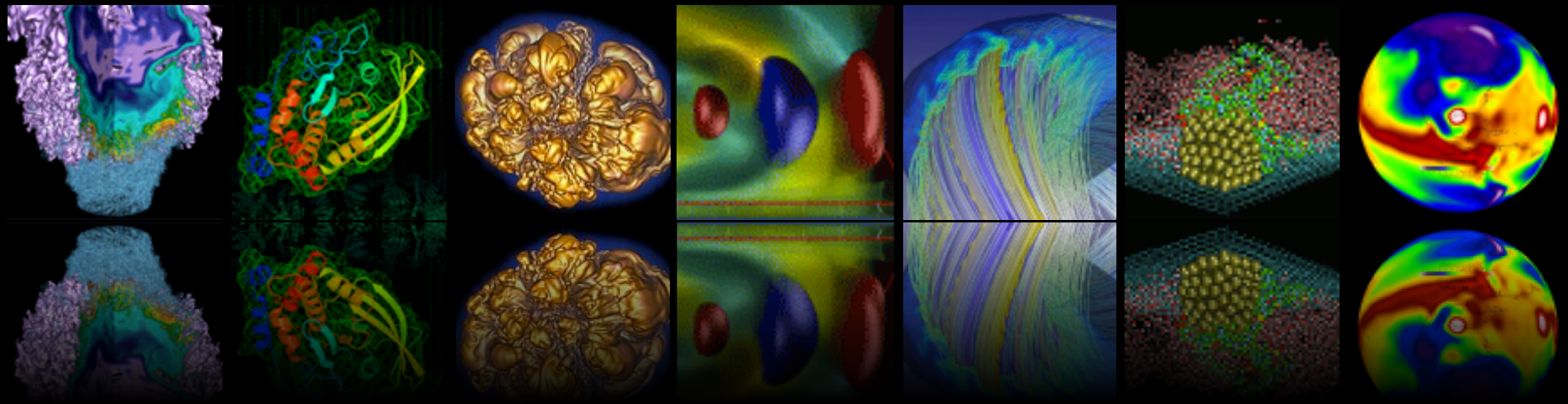


More Data, More Science and… Moore's Law

Kathy Yelick

Professor of Electrical Engineering and Computer Sciences
University of California at Berkeley
Associate Laboratory Director for Computing Sciences
Lawrence Berkeley National Laboratory

Three parts

- Science opportunities
- Technology challenges
- Productivity approaches

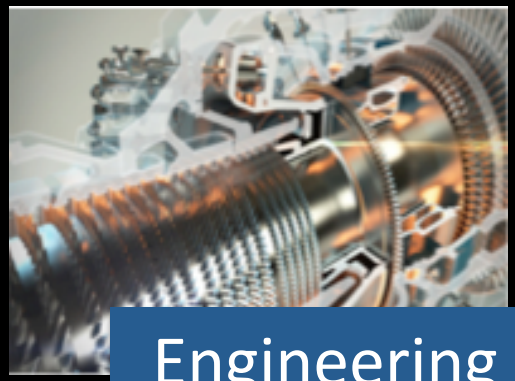
Part I

Science is poised for transformation

and Science in the broadest sense

Medicine

Basic Science

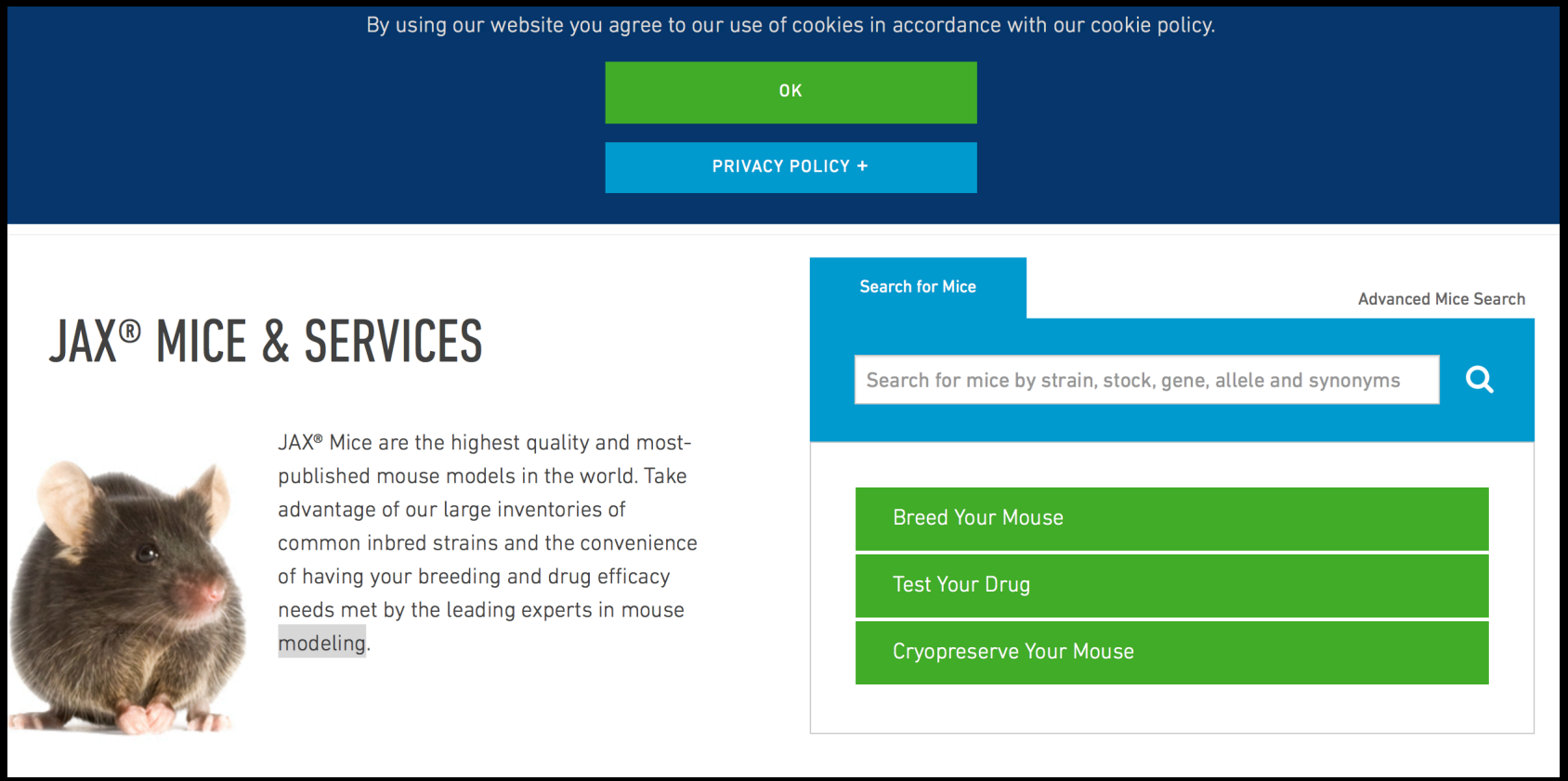


Engineering

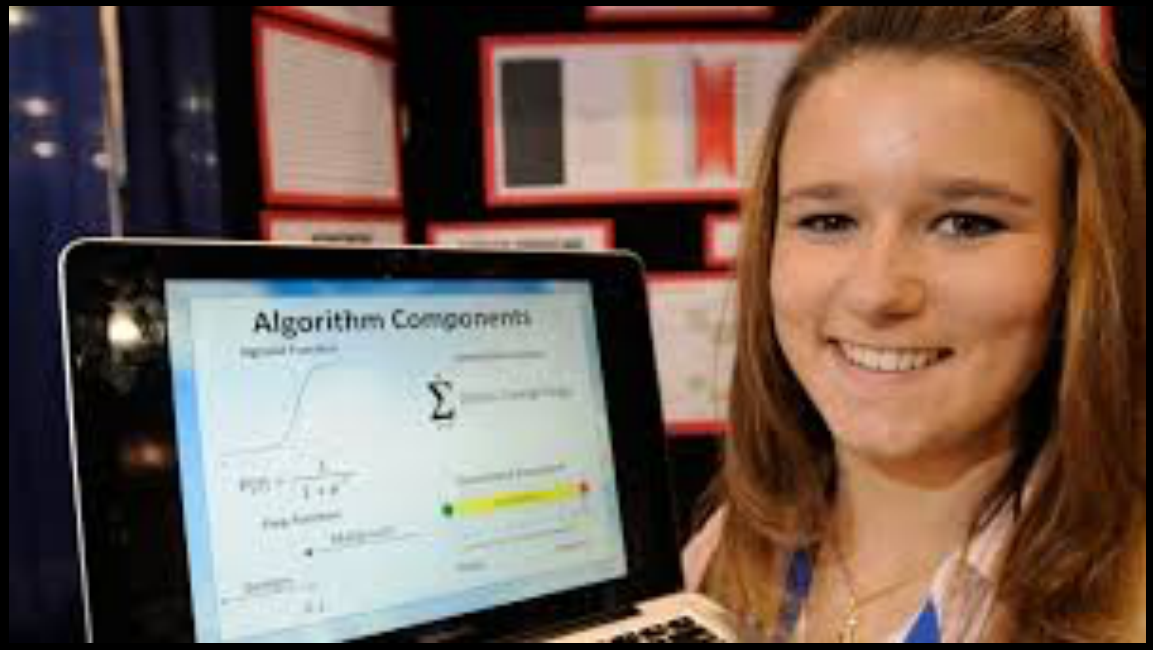
Environment

Security

Experimental Science is Changing



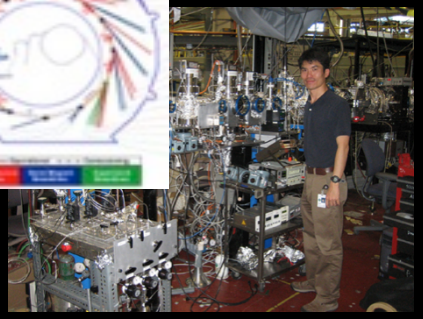
Empowering New Scientists

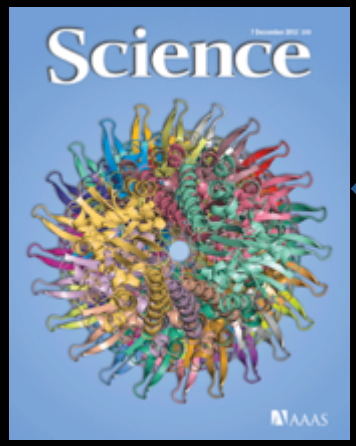


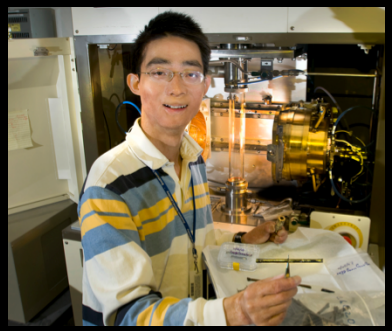
17-year-old Brittany Wegner creates breast cancer detection tool that is 99% accurate on a minimally invasive, previously inaccurate test.

Machine Learning + Online Data + Cloud Computing

Old School Experimental Science





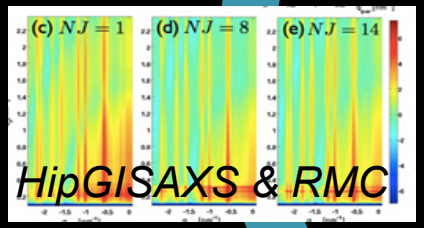


Computing, experiments, networking and expertise in a "Superfacility" for Science

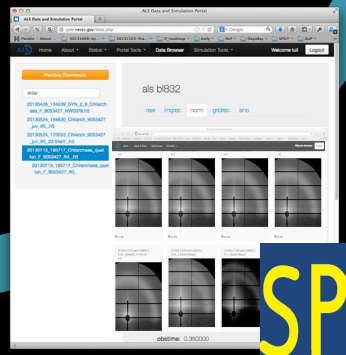
ALS ADVANCED LIGHT SOURCE

Special Purpose processing here?

ESnet

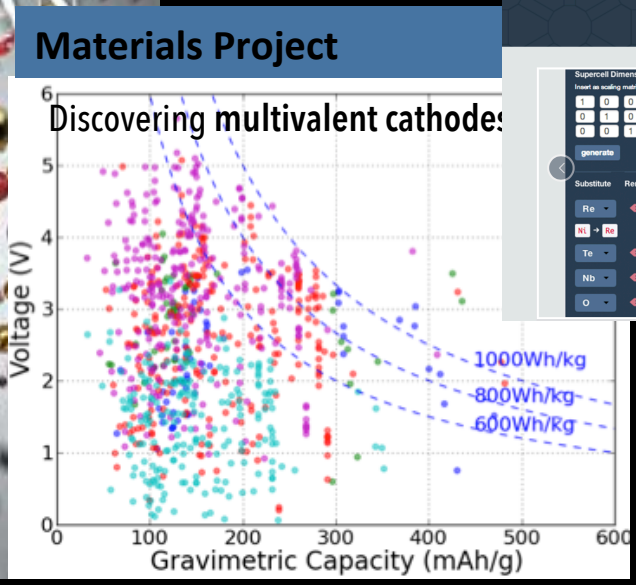


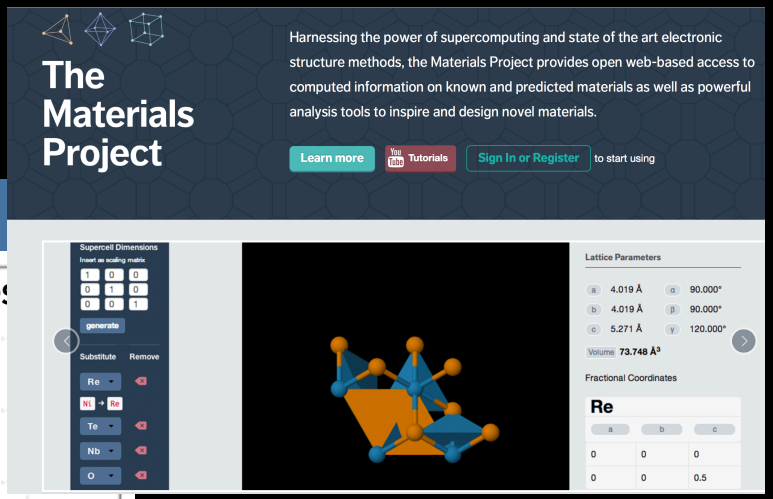
CAMERA
Applied Math



SPOT

Computational Scientists May Never Login

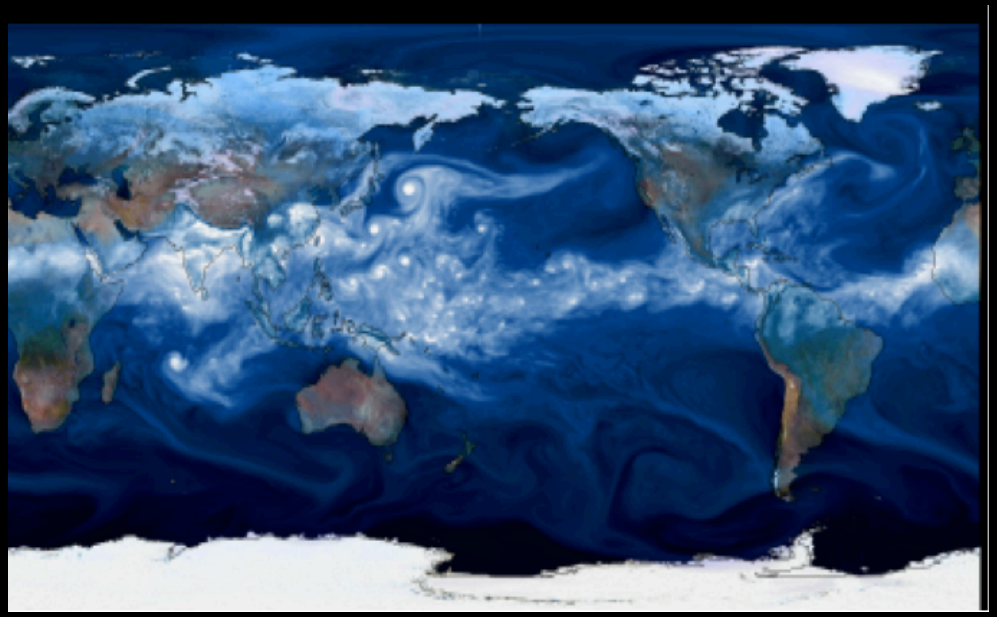


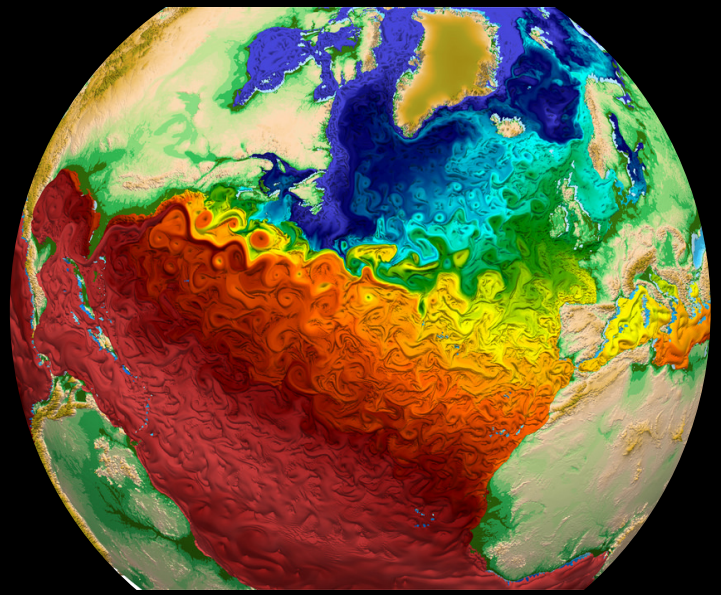


Novel computing technology here?

- Over 10,000 users of the Materials Project since 2011
- 1000x more computing needed for more materials, environments, and details

Simulations both consume and produce massive data



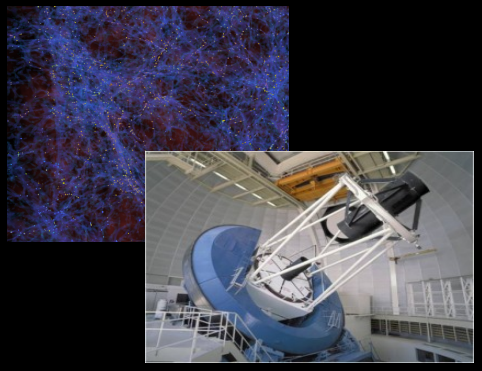


- **Simulation: Resolve clouds, identify tipping points, quantify confidence, quantify agriculture and economics**
- **Observation: Use environmental sensors and samples to refine models**
- **Analysis: Machine learning identifies extreme events; Reanalysis of 100+ years of weather to improve**

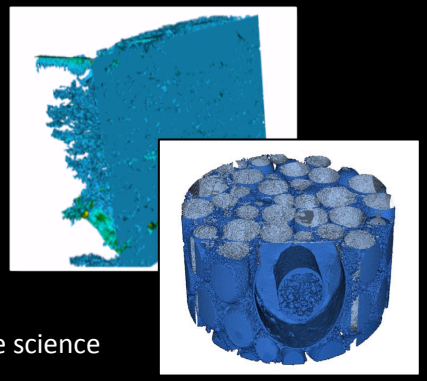
Breakthrough science will occur at the interface of observation and simulation

Materials and Chemistry

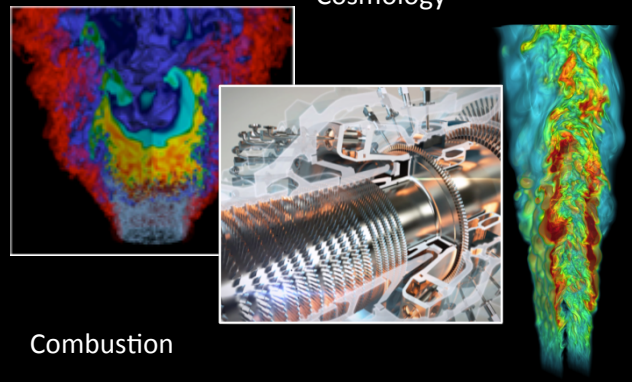
Environment and Climate



Cosmology



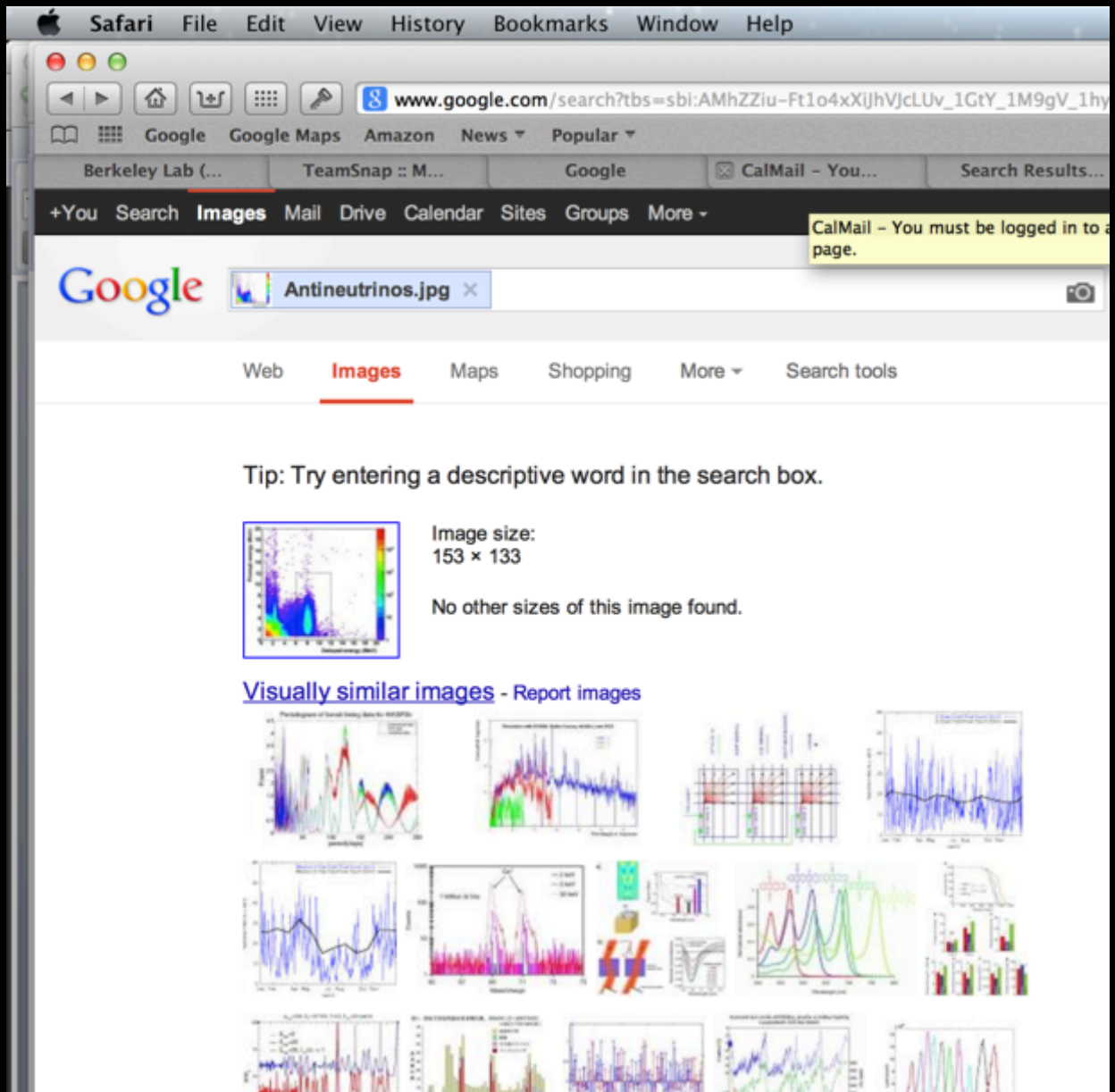
Subsurface science



Combustion

And it will need capable exascale systems, computer science, statistical machine learning and new mathematical models and algorithms

Challenge: Finding Scientific Data



Challenge: Filtering, De-Noising and Curating Data

AmeriFlux & FLUXNET: 750 users access carbon sensor data from 960 carbon flux data years

Arno Penzias and Robert Wilson discover Cosmic Microwave Background in 1965

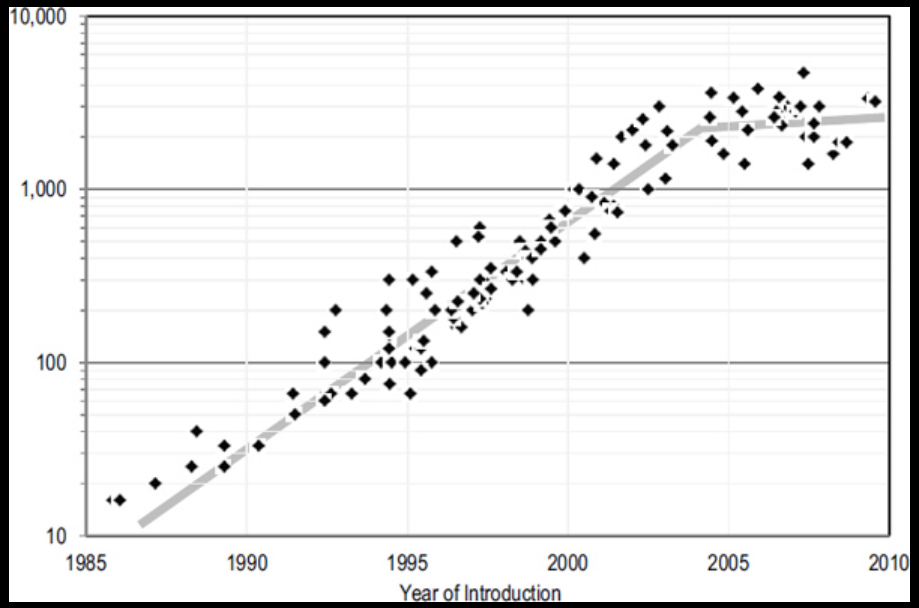
The Black Swans of Computing with 1990s Technology

Google™

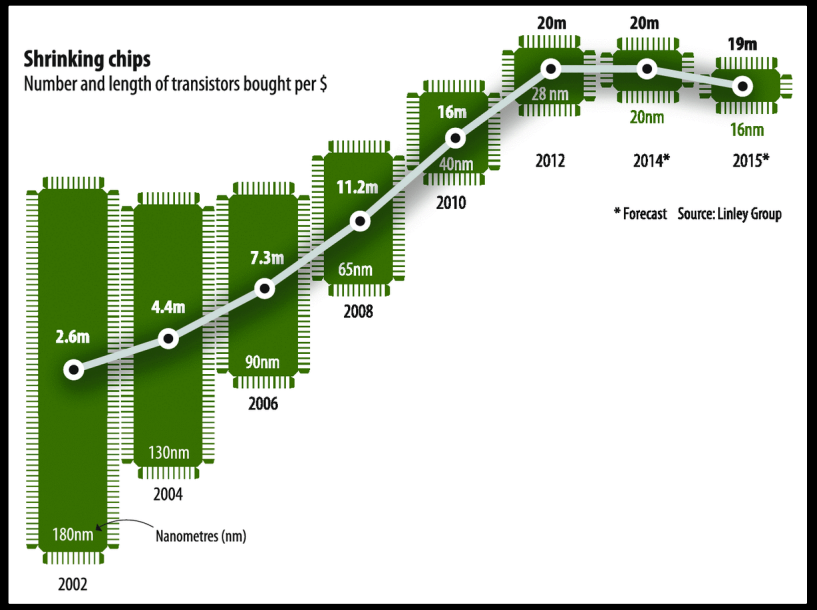
Part II

- Science is poised for transformation
- Computing performance growth is slowing

Computing Performance Stalls



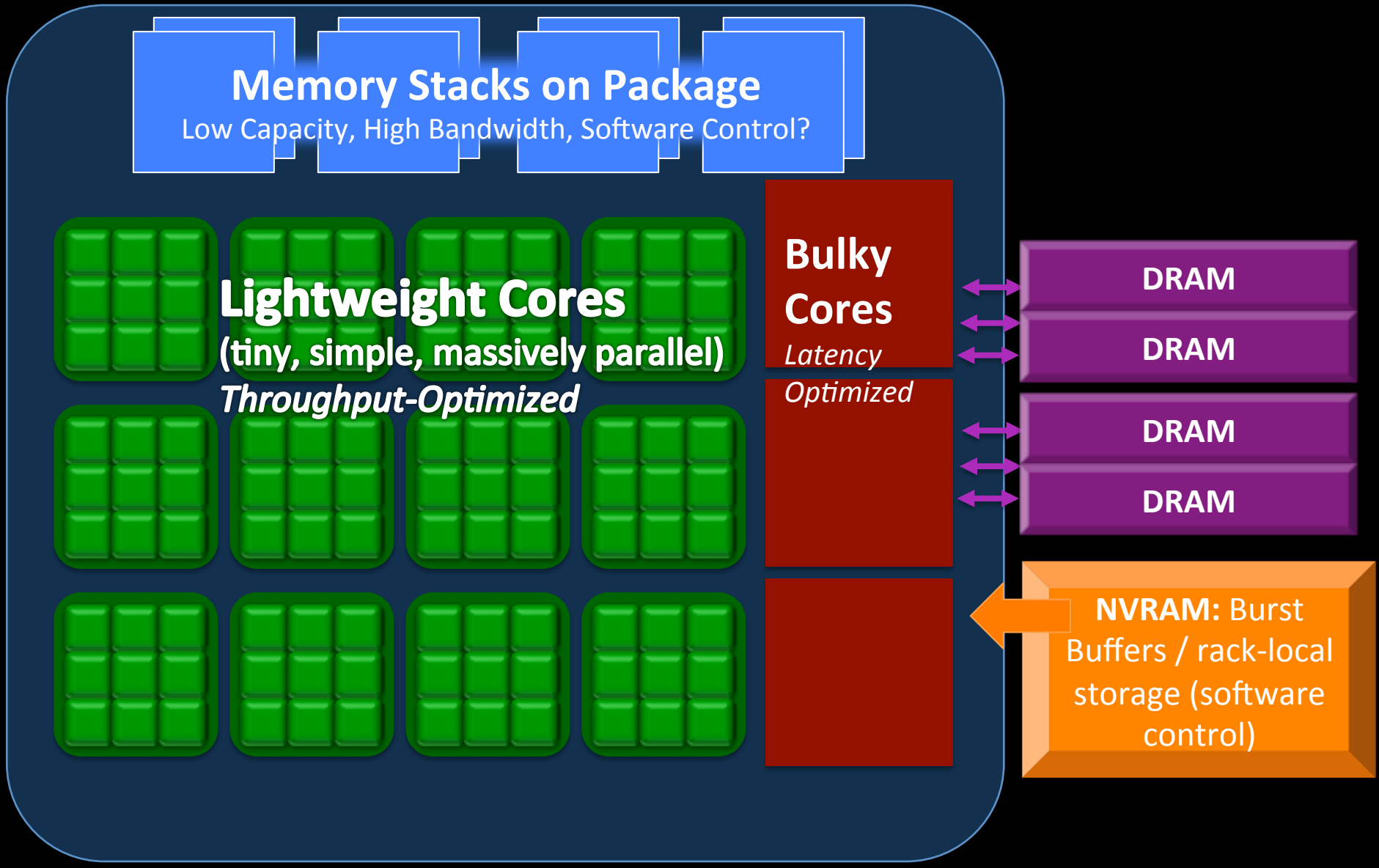
2004: End of clock speed scaling



202?: End of transistor density scaling

We need an end game for Moore's Law and a new game after that

A Generic 2020+ Node Architecture

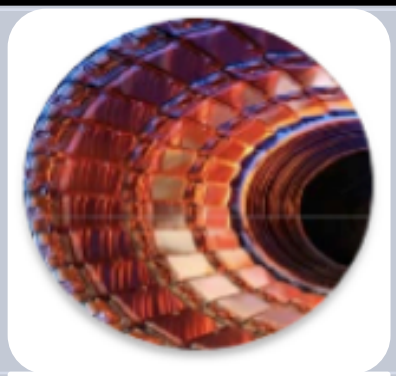


Part III

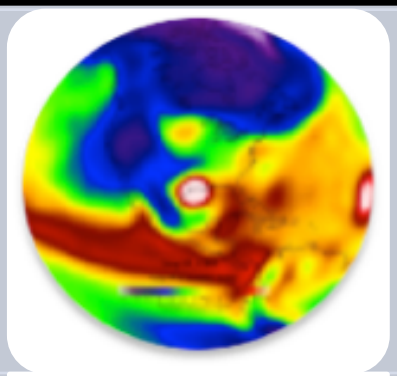
- Science is poised for transformation
- Computing performance growth is slowing
- How more productive tools can help

(What productivity means to me)

The right machine for the problem

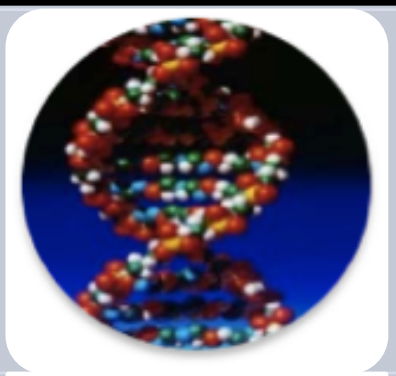


Massive Independent Jobs for Analysis and Simulations



Nearest Neighbor Simulations

All-to-All Simulations



Random access, large data Analysis

Applications often fit in multiple categories

The right systems software for the problem

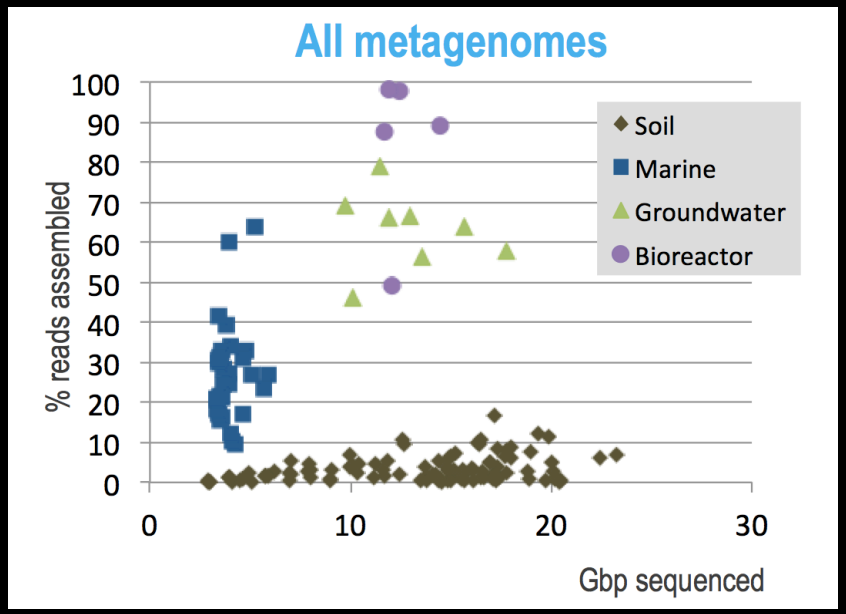
Cloud Computing
Map-Reduce w/ Resilience
Flexible software stack
On-demand scheduling
Price ≠ Cost

High Performance Computing
General programming
High speed networks
Co-scheduling large jobs
Algorithms and libraries
Performance expertise

Ideas and technologies have transferred in both directions

Ultimately, cloud computing is business model not a technology

A convenient programming model for the problem



De Novo Genome assembly for the environment and health

"Scaffolding" step on shared memory

- Human: 44 hours to 20 secs
- Wheat: "doesn't run" to 32 secs

All 2015 human sequences could have been assembled on 10% of annual NERSC resources; metagenomes are much harder

A programming model (UPC) that encourages the right algorithmic thinking (hash tables, atomics,…)

A Network that Makes Location Irrelevant

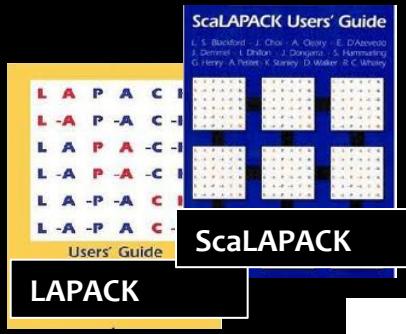
Technology, policy, and social challenges across:

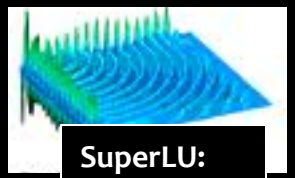
- Networks and countries
- Facilities
- Universities (Science DMZ)
- Systems (SDN)

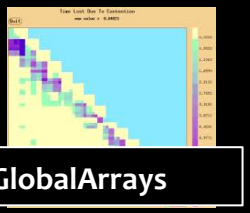
ESnet Goal: Scientific progress **completely unconstrained by the location** of people, instruments, computational resources, or data.

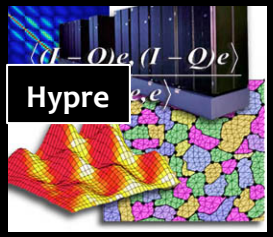
Use scientific libraries and frameworks

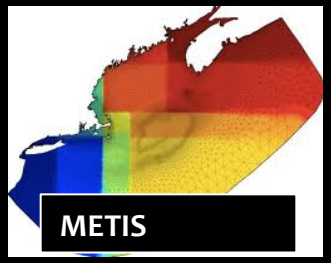
- *Encapsulate the best algorithms*
- *Hide machine details*

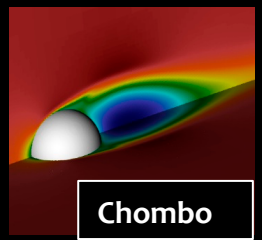


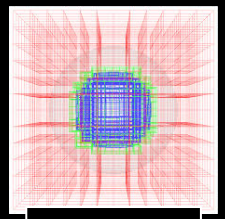


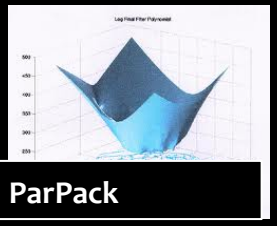


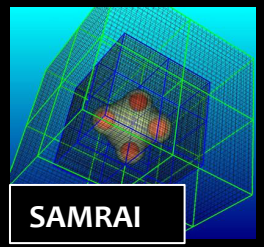










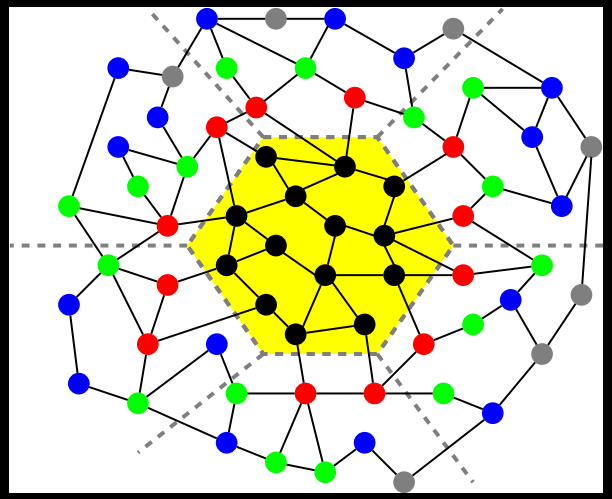


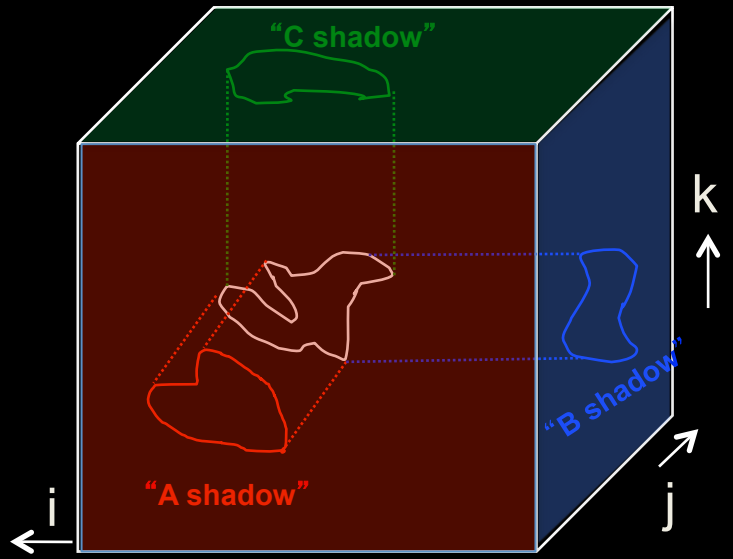
Enabled by portable programming models like MPI

The right libraries for the problem

7 Motifs of Simulation	7 Motifs of Data
Monte Carlo methods	Basic statistics
Particle methods	Generalized N-Body
Unstructured meshes	Graph-theory
Dense Linear Algebra	Linear algebra
Sparse Linear Algebra	Optimizations
Spectral methods	Integrations
Structured Meshes	Alignment

The right algorithms, models and methods





Optimize for data movement as first-class concern

- Change numerical methods (e.g., high order)
- Design algorithms that minimize data movement

High level programming abstractions

Where is Performance Portability?

- **Multiple generations of GPUs, Xeon Phi, …**
  - Not performance portable across systems
- **How do applications address this problem?**

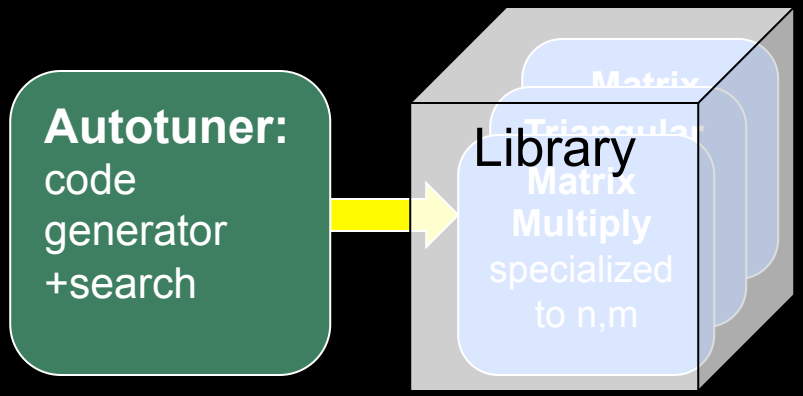
#ifdef

Approach 1: Use of libraries, templates and preprocessors

Approach 2: OpenMP4 absorbs necessary accelerator features (likely), but code still diverges

Approach 3: Architectures converge “enough” by 2023, perhaps with co-design help

Write code generators and autotuners



What code generators do we have?

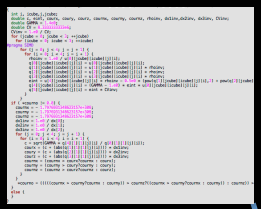
Dense Linear Algebra	**Atlas**
Spectral Algorithms	**FFTW, Spiral**
Particle	**TCE**
Sparse Linear Algebra	**OSKI**
Structured grids	**TBD**
Unstructured Grids	
Monte Carlo	

Two unsolved compiler problems

- Dependence analysis
- Optimization selection

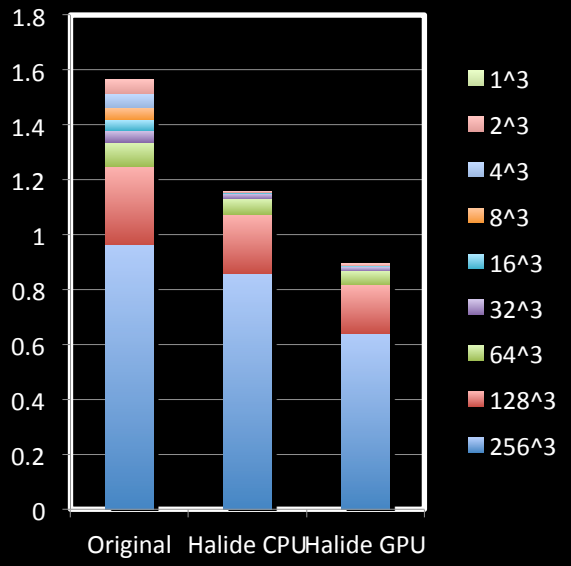
Develop and influence languages and compilers

Rose compiler input

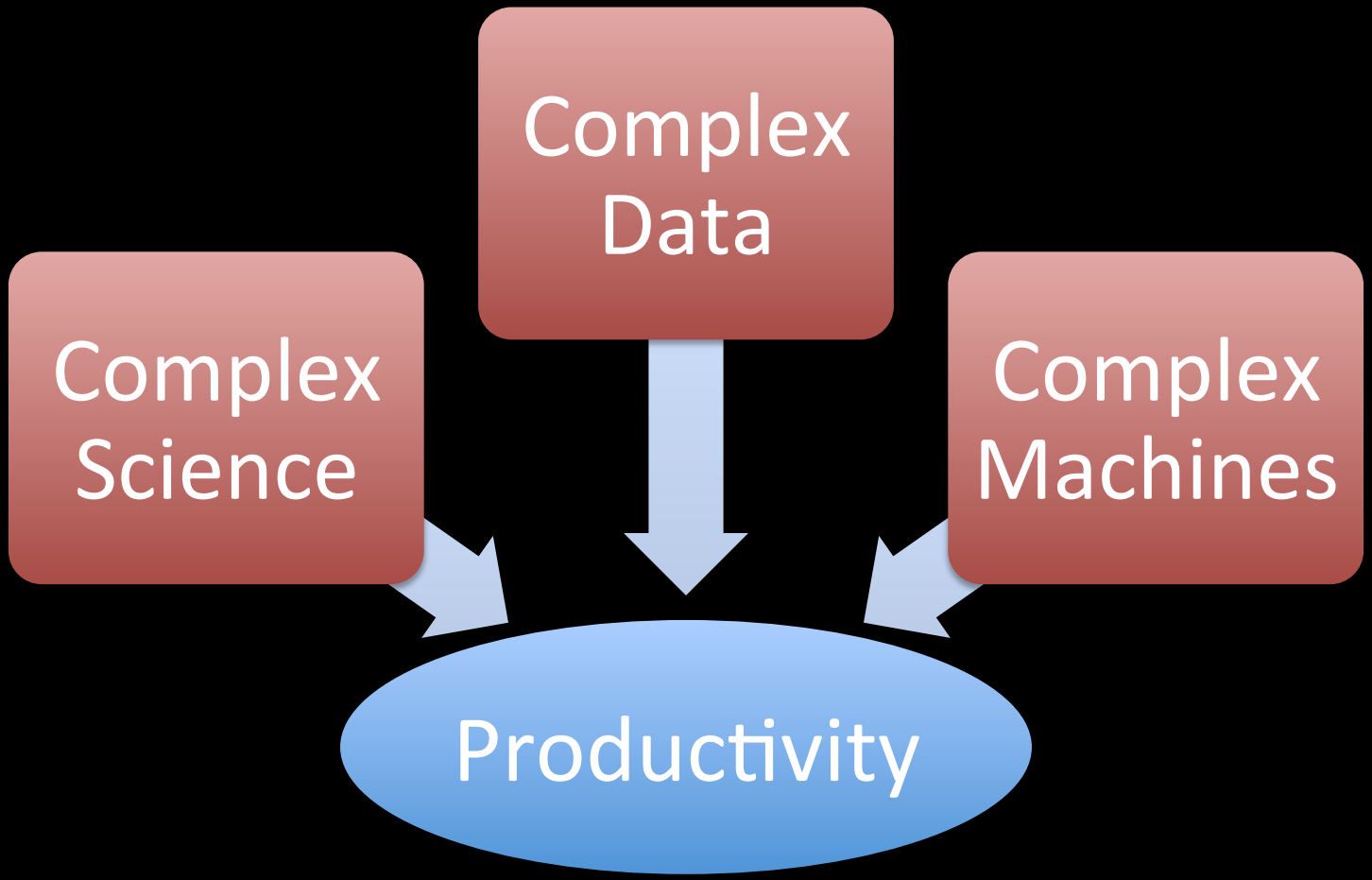


Rose compiler output

Halide DSL has 10+ developers and 50+ users in industry



Productivity will be key to dealing with complexity



Conclusions

NSCI is not:

- limited to the high end
- about integrating enough chips to reach exaflop

It is about:

- The cost of computing
- Power requirements, size, and speed of computers
- Solving real problems that require more computing