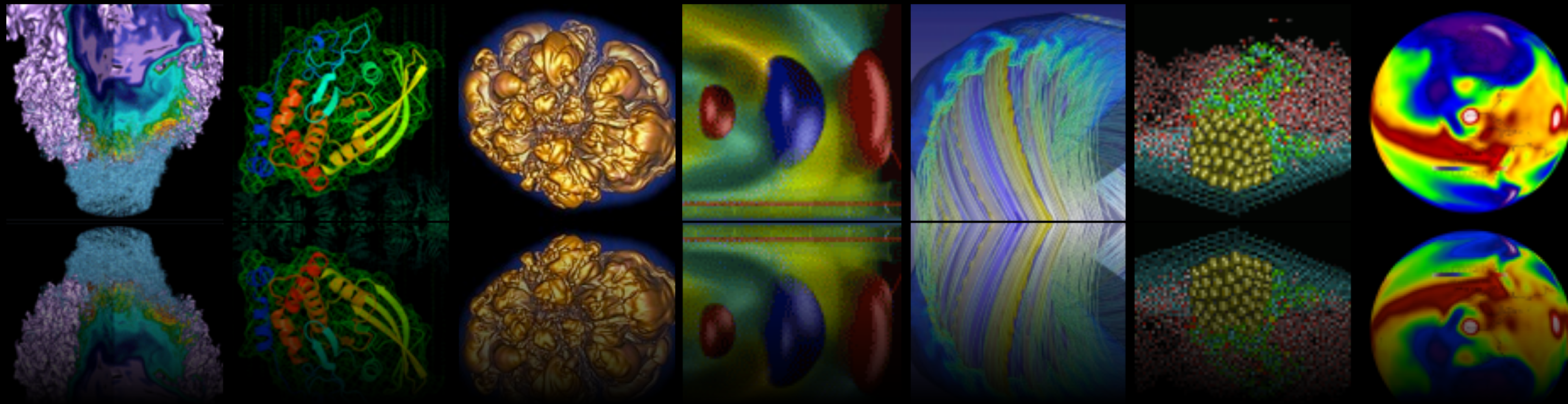


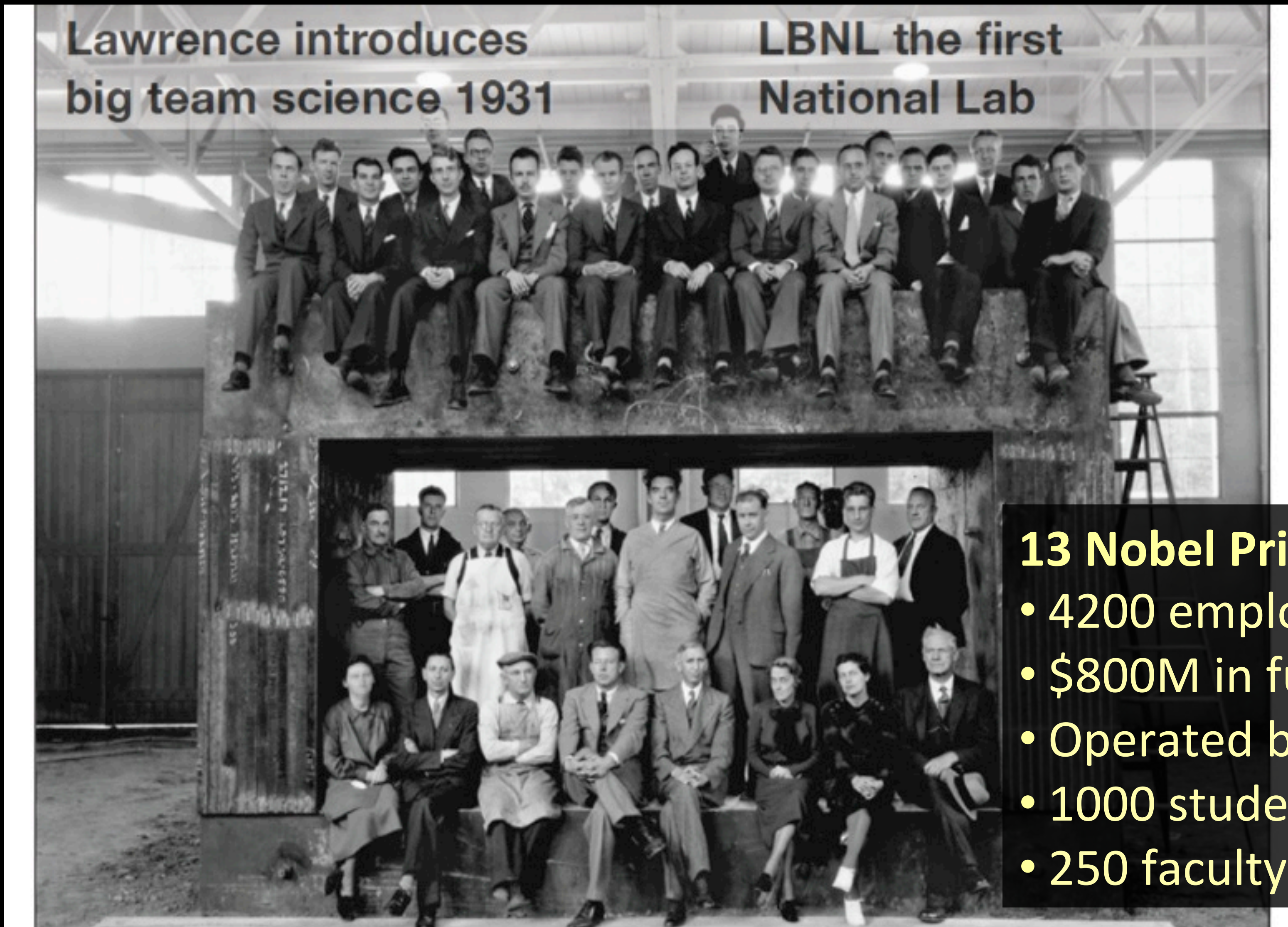
More Data, More Science and… Moore's Law

Kathy Yelick

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

“Big Data” Changes Everything…What about Science?

The Essence of Berkeley Lab: Team Science

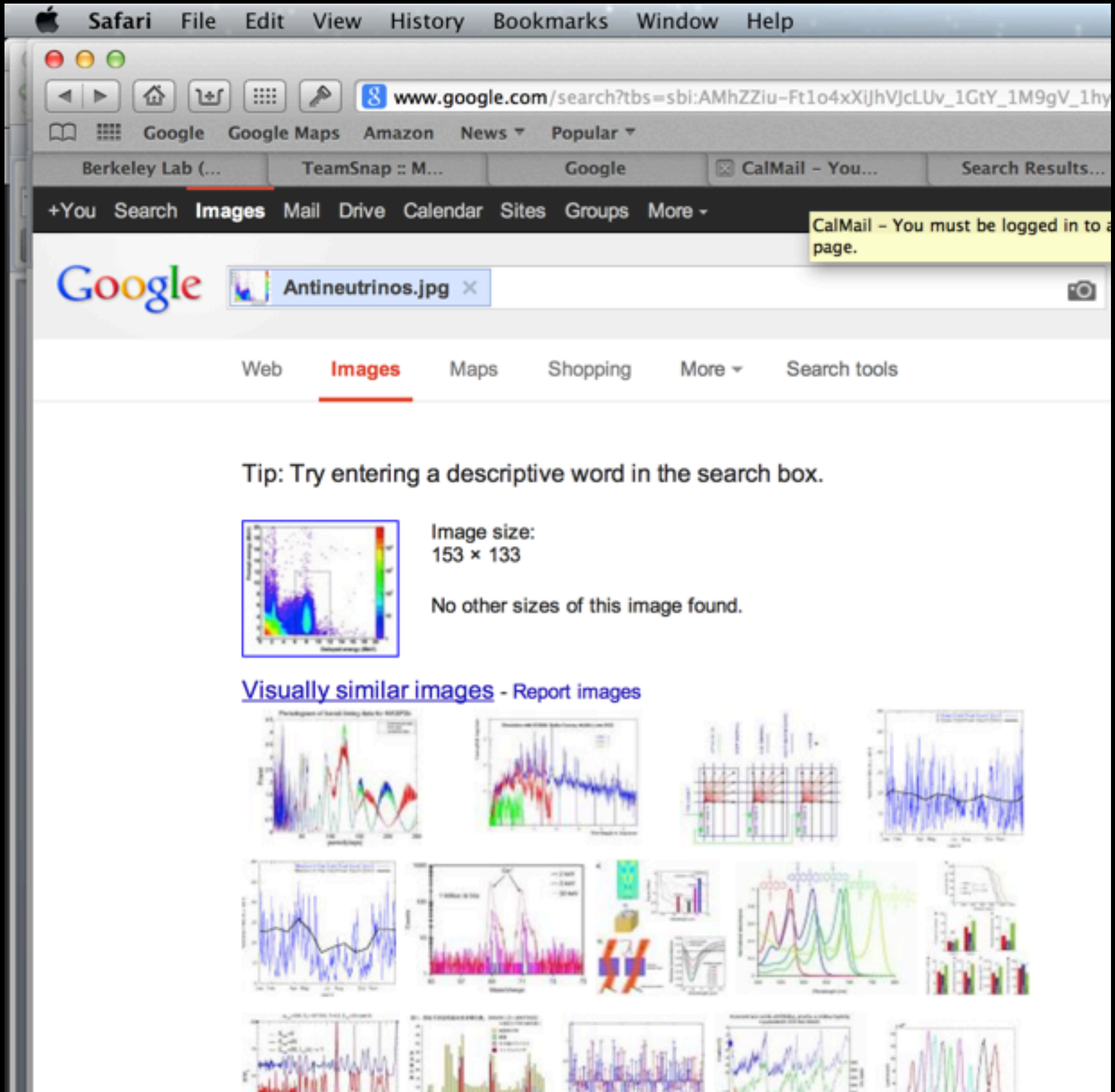


13 Nobel Prizes
- 4200 employees
- $800M in funding
- Operated by UC
- 1000 student
- 250 faculty

Berkeley Lab's Advanced Facilities Enable World-Leading Science

About 10,000 visiting scientists (~2/3 from universities) use Berkeley Lab research facilities each year, which provide some of the world's most advanced capabilities in materials science, biological research, computation and networking

Transforming Science: Finding Data



Scientific Workflow Today

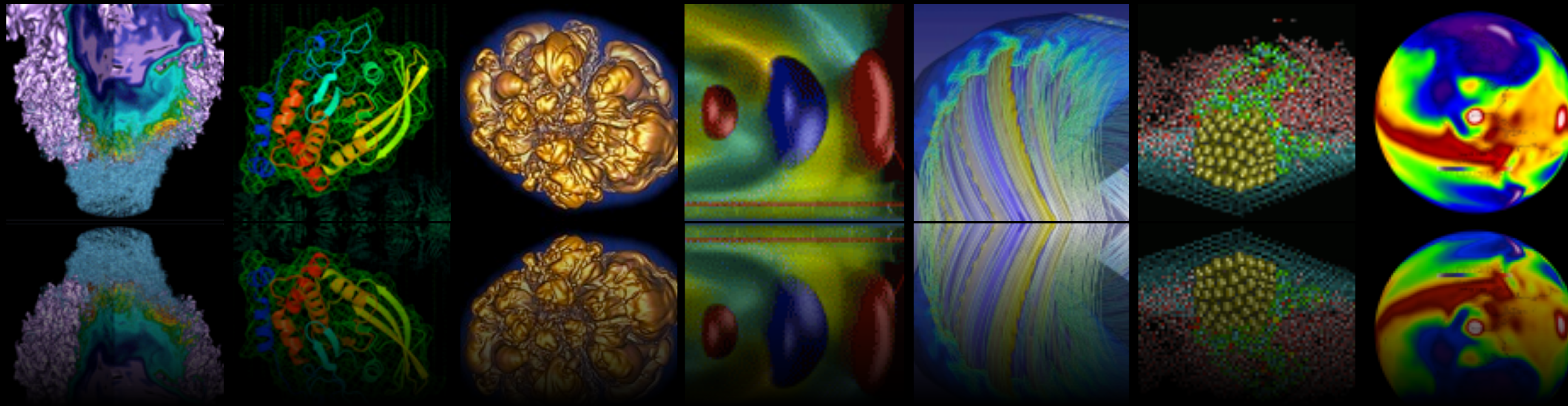
KATHY YELICK'S

2031:

a science odyssey

Life of a Scientist in 2031

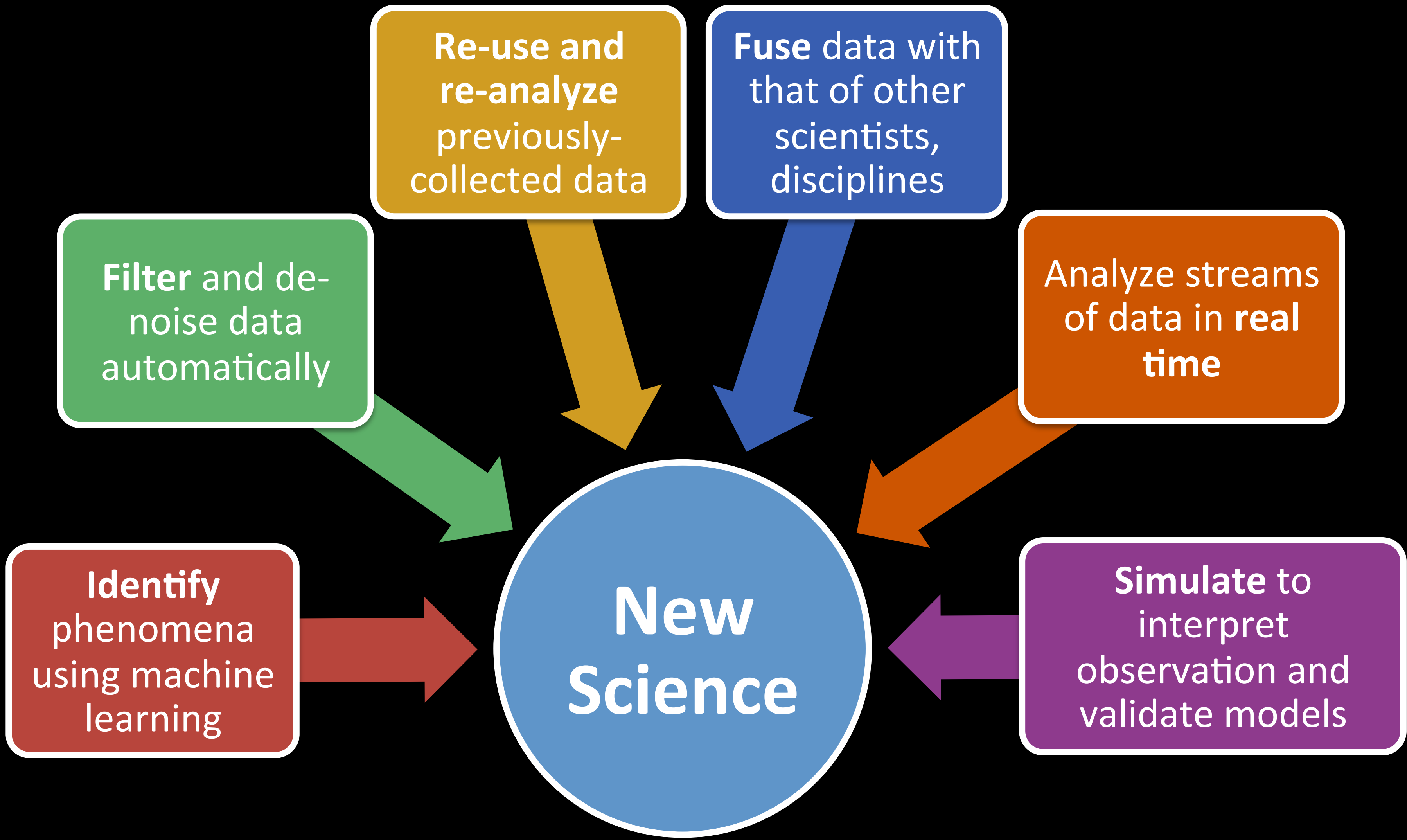
- No personal/departmental computers
- Users don't login to HPC Facilities
- Travel replaced by telepresence
- Lecturers teach millions of students
- Theorems proven by online communities
- Laboratory work is outsourced
- Experimental facilities are used remotely
- All scientific data is (eventually) open
- Big science and team science democratized



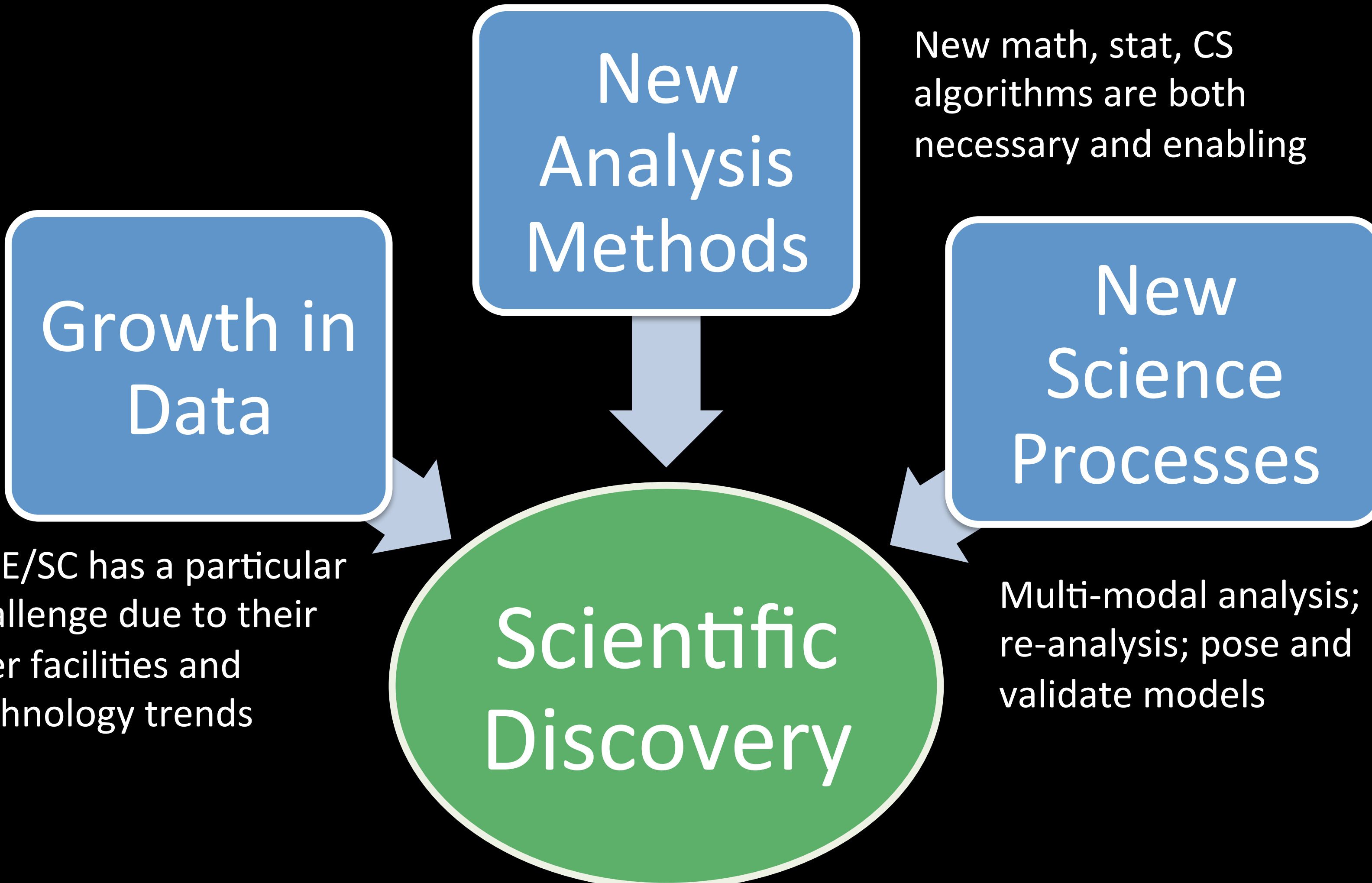
Extreme Data Science

The scientific process is poised to undergo a radical transformation based on the ability to access, analyze, simulate and combine large and complex data sets.

New Models of Discovery



Goal: To enable new modes of scientific discovery



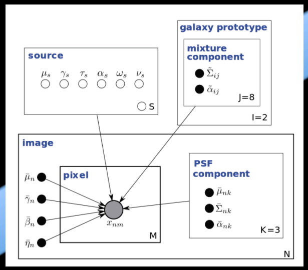
New math, stat, CS algorithms are both necessary and enabling

DOE/SC has a particular challenge due to their user facilities and technology trends

Multi-modal analysis; re-analysis; pose and validate models

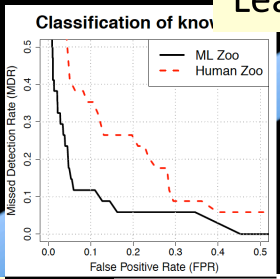
Cosmology: Machine Learning to Systematics

Graphical models



Machine Learning

New simulation models and AMR code (Nyx)

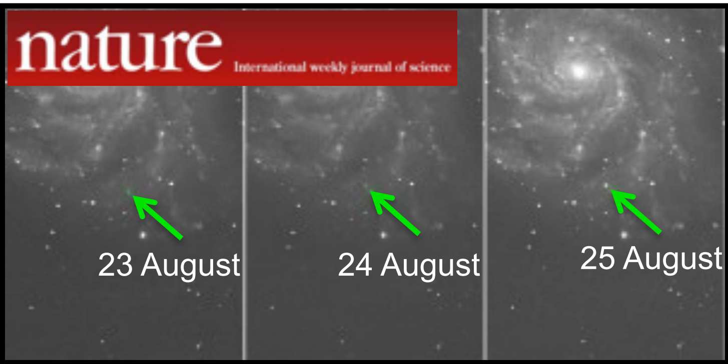


Crowd sourced

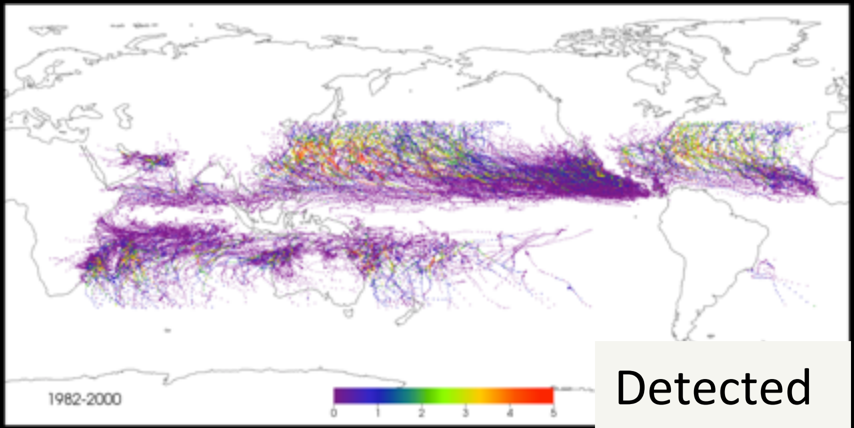
Filtered

GB per night
Manually analyzed

Example: Astrophysicists discover early nearby supernova



Climate: Machine Learning and Filtering



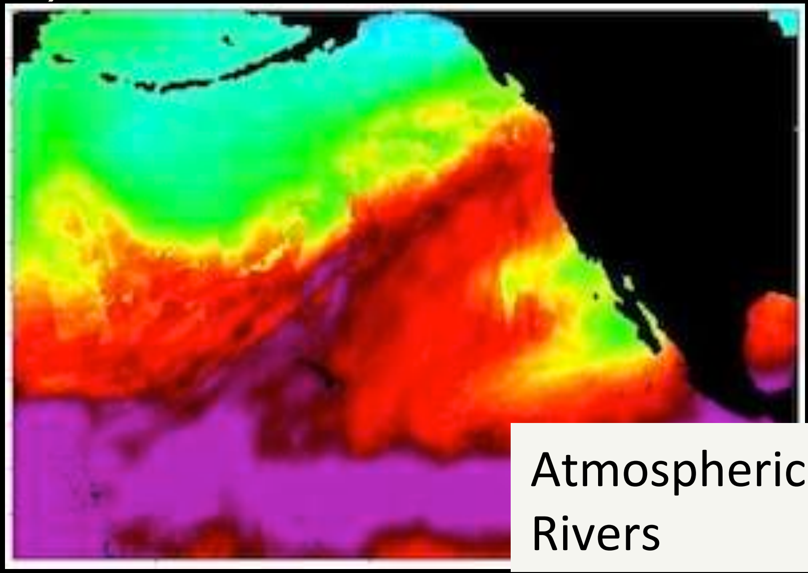
Detected cyclones

TECA Toolkit

- Automatic detection of cyclones, atmospheric rivers, and more
- Single data set is 100 TB
- Scalable analysis (80K cores): 9 years → 1 hour

Ongoing work

- Pattern detection using machine learning



Atmospheric Rivers

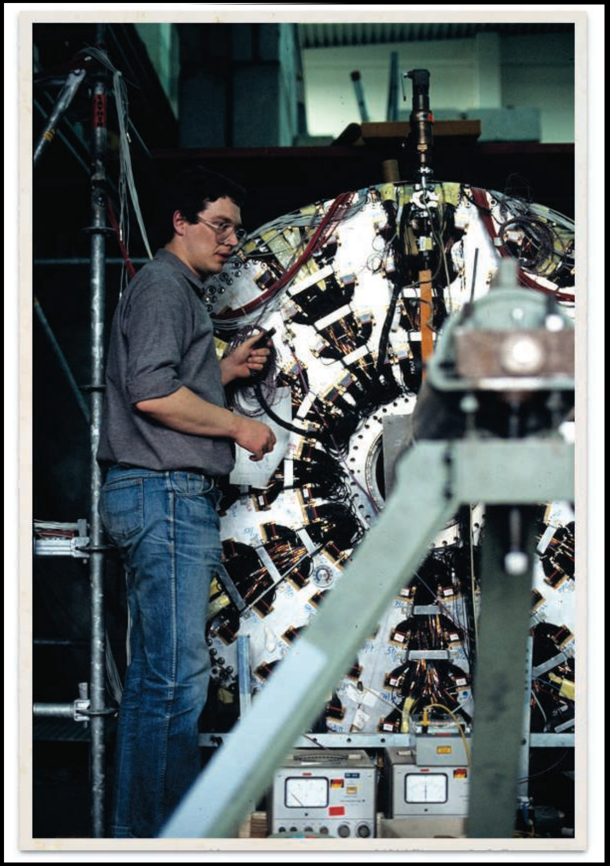
Mantissa Project, Prabhat

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

Particle Physics: Re-Use and Reanalyze



Sigfried Bethke working on the JADE detector in 1984.

- 1986: JADE experiment at DESY closed
- 1996: new theories could be validated with low energy data from JADE
- The struggle to recover the data…
  took 2 years, including:
  - A student working 1 year recreating code
  - 4 weeks of manual data entry from printouts
  - A lot of **luck**

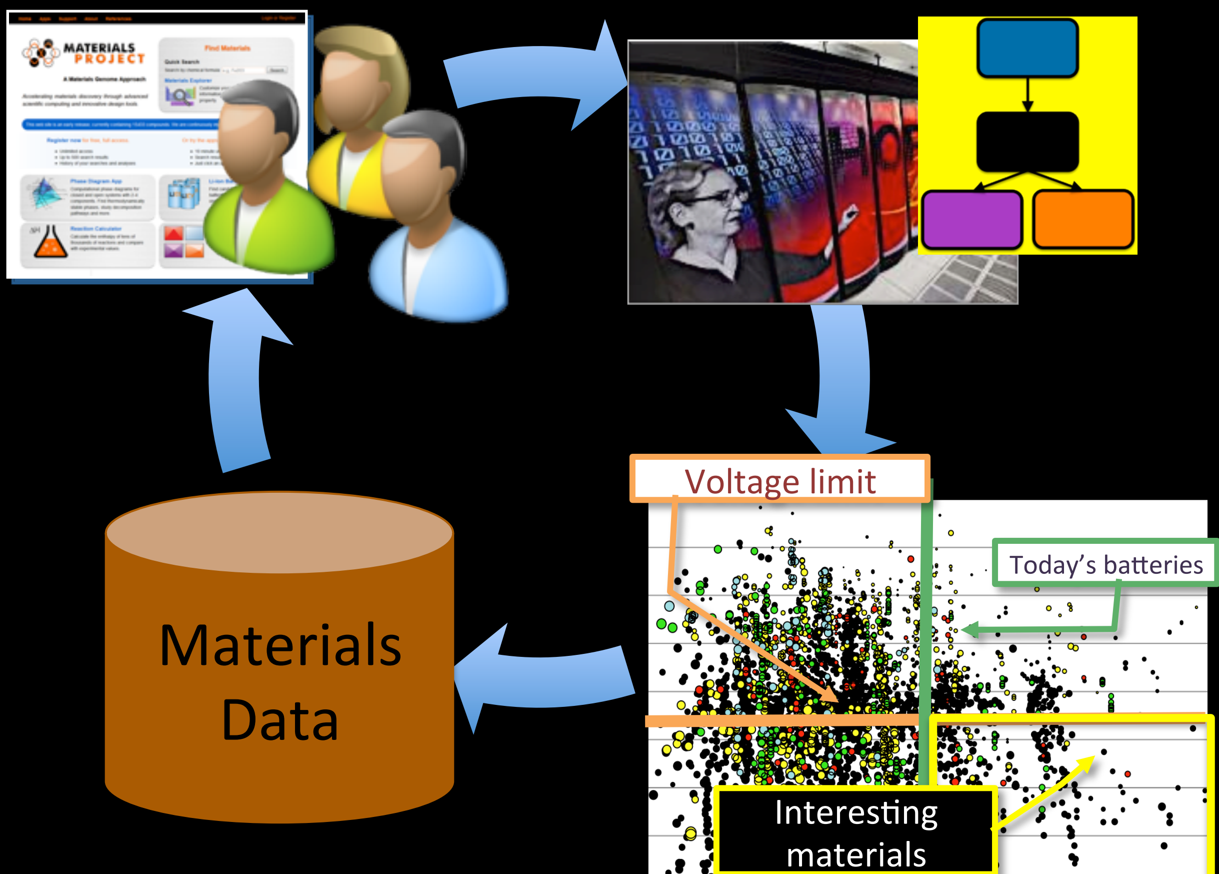
Result:

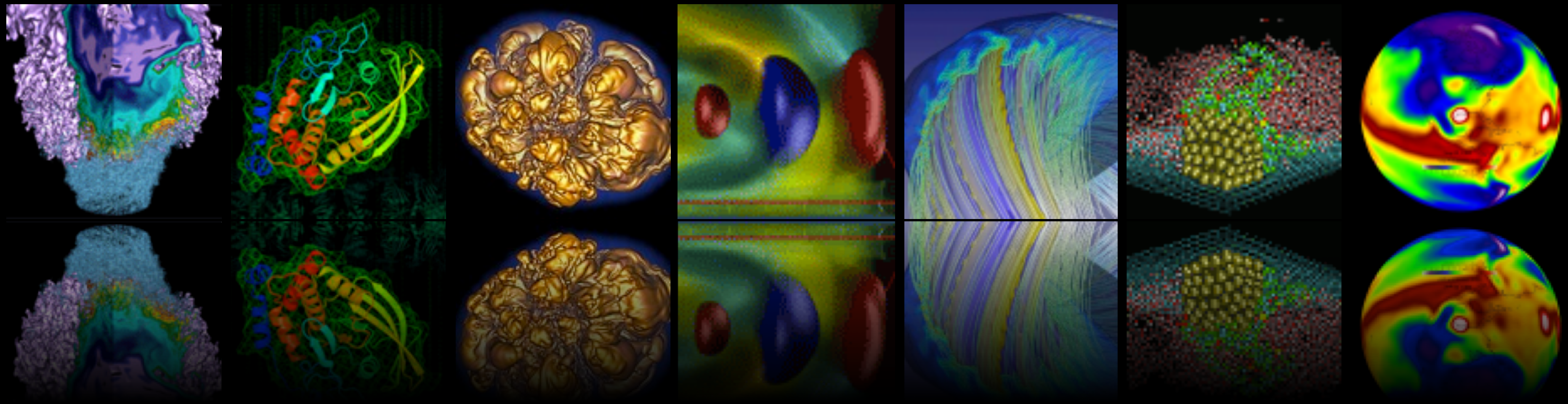
- A dozen high-impact publications.
- Cited by 2004 Physics Nobel committee

Curry, A. Rescue of old data offers lesson for particle physicists. *Science* (New York, N.Y.) 331, 694-5(2011).
Slide source: Filip Maia

Materials Model: Re-Use, Re-Analyze and Fuse

- **Materials Genome Initiative**
  - Materials Project: 4500 users 18 months!
  - “World-Changing Idea of 2013”

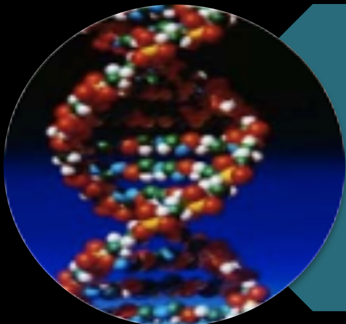




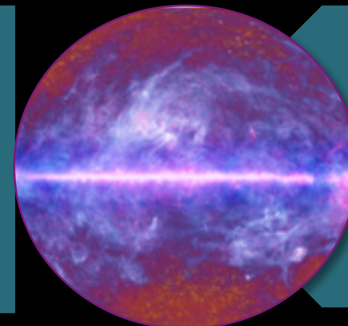
Science Data is Big (and Growing)

"Big Data" Challenges in Science

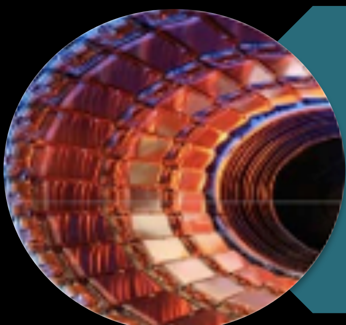
Volume, velocity, variety, and veracity



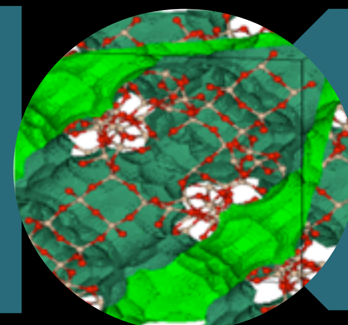
Biology
- *Volume:* Petabytes now; computation-limited
- *Variety*: multi-modal analysis on bioimages



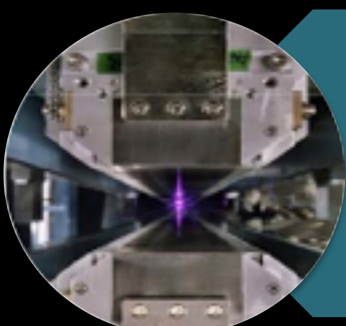
Cosmology & Astronomy:
- *Volume:* 1000x increase every 15 years
- *Variety:* combine data sources for accuracy



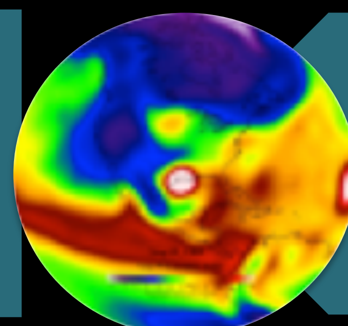
High Energy Physics
- *Volume*: 3-5x in 5 years
- *Velocity*: real-time filtering adapts to intended observation



Materials:
- *Variety:* multiple models and experimental data
- *Veracity:* quality and resolution of simulations

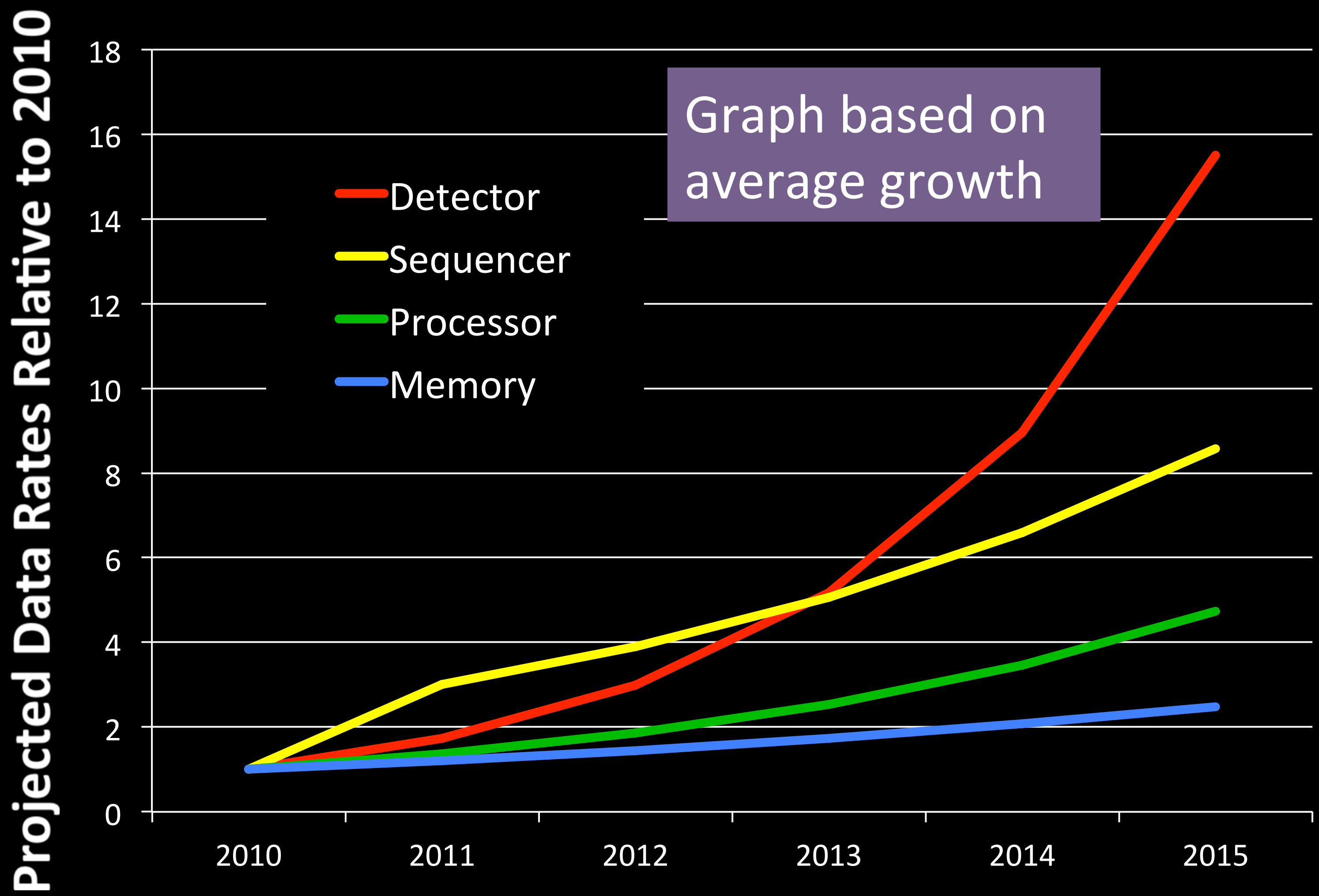


Light Sources
- *Velocity:* CCDs outpacing Moore's Law
- *Veracity:* noisy data for 3D reconstruction

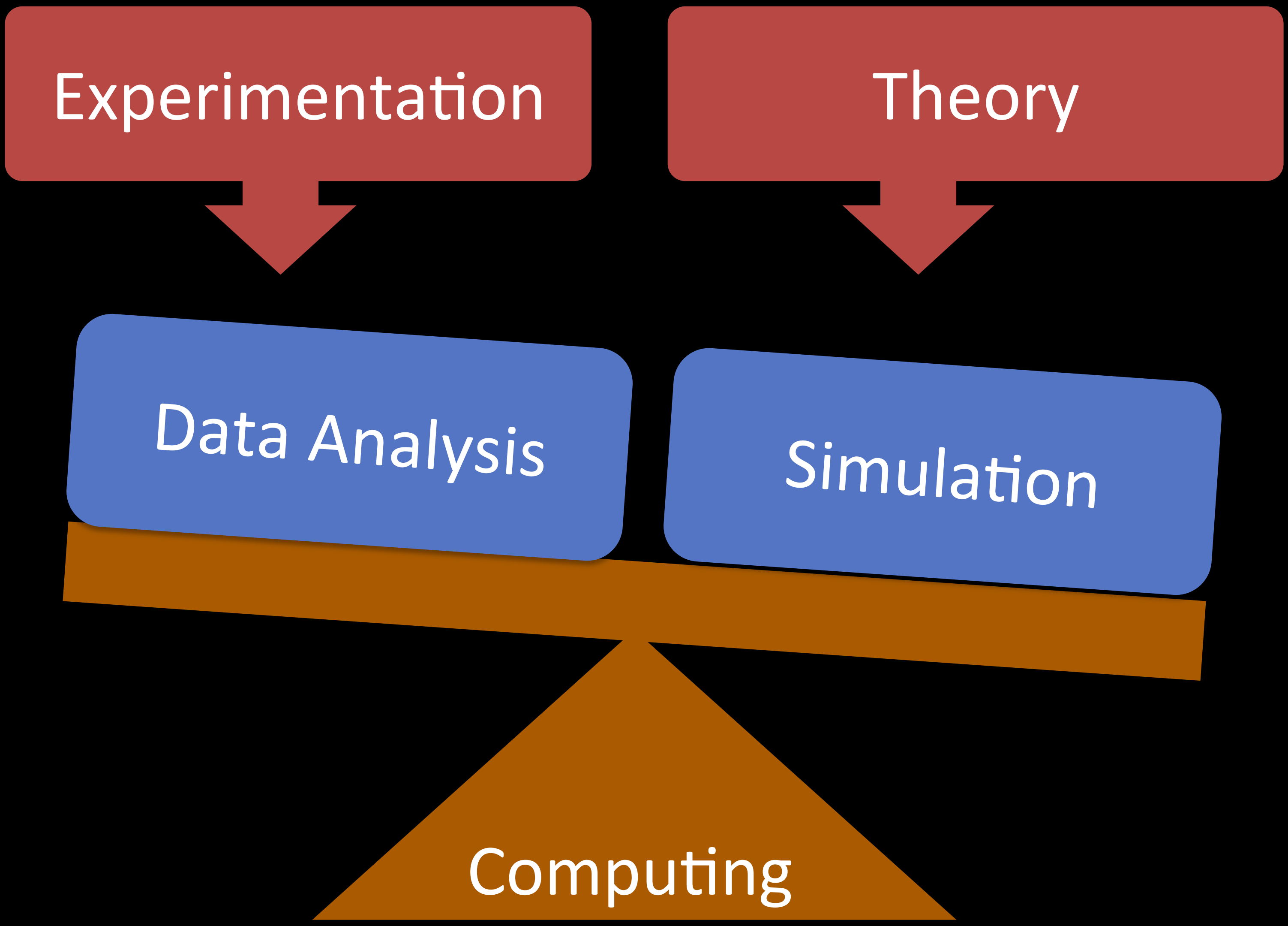


Climate
- *Volume:* Hundreds of exabytes by 2020
- *Veracity:* Reanalysis of 100-year-old sparse data

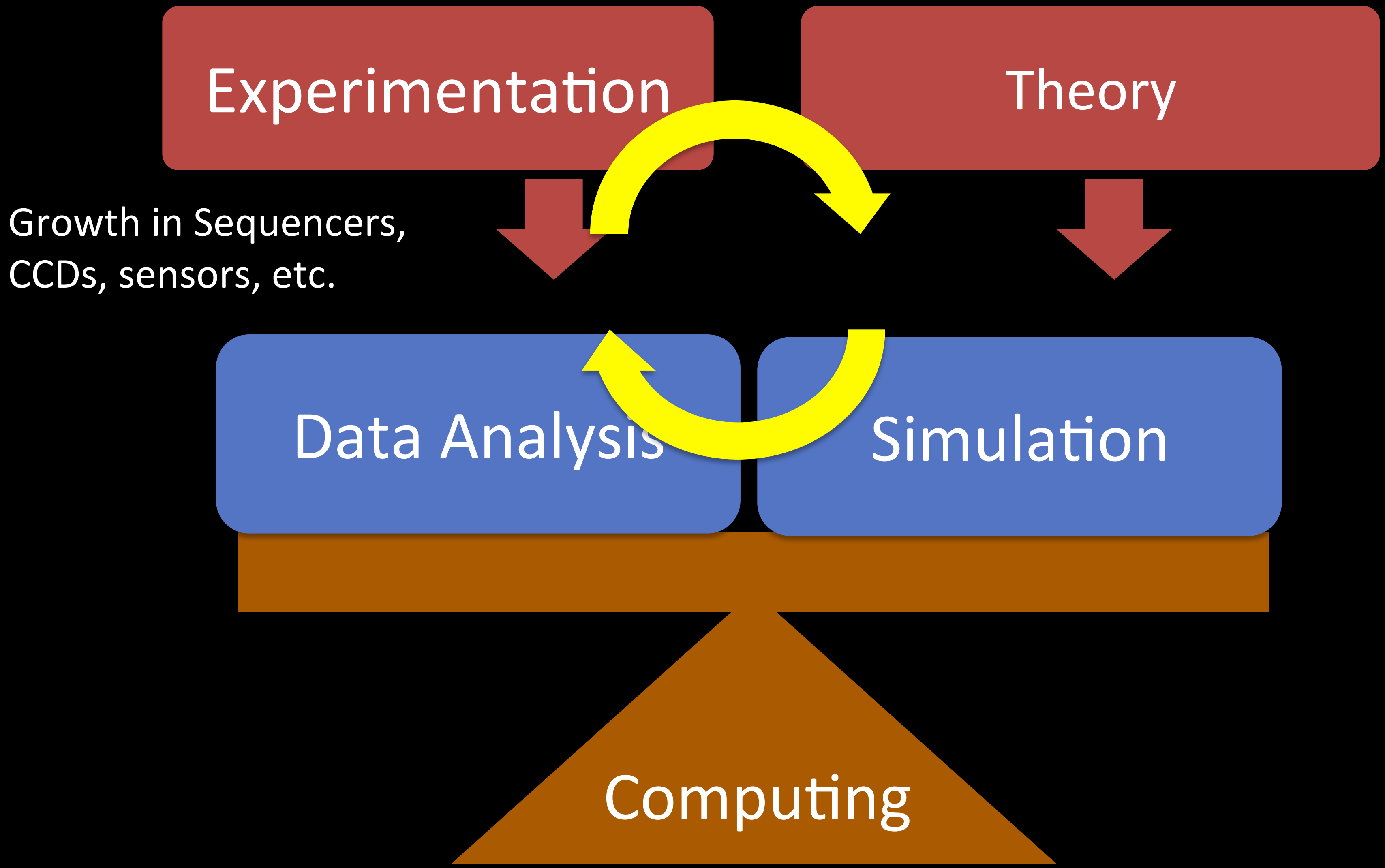
Data Growth is Outpacing Computing Growth

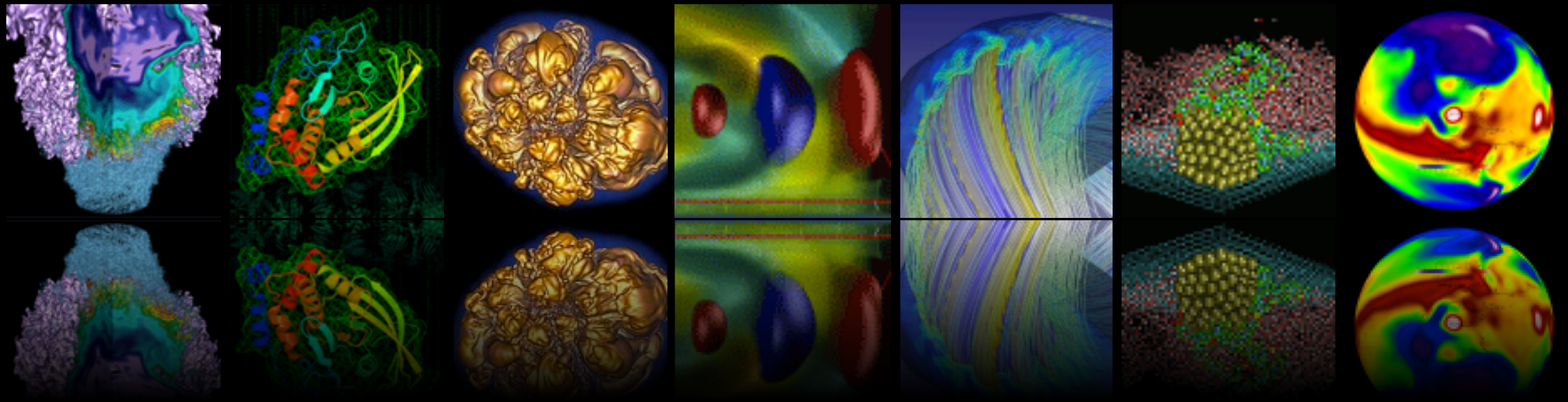


High end computing has focused on simulation



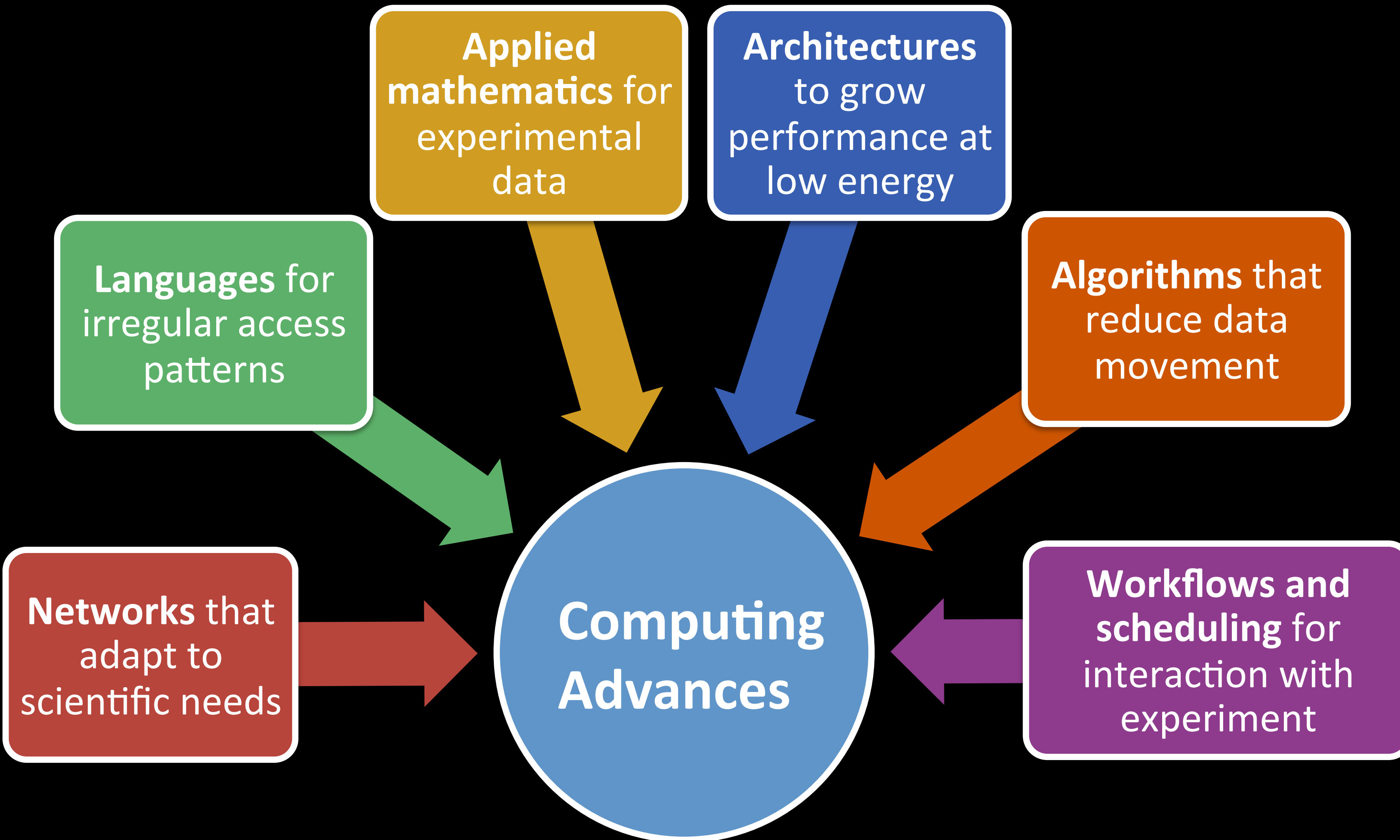
Data analysis is equally important in Science



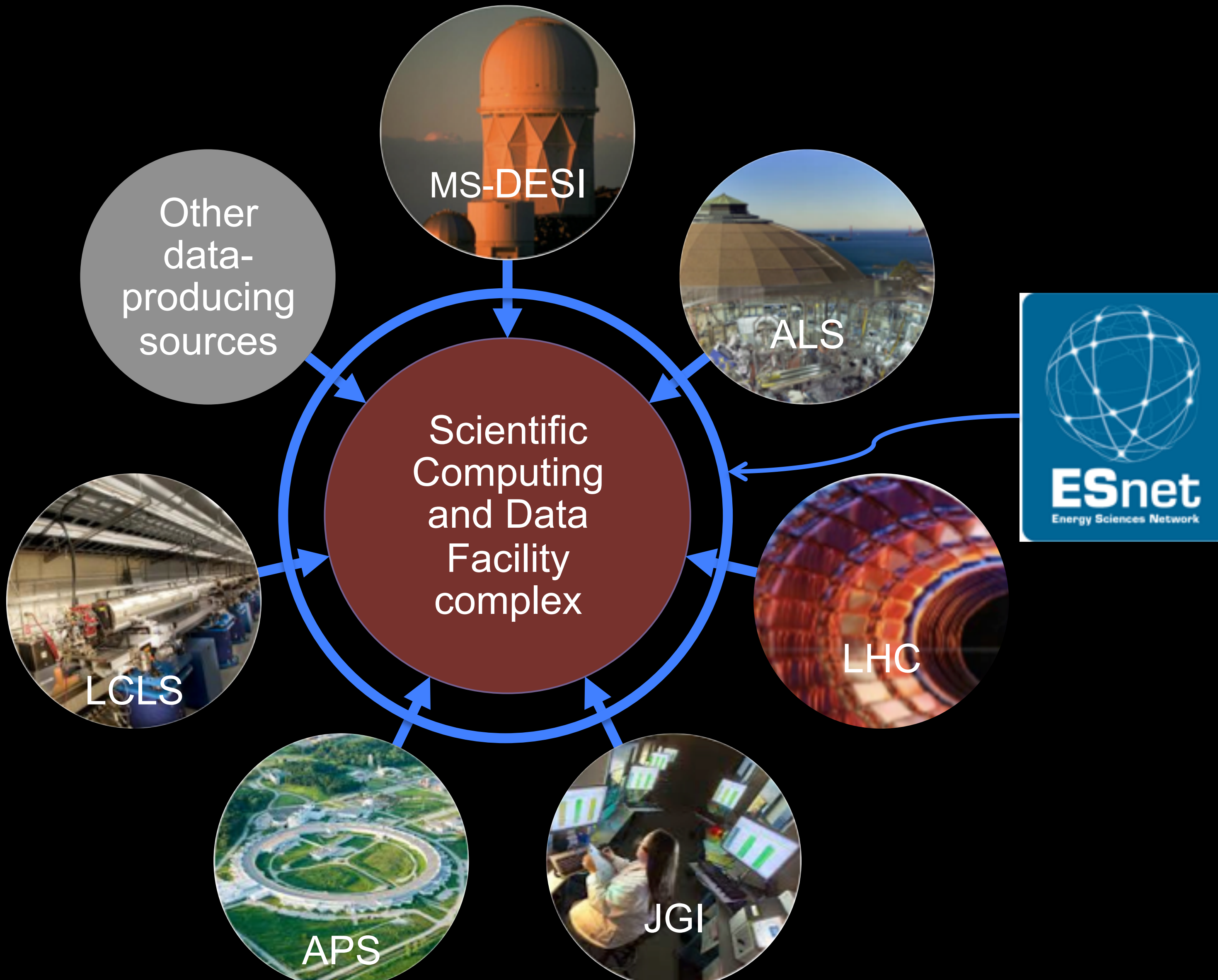


Are there Computer Science Research Challenges?

Computational Challenges

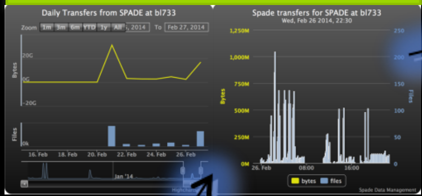


“Super Facility” Concept

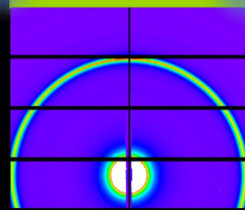


Light Source: Envisioning a "Super Facility"

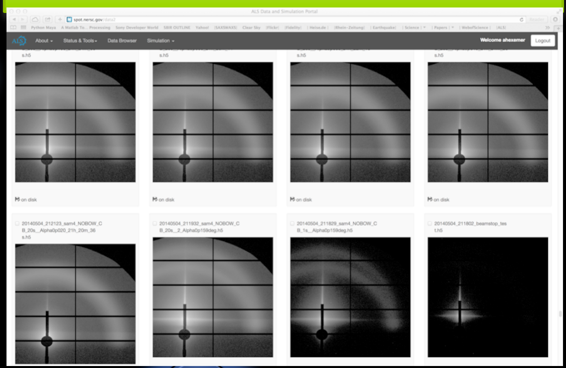
Transfer to NERSC



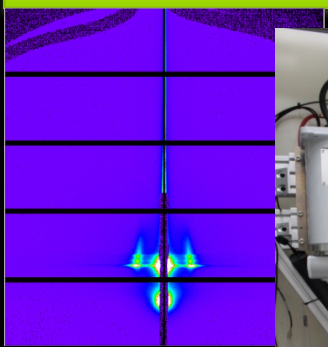
On-the-fly calibration, processing

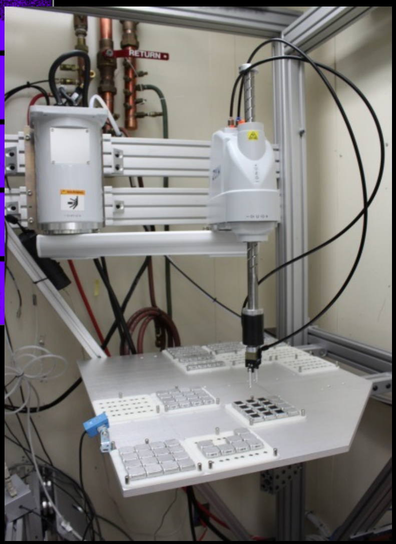


SPOT Suite: Real-time access via web portal



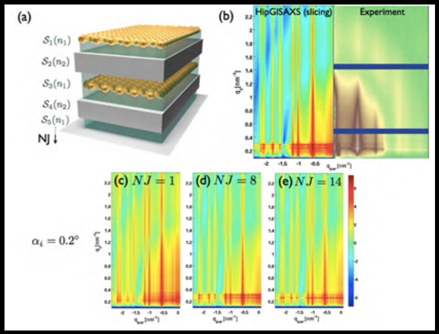
Data collection with robots and high resolution detectors (GISAX)

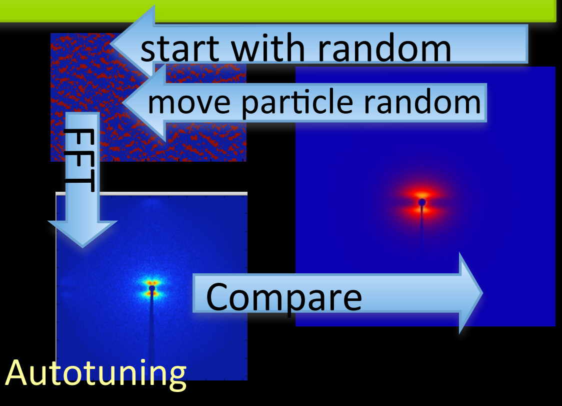




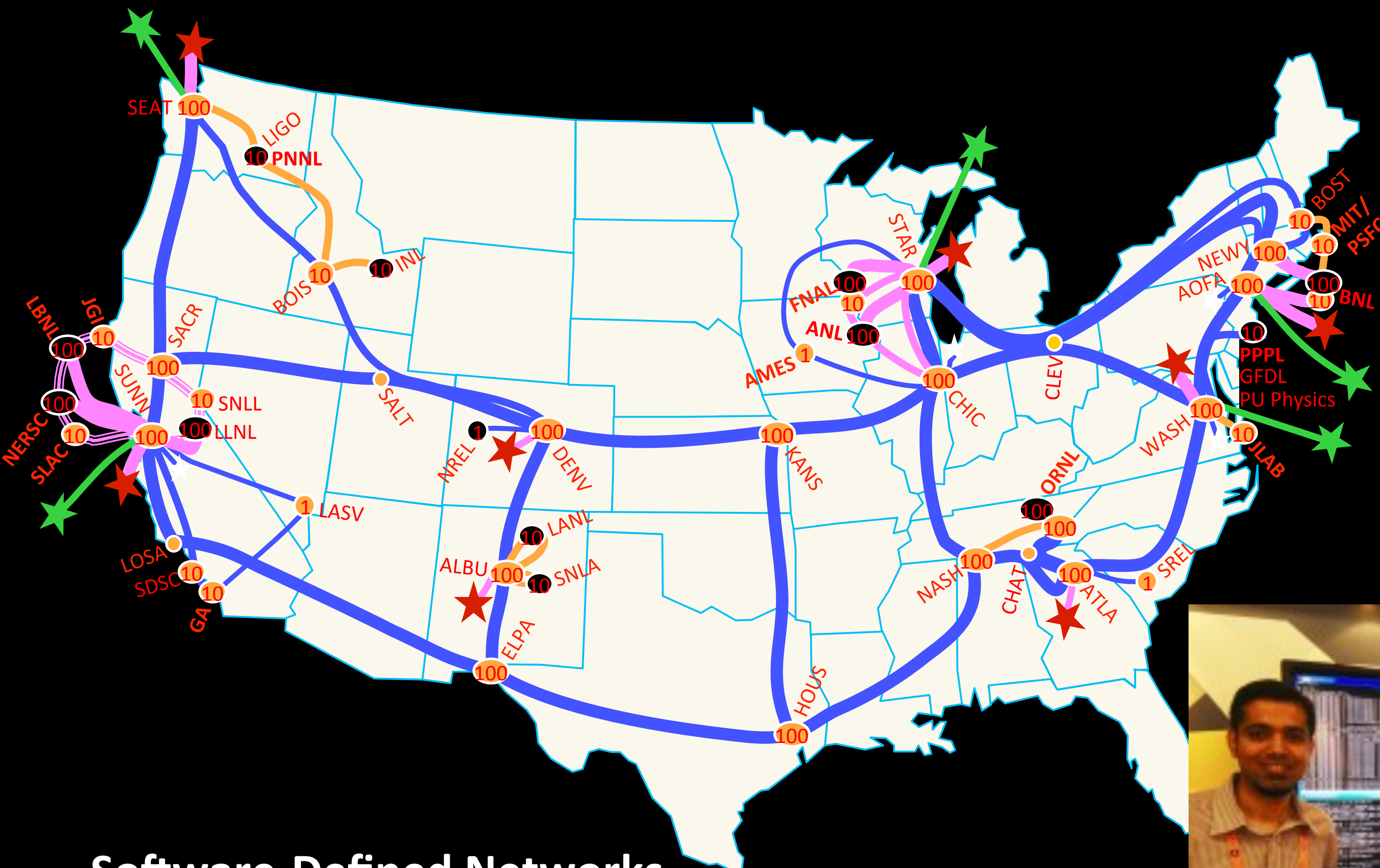
Analysis and modeling at NERSC, ORNL, or ANL

HipGISAXS simulation *HipRMC fitting*



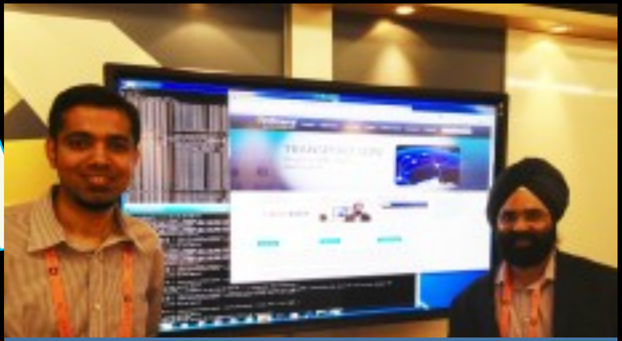


Network as ~~Infrastructure~~ *Instrument*



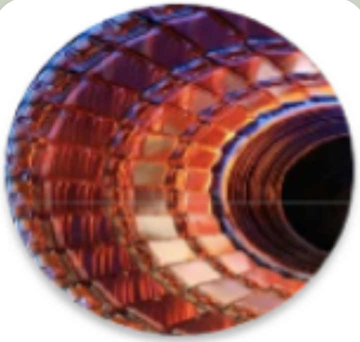
Software-Defined Networks

- Infrastructure: black box with complex internals
- Instrument: fast, adaptive, programmable

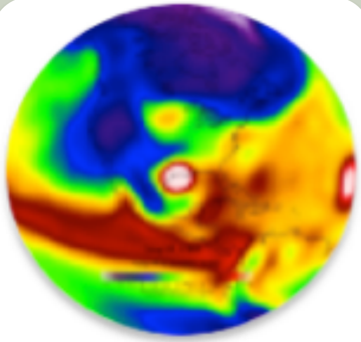


ESnet, Infinera, and Brocade demo transport SDN

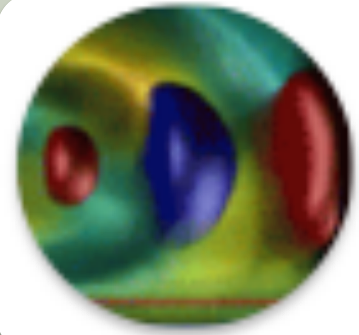
Programming Challenge? Science Problems Fit Across the "Irregularity" Spectrum



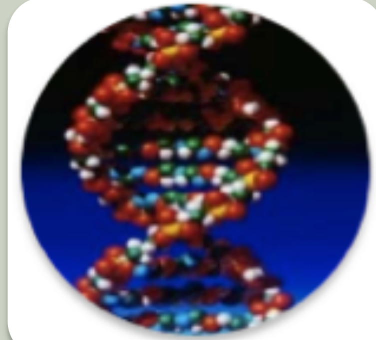
Massive Independent Jobs for Analysis and Simulations



Nearest Neighbor Simulations



All-to-All Simulations

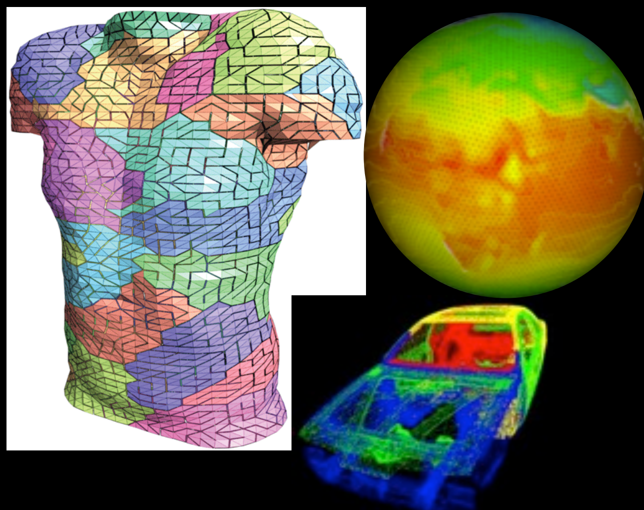


Random access, large data Analysis

… often they fit in multiple categories

How to Program (Irregular) Data Analics

More Regular



Message Passing Programming

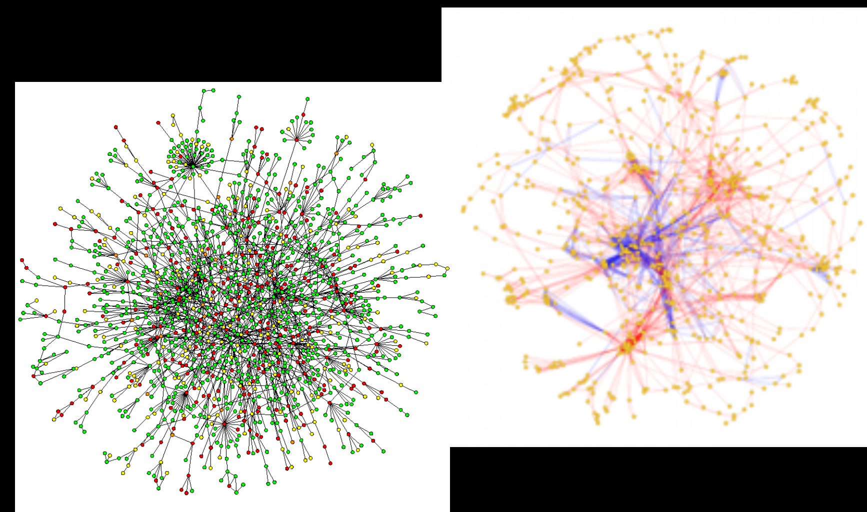
Divide up domain in pieces

Compute one piece

Send/Receive data from others

MPI, and many libraries

More Irregular



Global Address Space Programming

Each start computing

Grab whatever / whenever

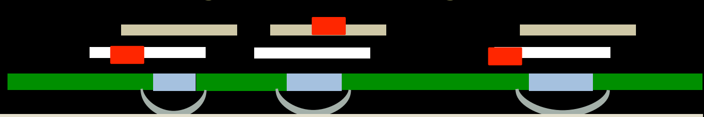
UPC, CAF, X10, Chapel, GlobalArrays

Languages for Random Access to Large Memory

Meraculous Assembly Pipeline

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

Scaffolds using Scalable Alignment



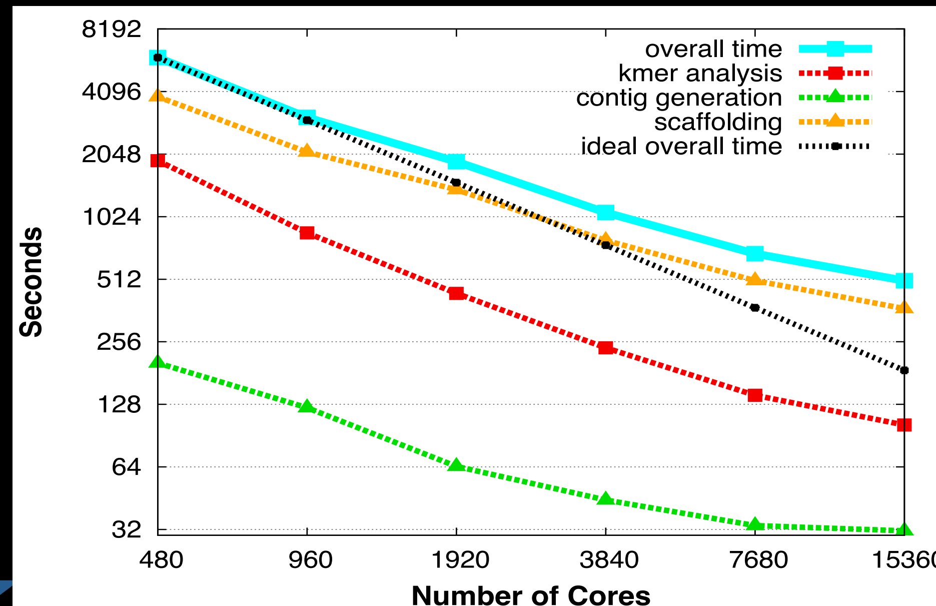
Combines with new algorithm to anchor 92% of wheat chromosome

Transforms process of discovery for de novo assembly

Perl to PGAS: Distributed Hash Tables

- Remote Atomics
- Dynamic Aggregation
- Software Caching (sometimes)
- Clever algorithms and data structures (bloom filters, locality-aware hashing)

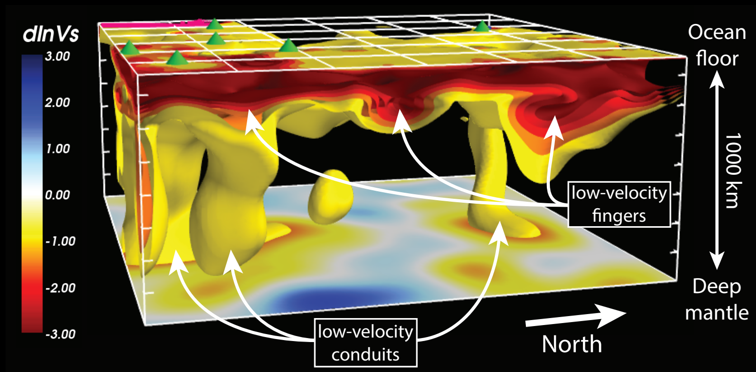
→ **Hash Table with "tunable" runtime**



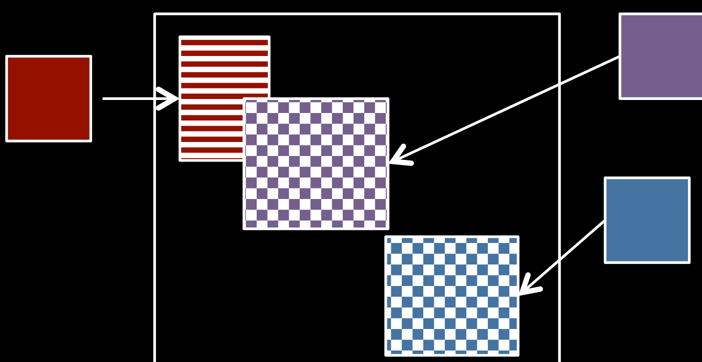
DEGAS

Evangelos Georganas, Aydin Buluc (MANTISSA), Lenny Oliker, Jarrod Chapman (JGI), Dan Rokhsar (JGI), Kathy Yelick

Languages for Irregular Access: Data Fusion in UPC++



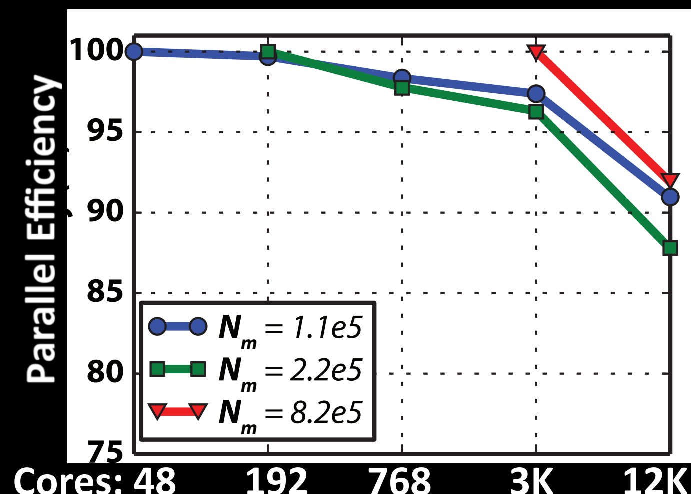
- Seismic modeling for energy applications "fuses" observational data into simulation
- With UPC++, can solve larger problems



Distributed Matrix Assembly

- Remote asyncs with user-controlled resource management
- Team idea to divide threads into injectors / updaters
- 6x faster than MPI 3.0 on 1K nodes

→ **Improving UPC++ team support**



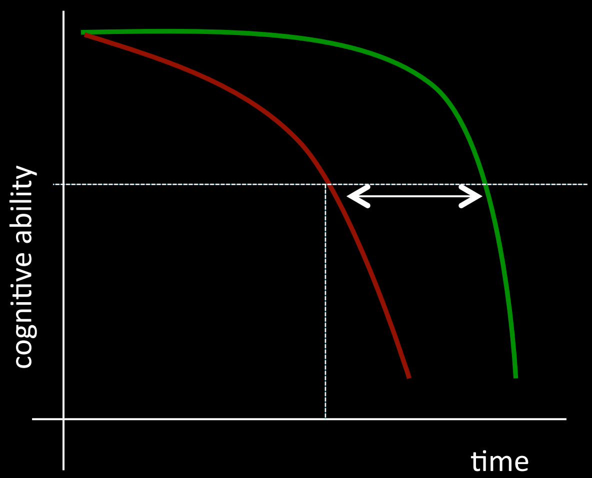
French and Romanowicz use code with UPC++ phase to compute *first ever* whole-mantle global tomographic model using numerical seismic wavefield computations (F & R, 2014, GJI, extending F et al., 2013, Science). See F et al, IPDPS 2015 for parallelization overview.

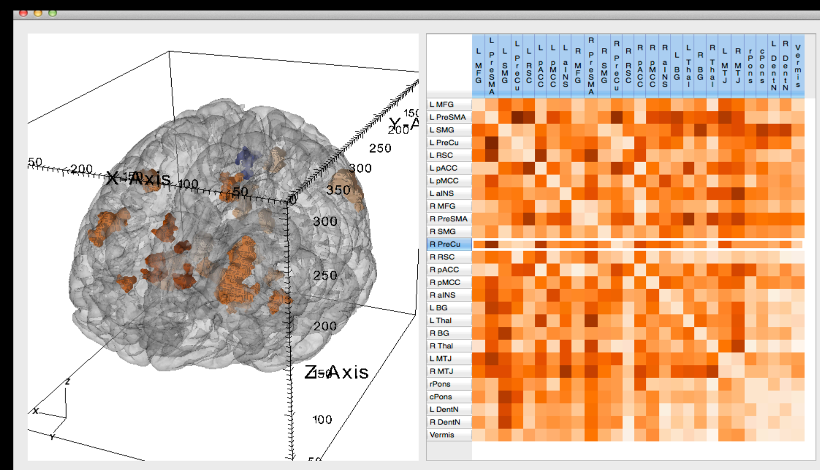
Analytics vs. Simulation Kernels:

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

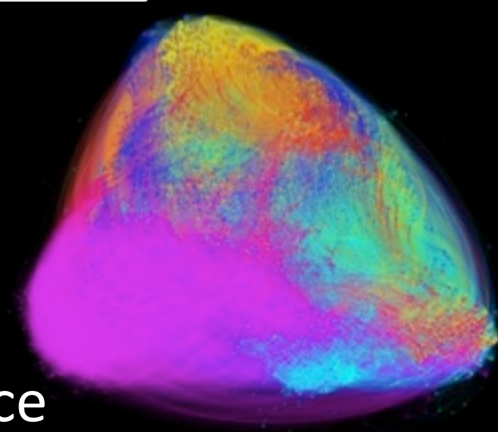
Brain Imaging: Multi-Modal Analysis and Data Fusion

Analyze brain using multiple modalities and scales





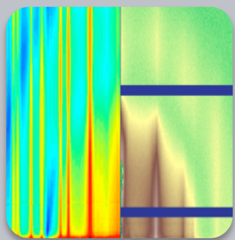
- **Detection of regions across community**
  - 100 individuals takes 18 days right now
- **Graph to classify disease**
  - Features: biomarkers, image modalities
  - Use hierarchy of regions from Pearson distance



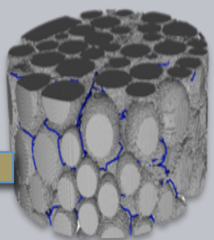
CAMERA leverages state-of-the-art mathematics to transform experimental data into understanding

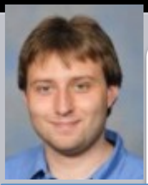
CENTER FOR APPLIED MATHEMATICS FOR ENERGY RESEARCH APPLICATIONS

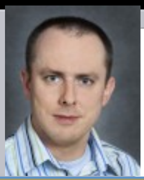

Sherry Li

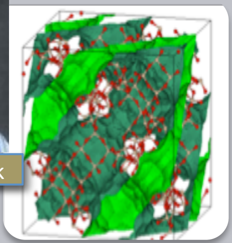



D. Ushizima




J. Donatelli


M. Haranczyk



X-ray scattering data analysis

HipGISAXS 400-1500x faster analysis for X-ray scattering data

Now: Nonlinear optimization, genetic algorithms, pattern recognition w/ noise

Micro-CT Sample Analysis

Quant-CT provides automated quantitative analysis

Now: 3D image segmentation, pattern recognition, classification algorithms, PDE- and graph-based analysis

X-ray Nano-Crystallographic Reconstruction

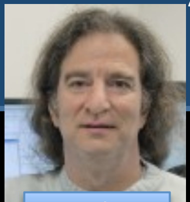
Indexing ambiguity resolved [PNAS13]

Now: Image orientation, find crystal shape/size; address orientation ambiguities; data variance reduction

Designing New Materials

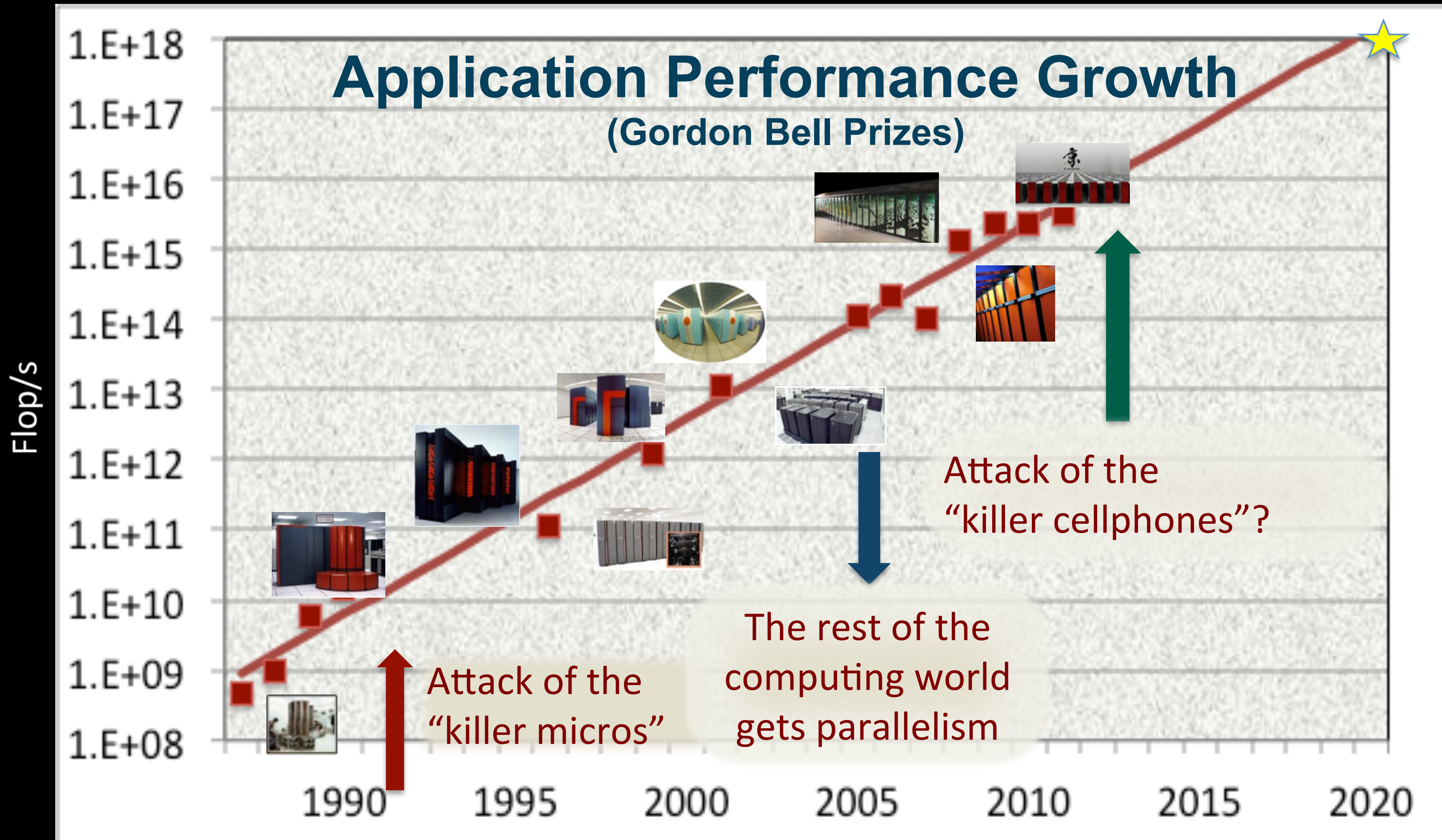
Designed record-breaking high-surface area materials

Now: 3D porous polymer model assembly; Zeo++ porosity characterization ; Optimal high-performing material designs


J. Sethian

Lead PI, DOE-funded

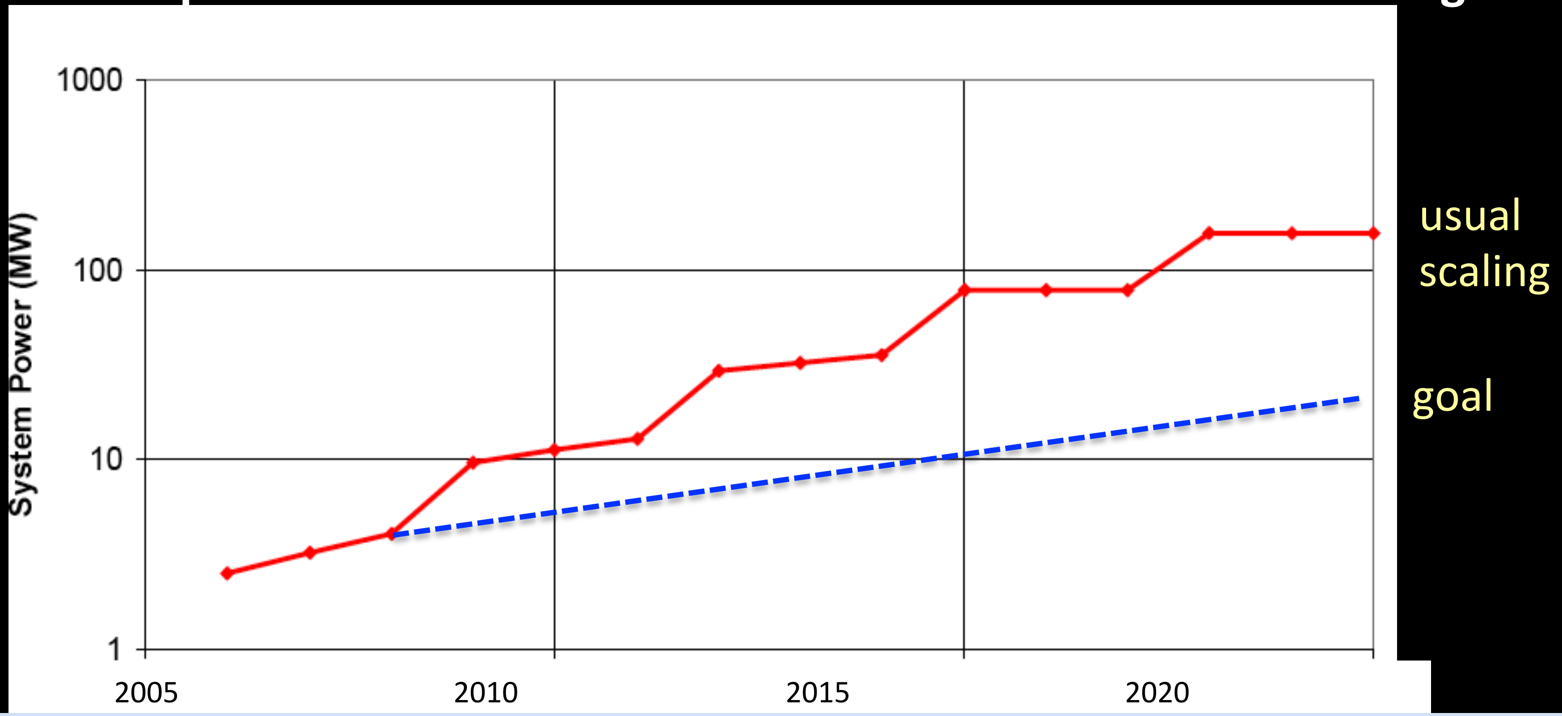
Architectures to Grow Computing Performance



Energy Efficiency is Key to Performance Growth

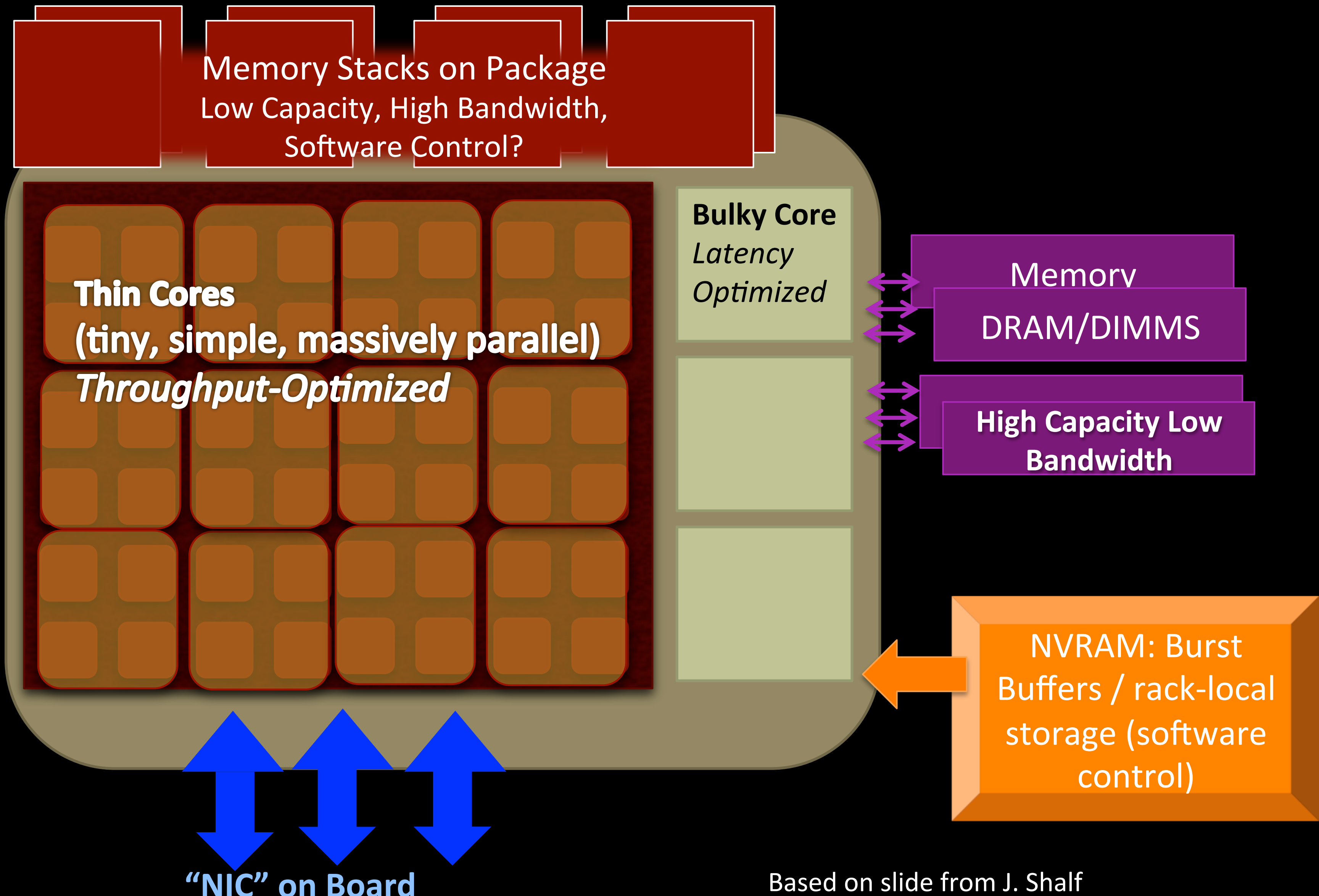
At $1M per MW, energy costs are substantial

- **1 petaflop in 2010 used 3 MW**
- **1 exaflop in 2018 would use 100+ MW with "Moore's Law" scaling**



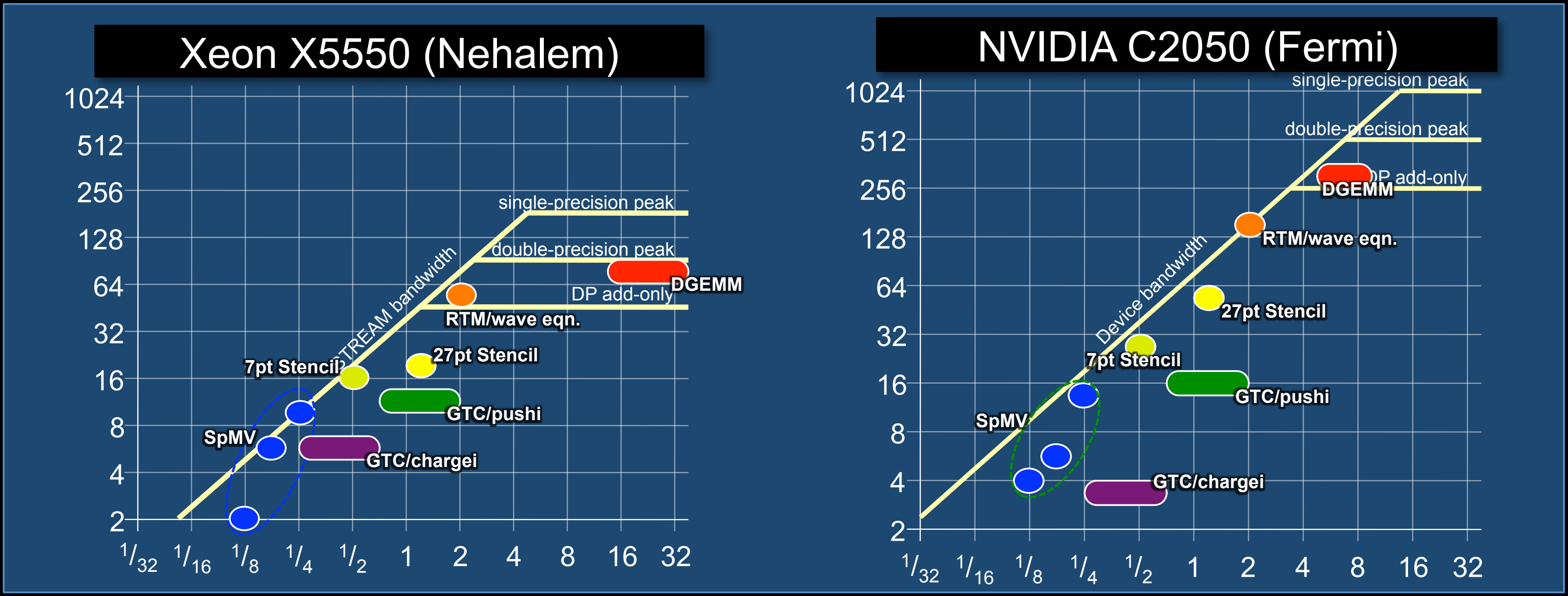
This problem doesn't change if we were to build 1000 1-Petaflop machines instead of 1 Exasflop machine. It affects every university department cluster and cloud data center.

Emerging Exascale Node Architecture



Based on slide from J. Shalf

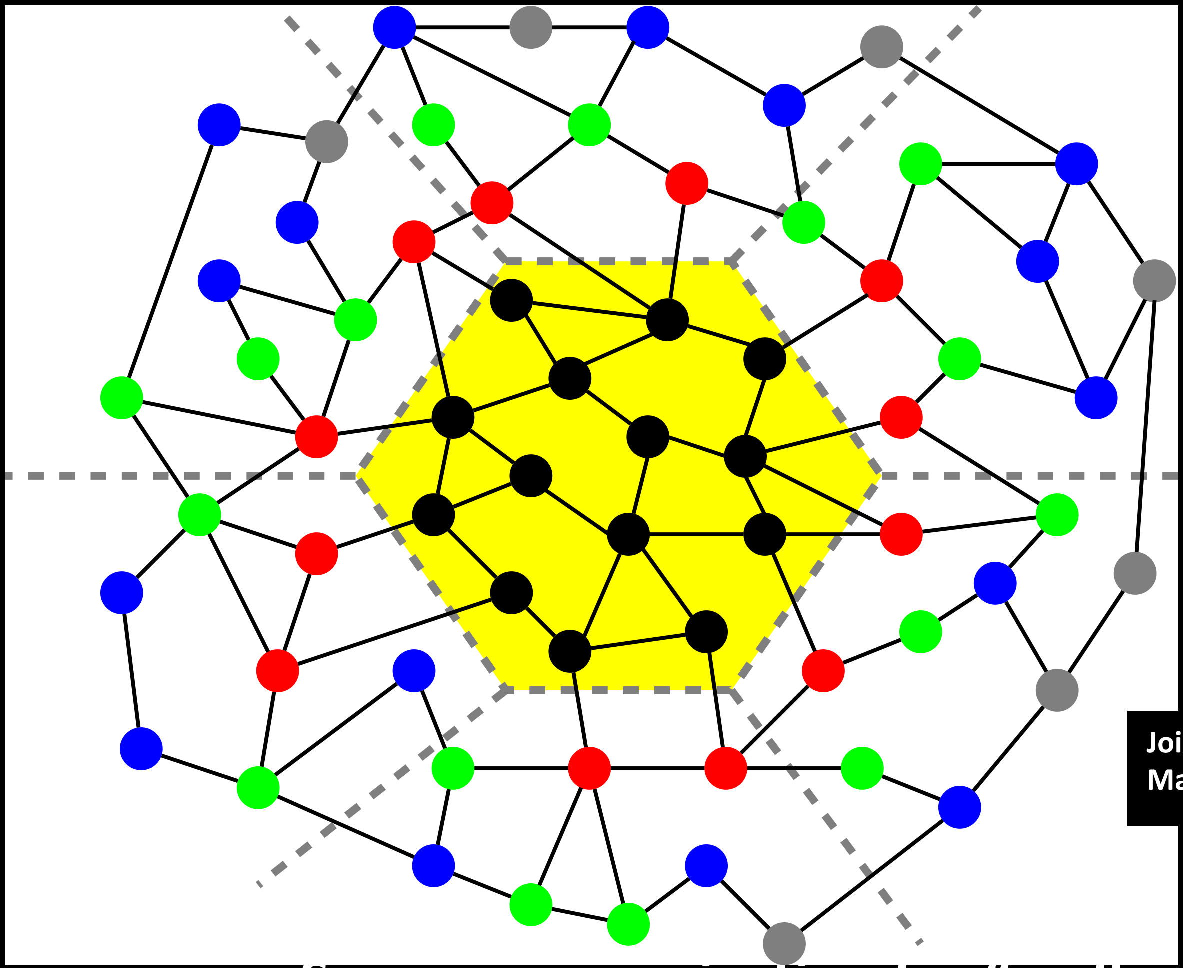
Exascale Programming: Memory System Structure



Communication wall will get worse (dominates energy and time)

- Optimizing for memory/network more important than ever
- Automatic data movement (caches, VM) can be wasteful
- Autotuning (search) helps reach bandwidth limits

Communication Avoiding Sparse Solvers

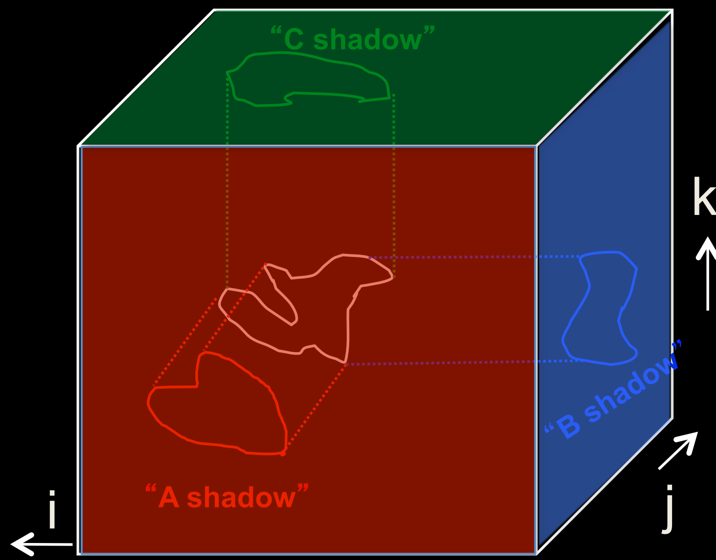


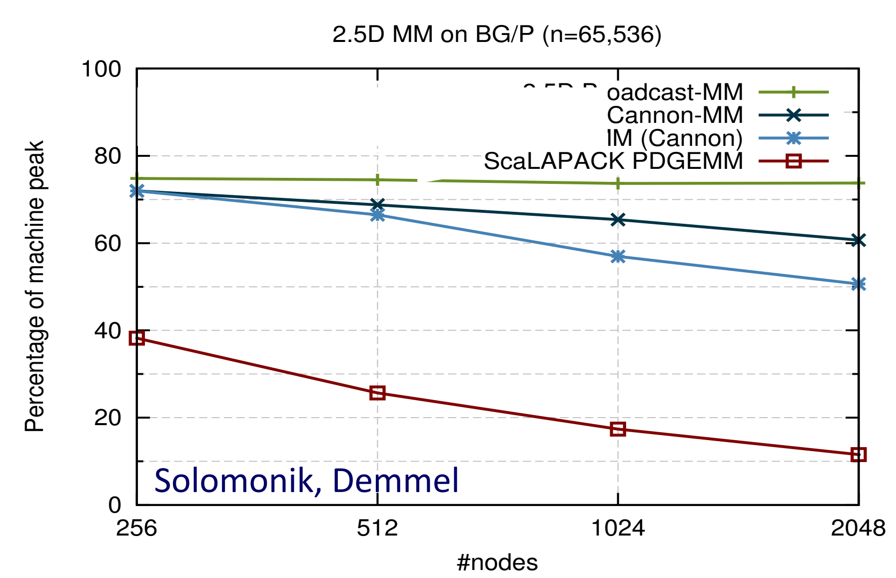
For implicit memory management (caches) uses a TSP algorithm for layout

Joint work with Jim Demmel, Mark Hoemman, Marghoob Mohiyuddin

- **Saves communication for “well partitioned” matrices**
  - Serial: O(1) moves of data moves vs. O(k)
  - Parallel: O(log p) messages vs. O(k log p)

Algorithms and Software: Avoid Communication and Synchronization





"Rethink" algorithms to optimize for data movement

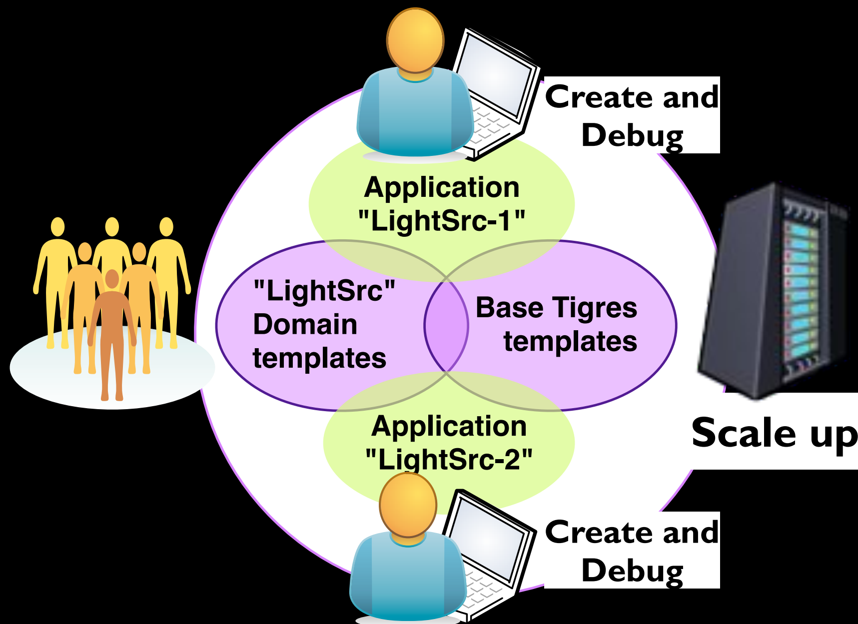
- New class of communication-optimal algorithms
- Most codes are not bandwidth limited, but many should be

Can they be automated and for what types of loops?

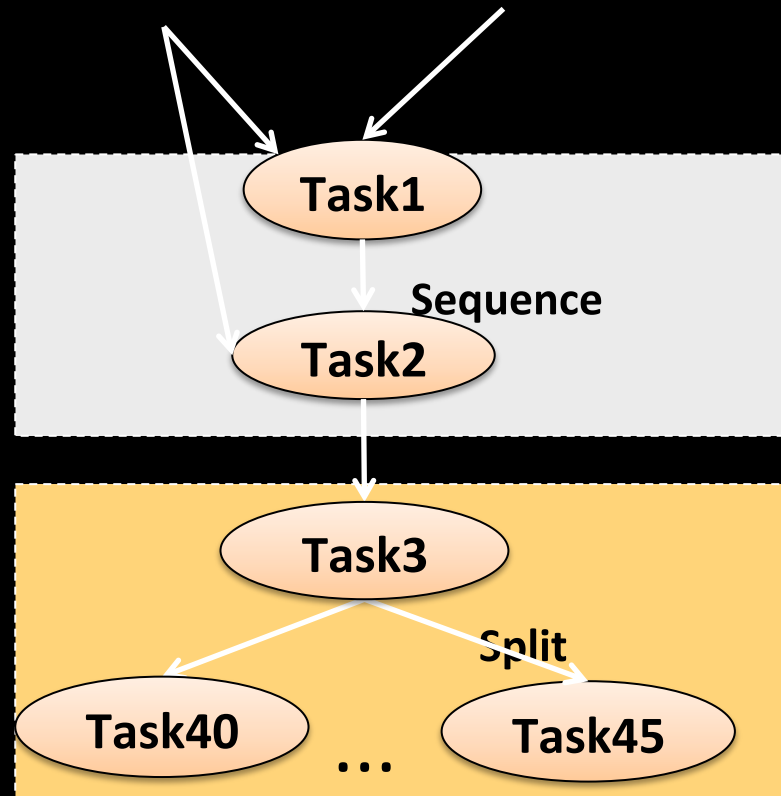
Technology for Scientific Data

	Compute Intensive	***Data Intensive***
Compute		
On-Package DRAM		
Capacity Memory		
On-node-Storage	**Goal:** *Maximize computational capacity and local bandwidth*	**Goal:** *Maximize data capacity and global bandwidth*
In-Rack Storage		
Interconnect		
Global Shared Disk		
Off-System Network		

Scientific Workflows

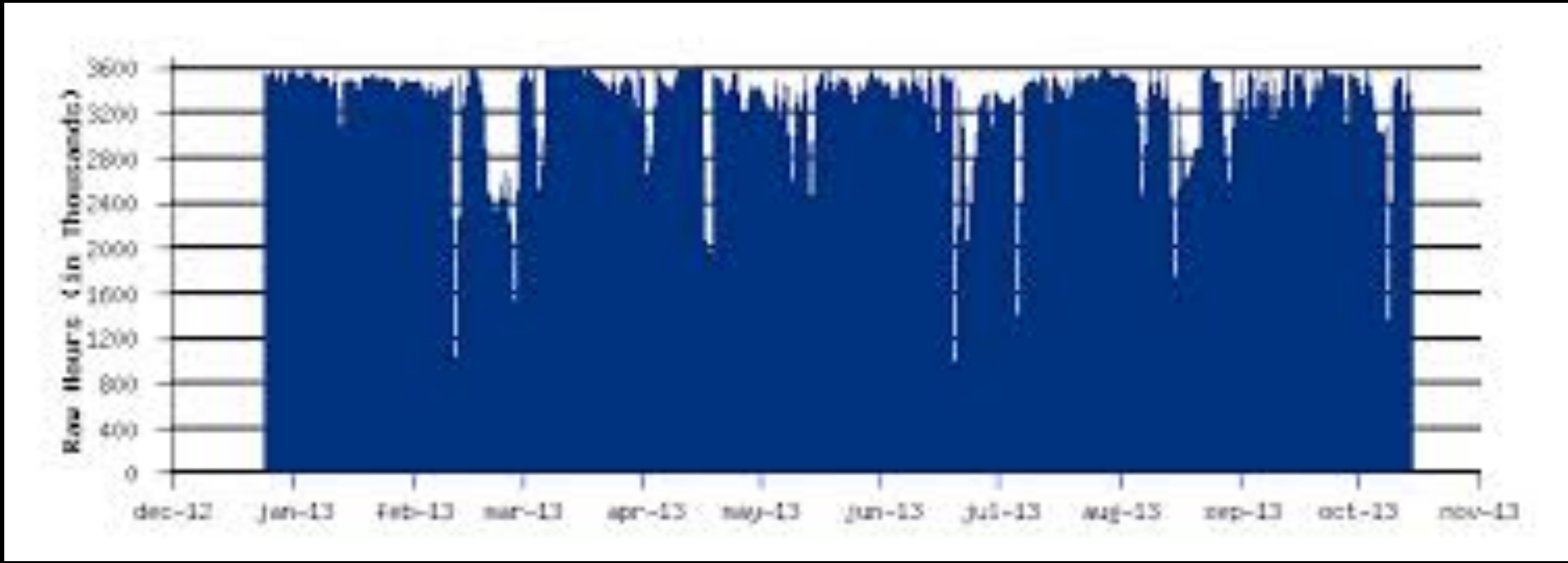


- **Tigres: Design *templates* for scientific workflows**
  - Explicitly support Sequence, Parallel, Split, Merge
- **Fireworks: High Throughput job scheduler**
  - Runs on HPC systems



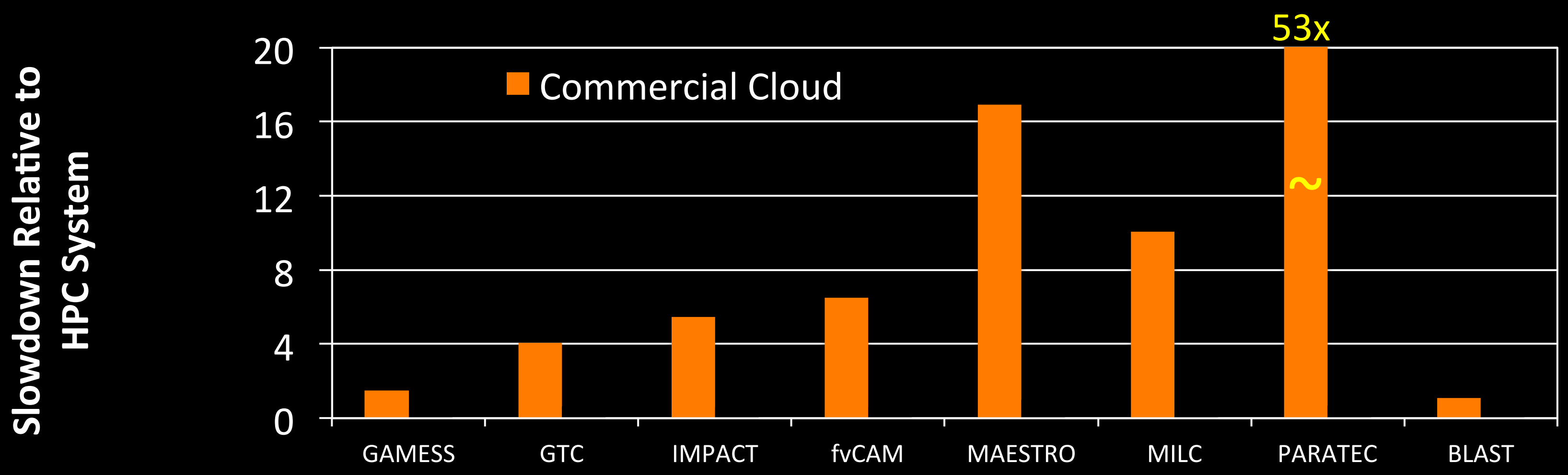
L. Ramakrishnan, V. Hendrix, D. Gunter, G.Pastorello, R. Rodriguez, A. Essari , D. Agarwal

Trade-offs in Utilization vs Response Time



- 95% utilization, but the users wait
- Real-time analysis on streams
- Interactive access to data

Cloud Computing: A Business Model, Not a Technology



NERSC features:

- NERSC runs at higher utilization (> 90%) and no profit.
- NERSC cost/core hours dropped 10x (1000%) from 2007 to 2011, while Amazon pricing dropped 15% in the same period
- No on-demand offered

Biggest differences: Resilience model and job scheduling

Take Home Messages

- **Science is increasingly computational**
- **Computational Science = Data Analysis + Simulation**
  - In many disciplines these will be tightly integrated
- **Needs computing capability**
  - Limited by energy; scale is one issue, but not the only one
- **Centralized computing is efficient**
  - The business model is a separate issue
- **Computer science and math research challenges**
  - In data analytics, computing performance, simulation…
- **Future: more computing → transform science**

Black Swans of Computing

Google

2012 Computing with 1992 Technology

Google™