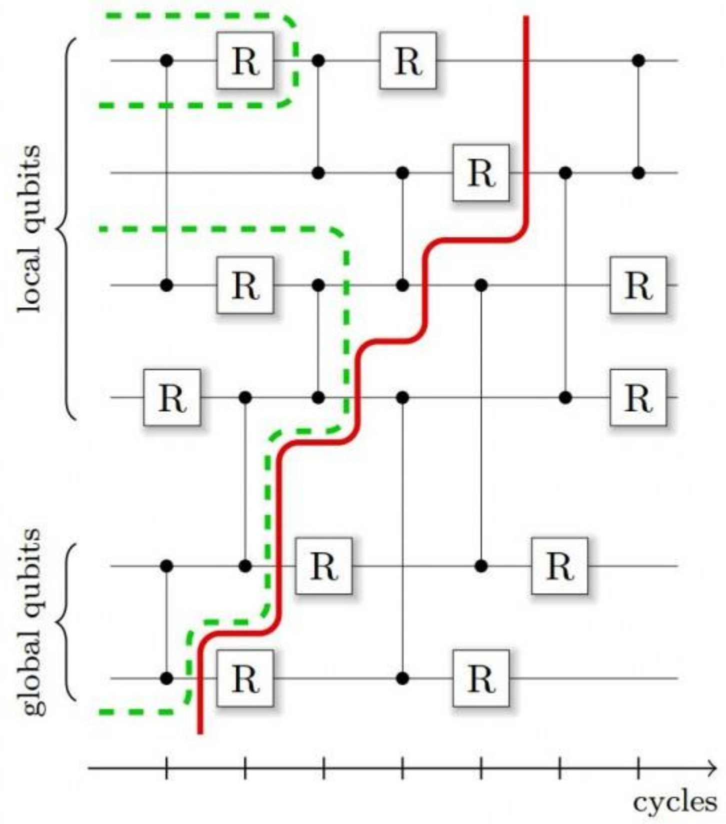
1726 Si atoms (~7K electrons) is largest GW calculation published

Quantum Supremacy Simulations

- 45 Qubit simulation is largest ever quantum computing simulation ever
- Previously largest calculation was 42 (complexity is exponential)
- Simulationas are important for validating prototype quantum computers devices
- Team lead by ETH scientists collaborators at Google, LBNL's Computational Research Divisio

Scale:

- >8000 KNL nodes
- 0.5 Petabytes of memory used (~2^45)
- 0.43 PetaFLOPS (Bandwidth bound)



Deep-Learning on Full Cori System

- Supervised Classification for LHC datasets
- Pattern discovery for climate datasets
- Production DL stack (IntelCaffe, MLSL)
- Convolutional architectures optimized on KNL with IntelCaffe and MKL
- Synch + Asynch parameter update strategy for multi-node scaling

Scale:

- 9600 KNL nodes on Cori
- 10 Terabyte datasets
- Millions of Images

Machine learning techniques can automatically detect patterns in simulation data. Applied to climate and particle-physics data from collider experiments.

In 2016 scientists at NERSC

used

410,000 single-CPU-years

3,600,000,000

MPP hours of compute time

Homo erectus
~300,000 years ago

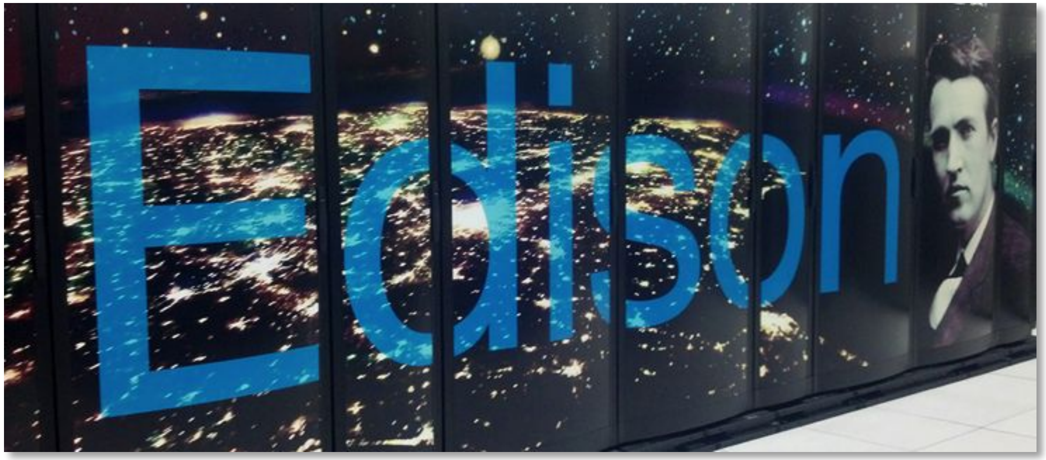
and currently store

90,000,000

2 million iPhones

Gbytes of data

Compute Hours



Edison

2,000 M hours

Cori

HSW - 1,000 M hours

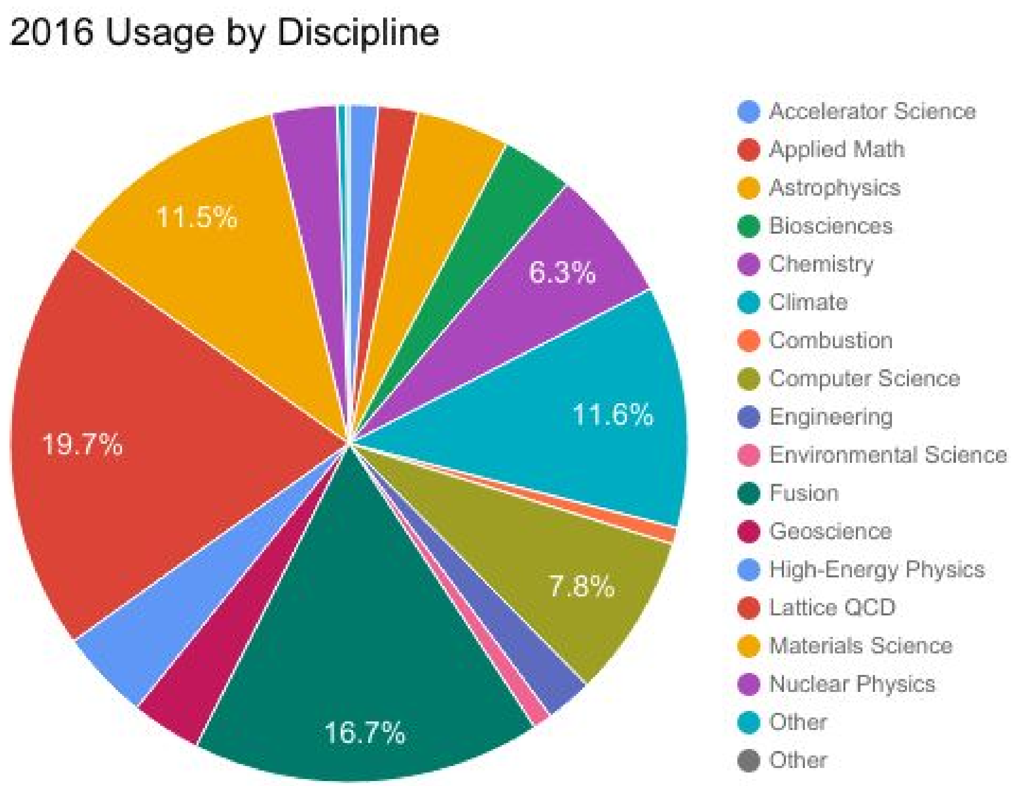
KNL - 6,000 M hours

Data Storage

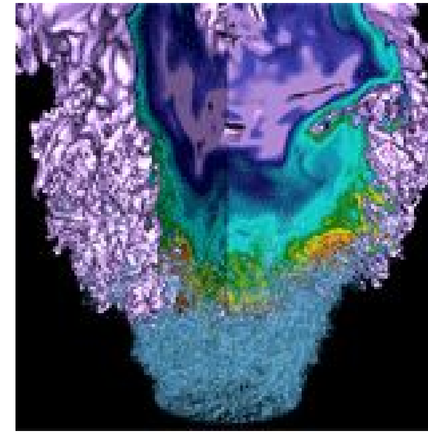
HPSS
105 Petabytes

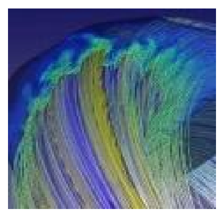
Project
4 Petabytes

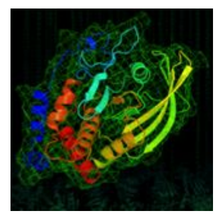
MPP Usage by Science Area

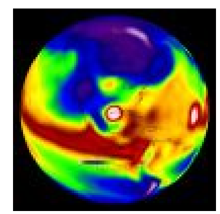


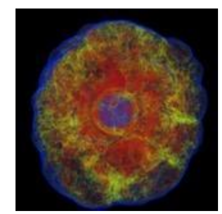
Challenges in HPC

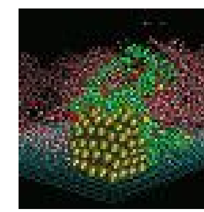






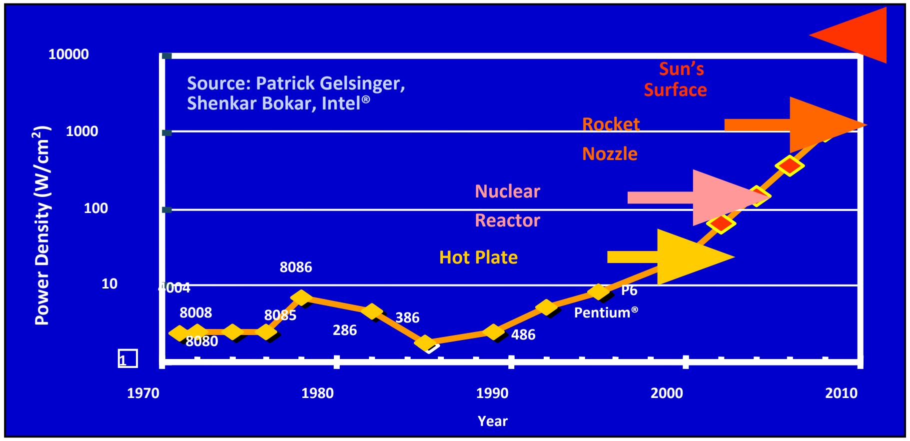


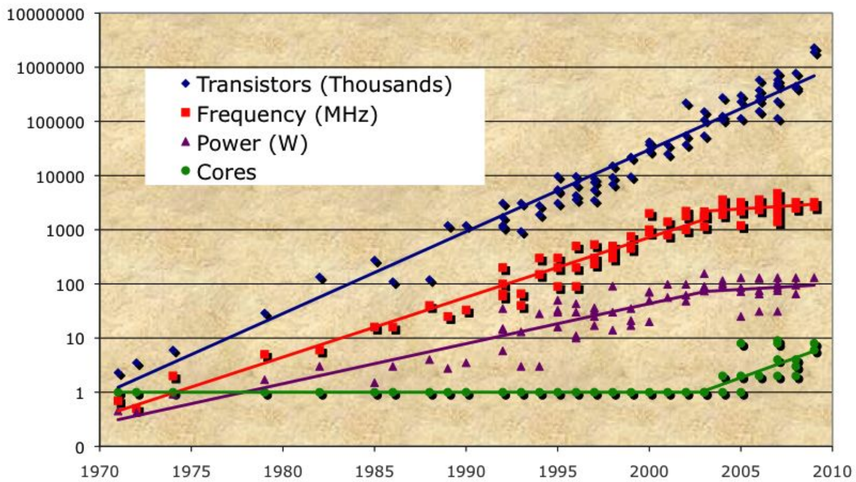




Power: the Biggest Architectural Challenge

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



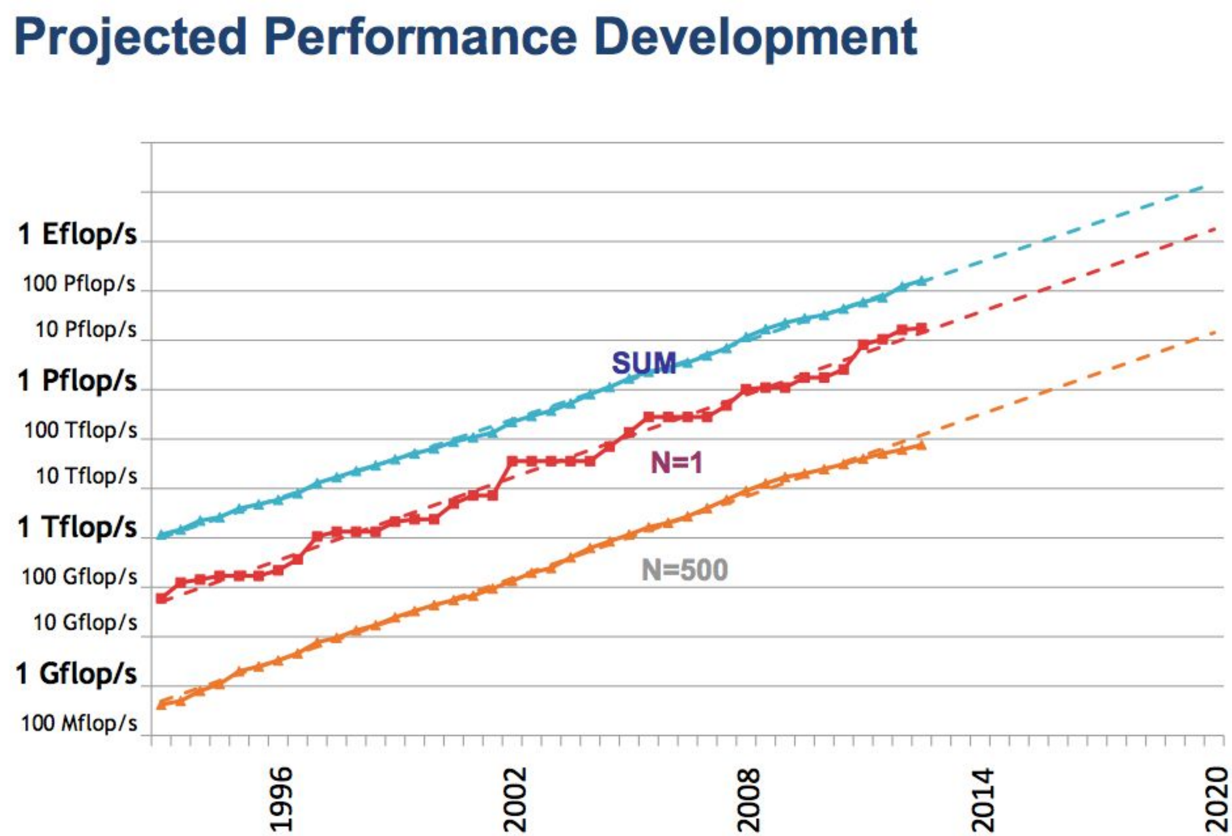


Now compute cores are getting slower and simpler, but we're getting lots more on a chip.

GPUs and Intel Xeon Phi have 60+ "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.



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
  - Must incorporate fine-grained (threaded) parallelism
- **Current programming methods for offload to accelerators are difficult and non-portable**
  - 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

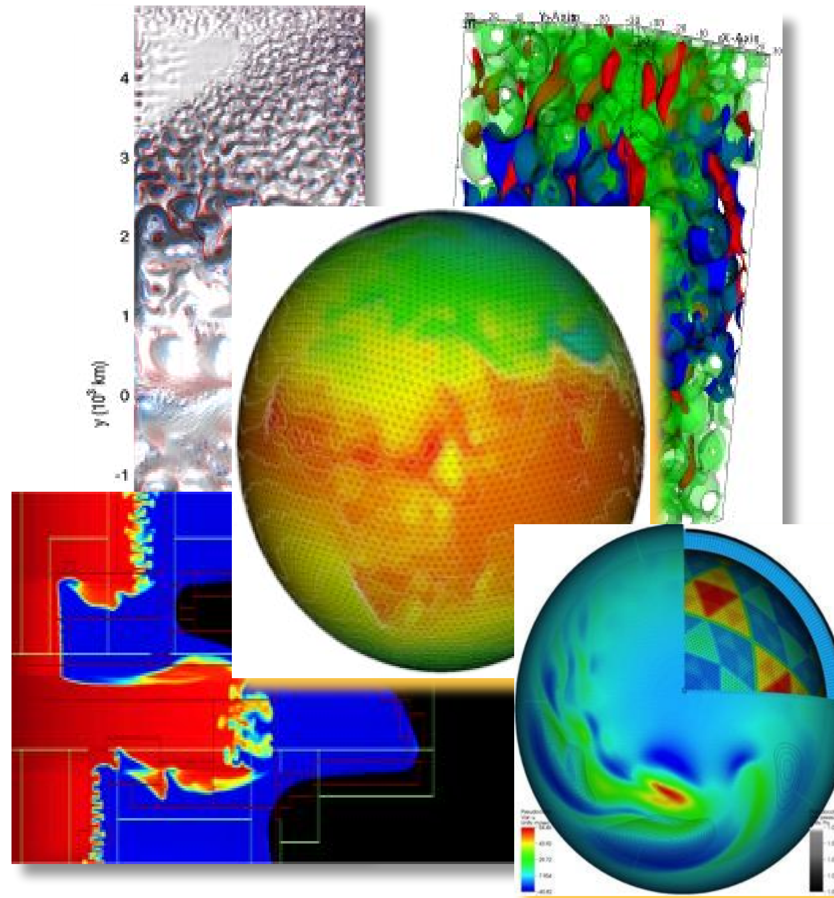
Data: Getting Bigger All the Time

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