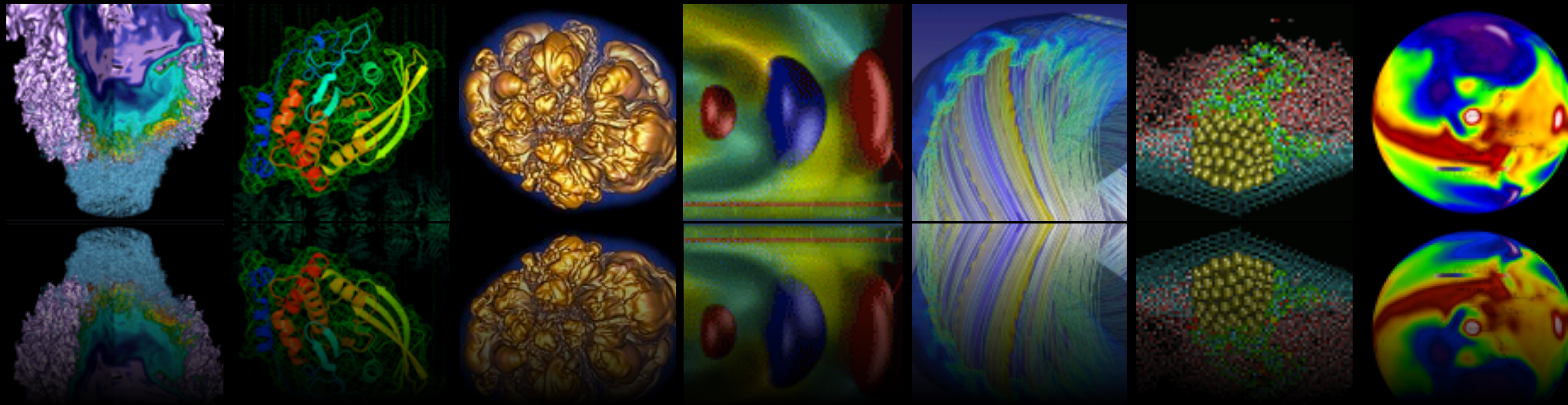


More Data, More Science and… Moore's Law

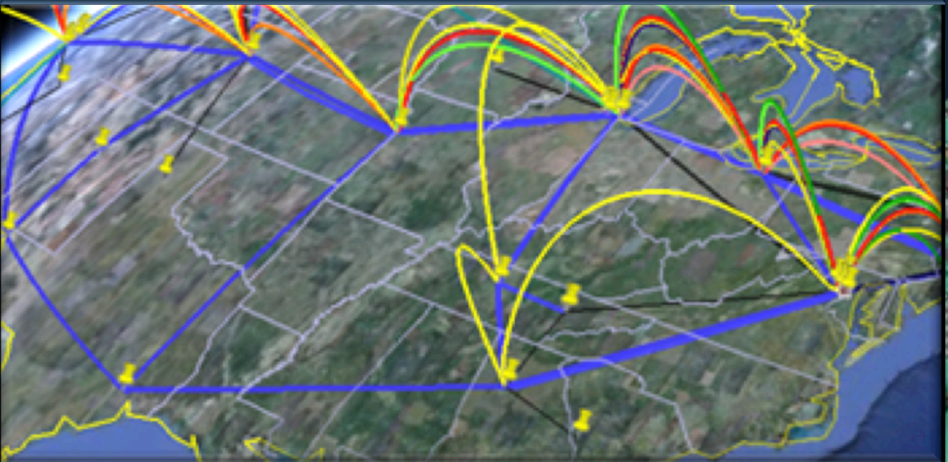
Kathy Yelick

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

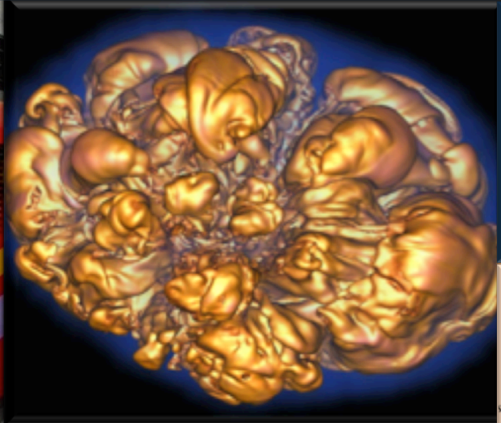
Computing Sciences at Berkeley Lab

NERSC Facility

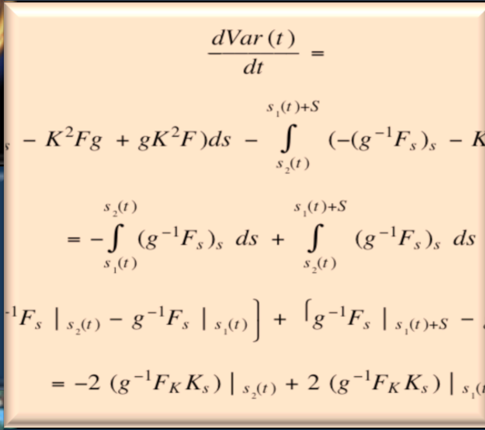
ESnet Facility



Computational Research



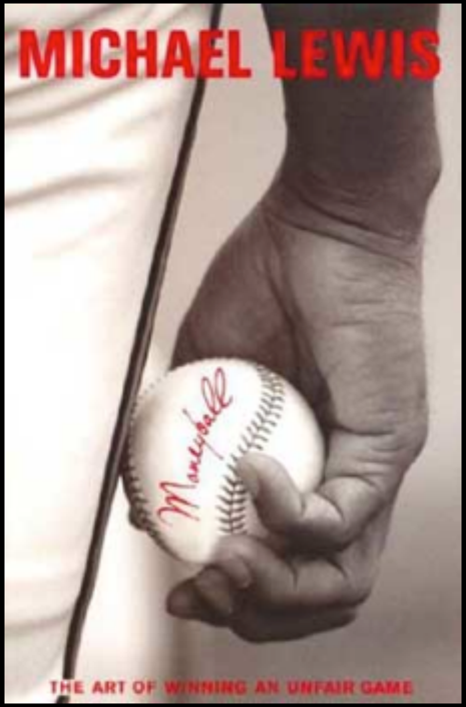
Computational Science



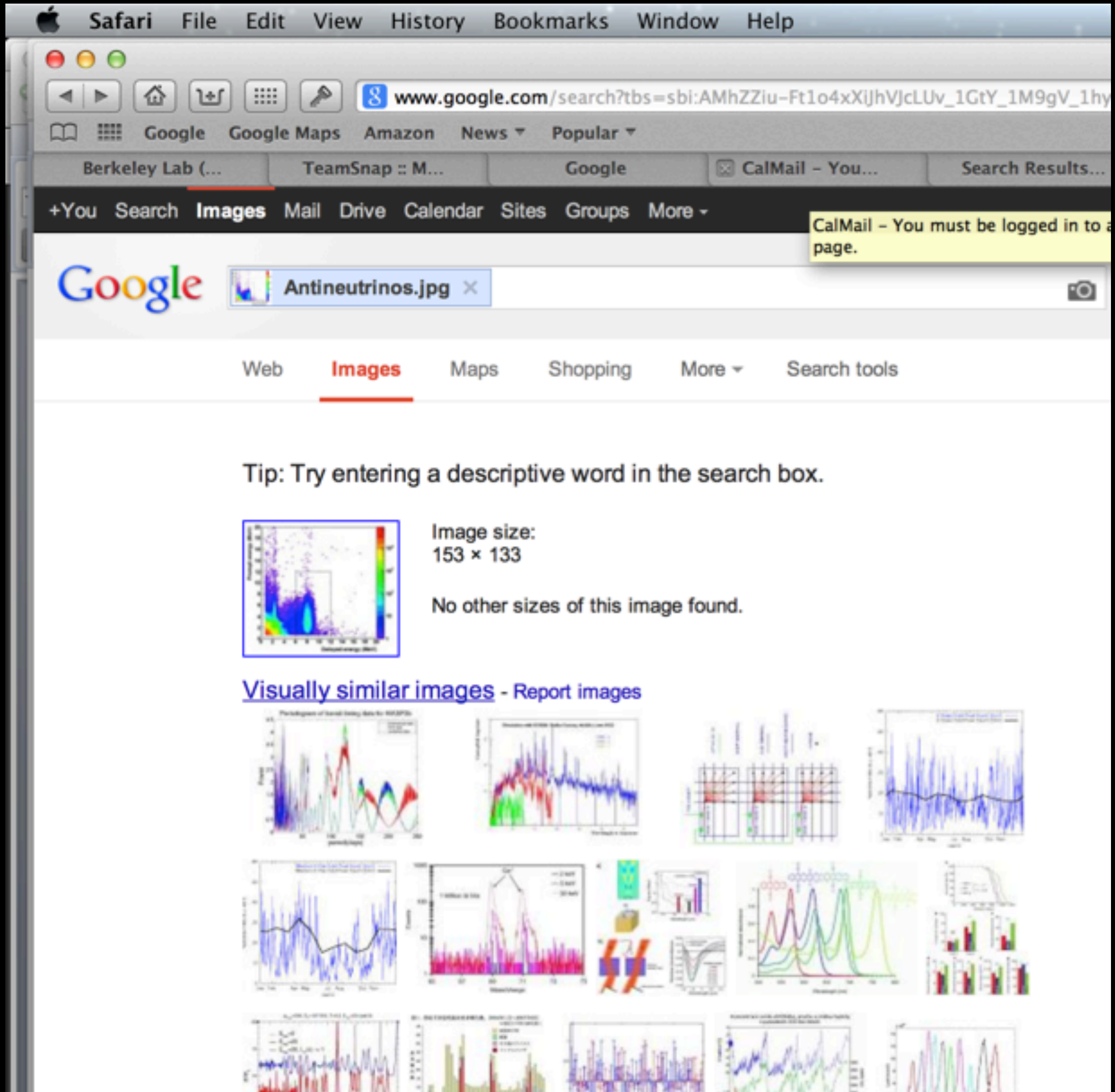
Applied Mathematics

Computer Science

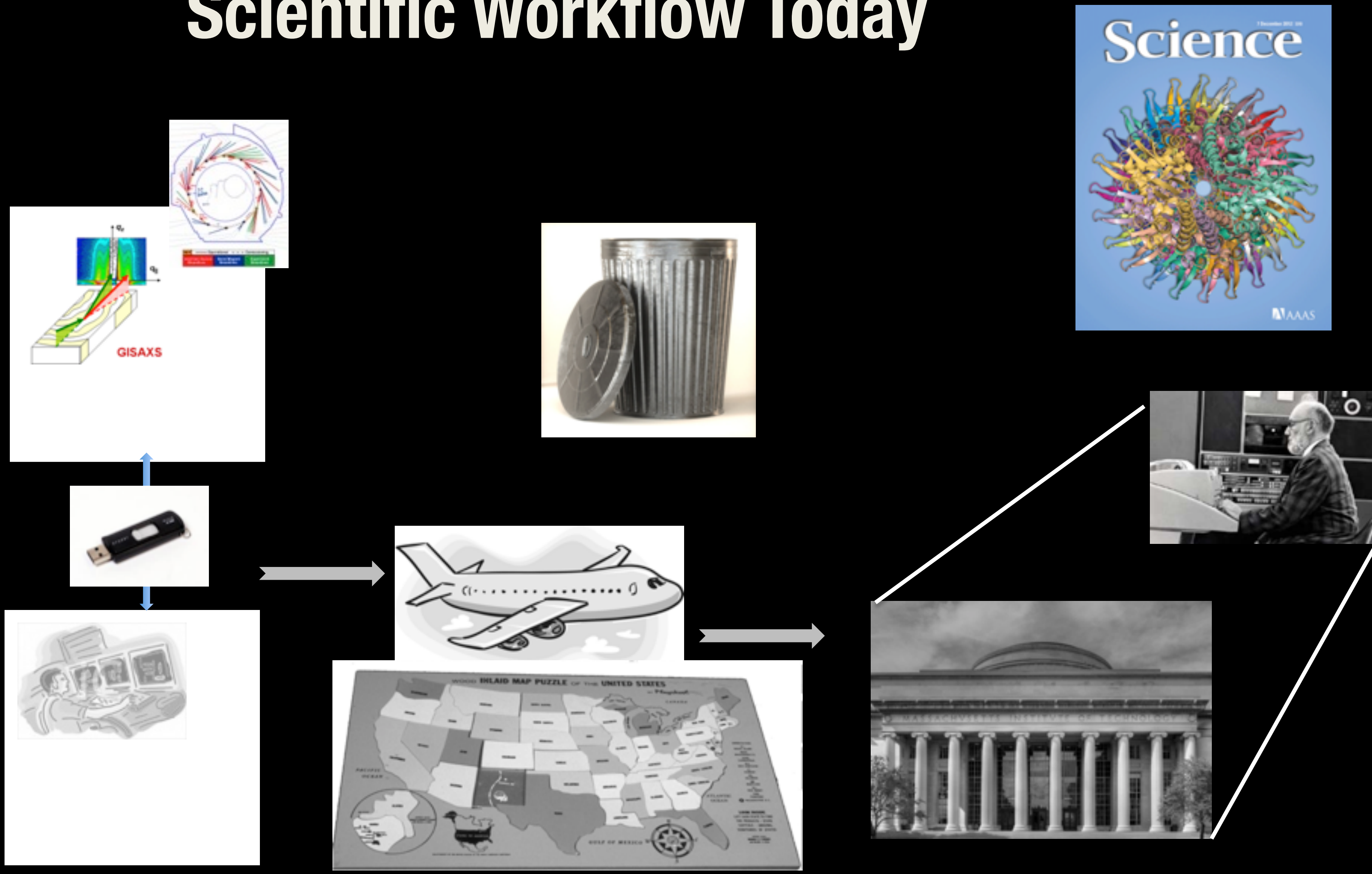
"Big Data" Changes Everything…What about Science?



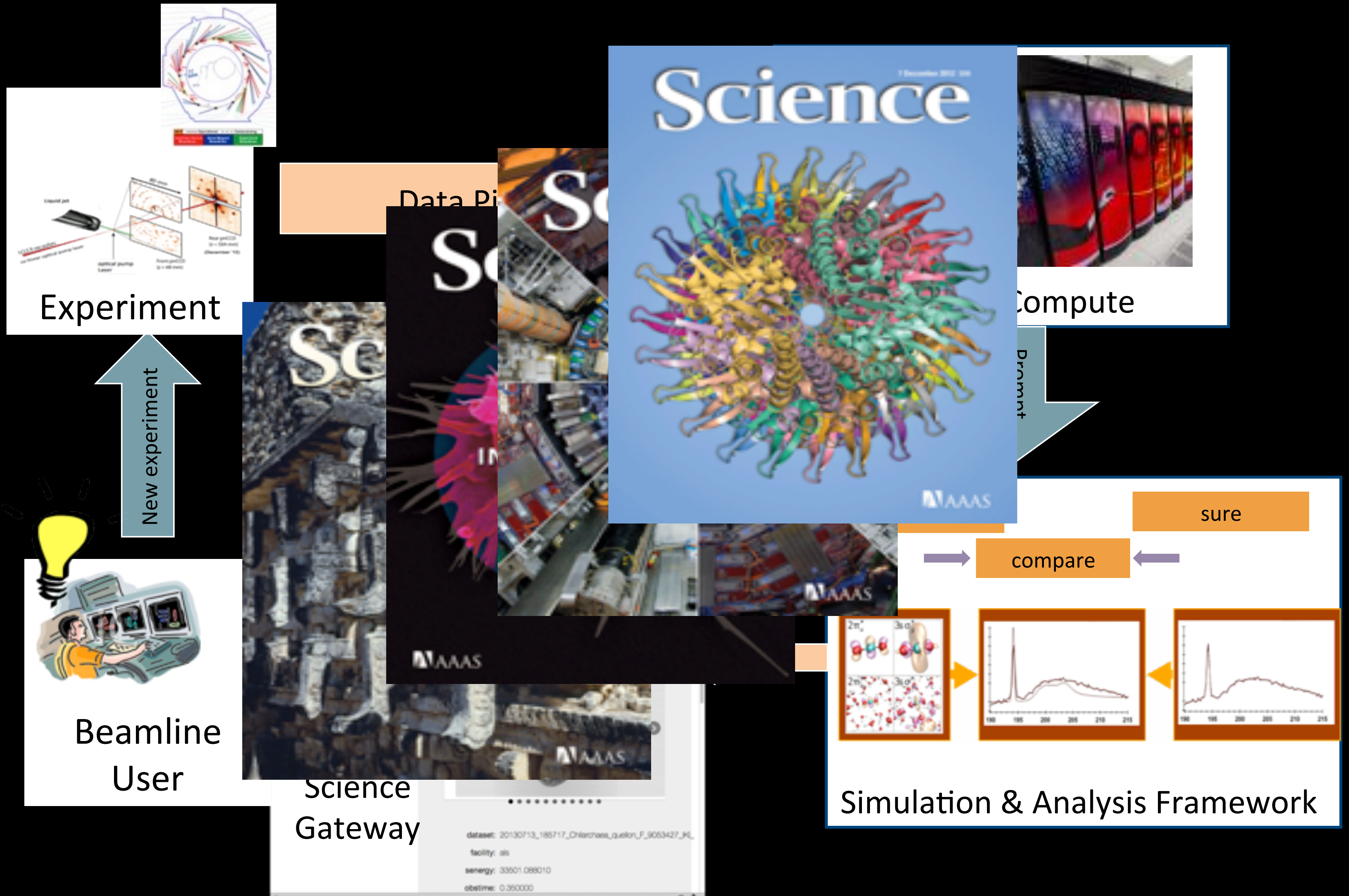
Transforming Science: Finding Data



Scientific Workflow Today



Scientific Workflow envisioned



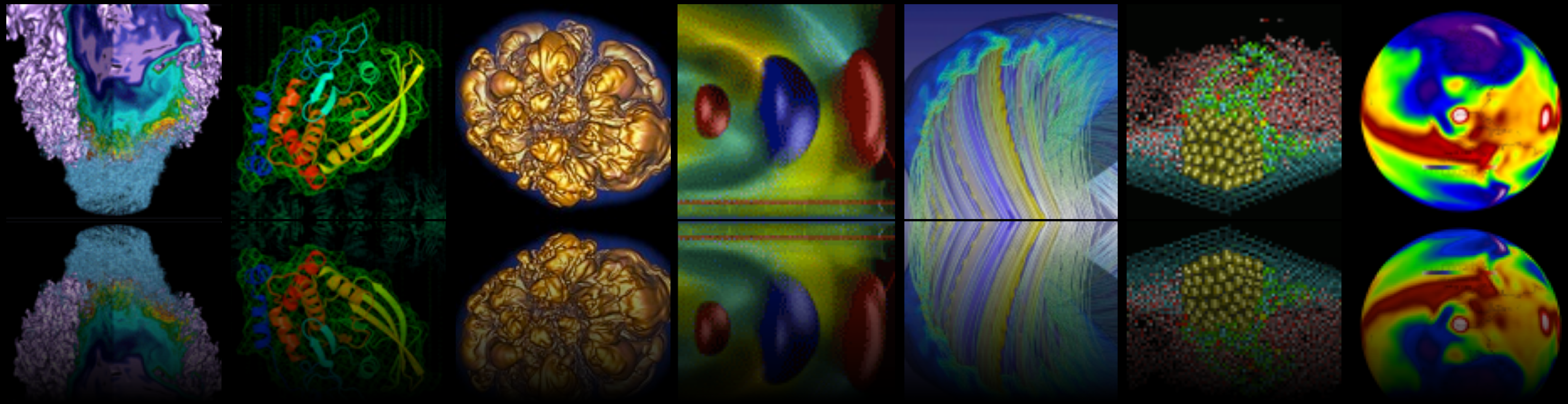
KATHY YELICK'S

2031:

a science odyssey

Life of a Scientist in 2031

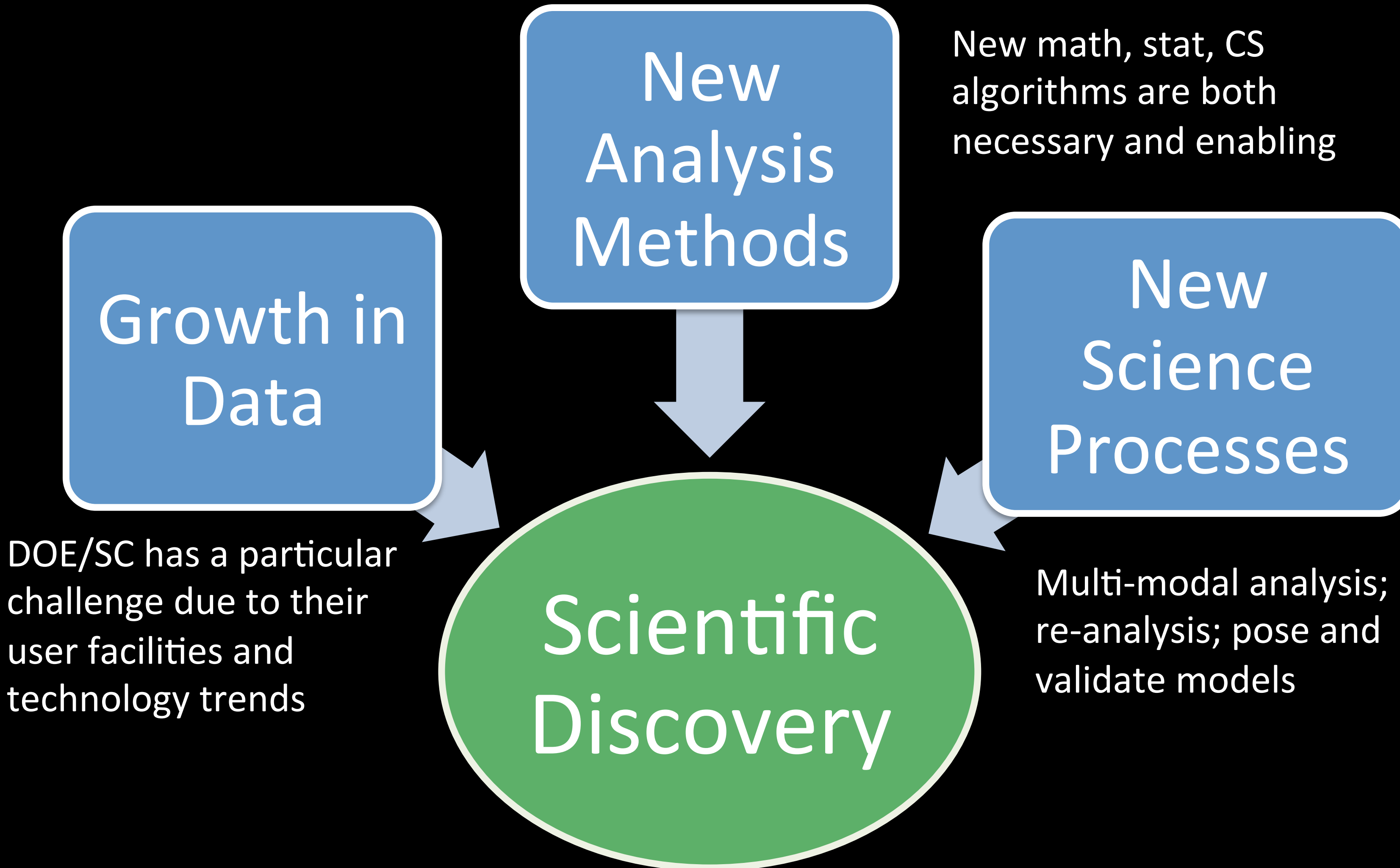
- No personal/departmental computers
- Users don't login to HPC Facilities
- Travel replaced by telepresence
- Lecturers teach millions of students
- Theorems proven by online communities
- Laboratory work is outsourced
- Experimental facilities are used remotely
- All scientific data is open
- Big science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



Extreme Data Science

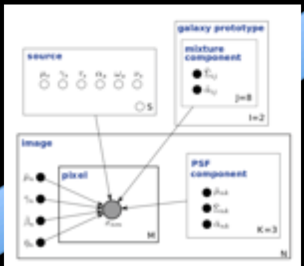
The scientific process is poised to undergo a radical transformation based on the ability to access, analyze, simulate and combine large and complex data sets.

Goal: To enable new modes of scientific discovery



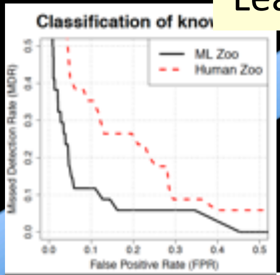
Data in Astrophysics: The Challenge is Systematics

Graphical models



New simulation models and AMR code (Nyx)

Machine Learning

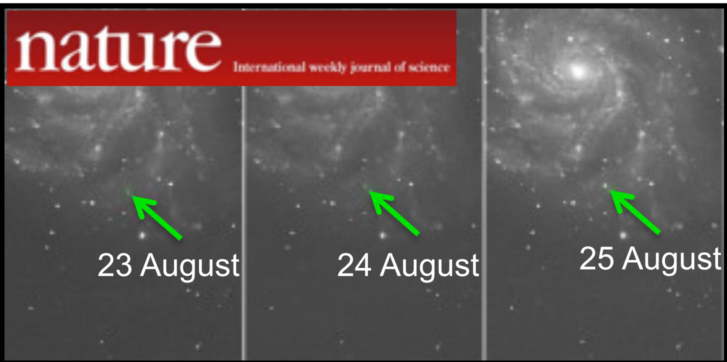


Crowd sourced

Filtered

GB per night
Manually analyzed

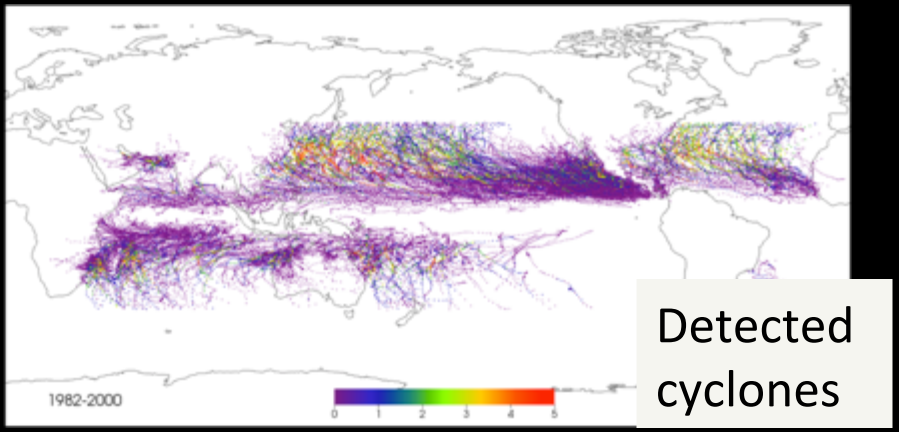
Example: Astrophysicists discover early nearby supernova

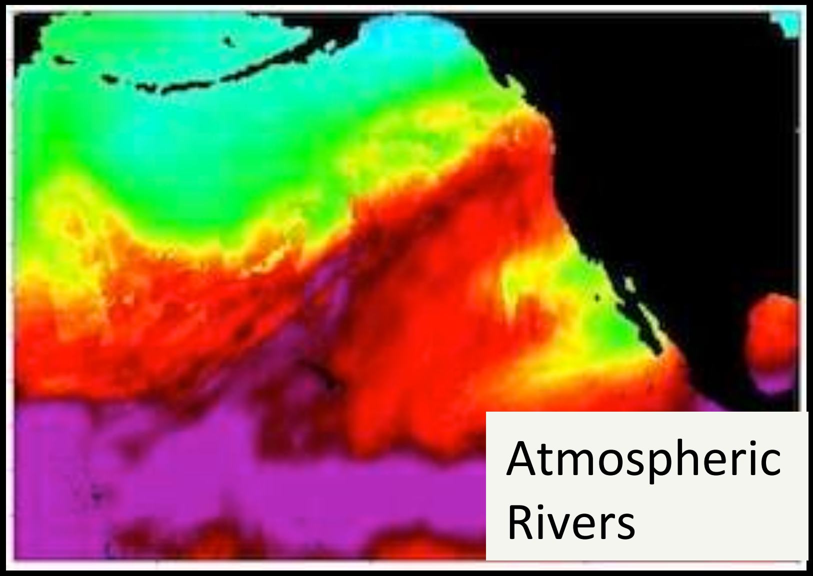


Identify Phenomenon using Machine Learning

- **TECA Toolkit today**
  - Automatic detection of cyclones, atmospheric rivers, and more
  - Analysis time years to minutes

- **Climate Analysis in 2031**
  - Machine learning for all events
  - Automatic metadata generation
  - Fusion of simulations, sensors, etc.
  - Real-time analysis and response





Filtering, De-Noise 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

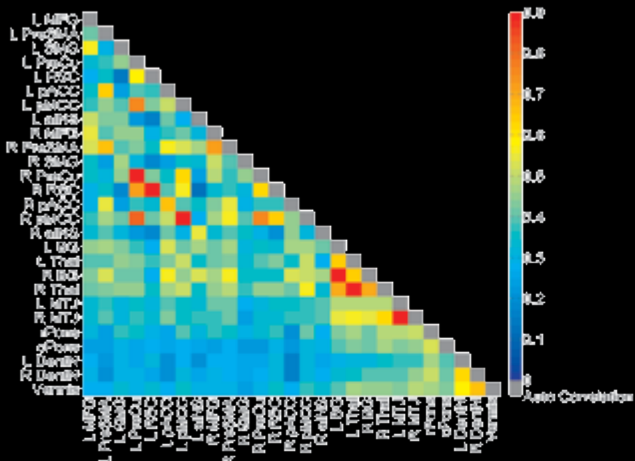
Re-Use and Re-Analyze Previously Collected Data

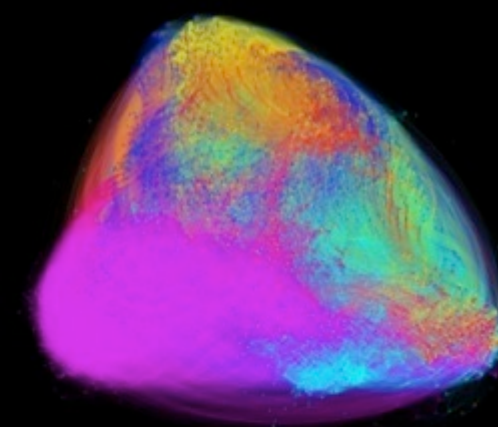
- **Materials Genome Initative**
  - Materials Project: 4500 users 18 months!
  - Scientific American “World Changing Idea” of 2013 – what about 2031?

Unbounded computing requirements for simulation and analysis

Multi-modal analysis of Brain Connectivity

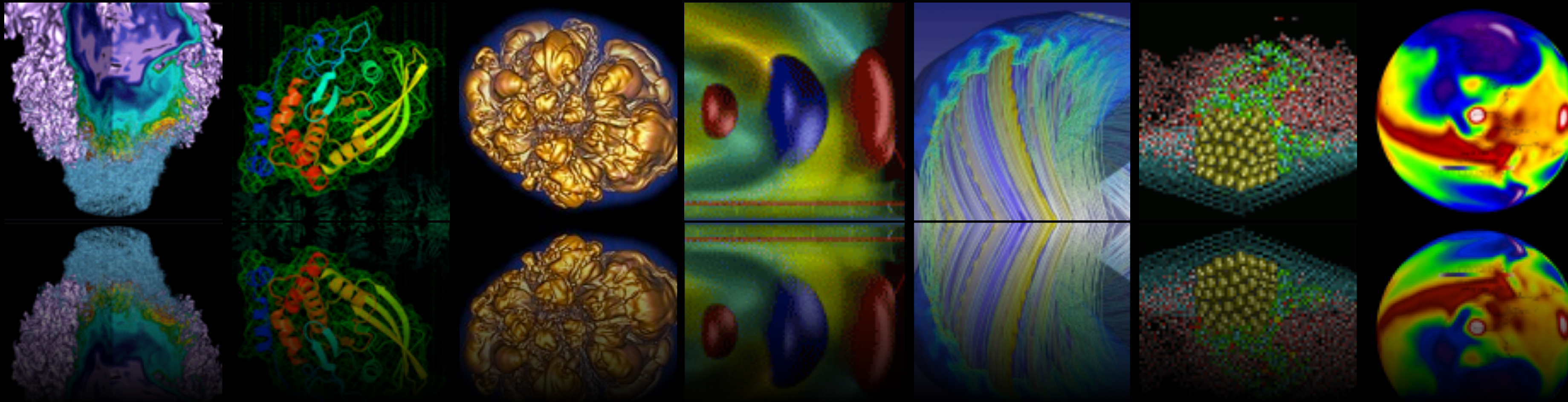
Analyze brain connectivity at multiple scales:
From cells and regions to complex neural circuits.





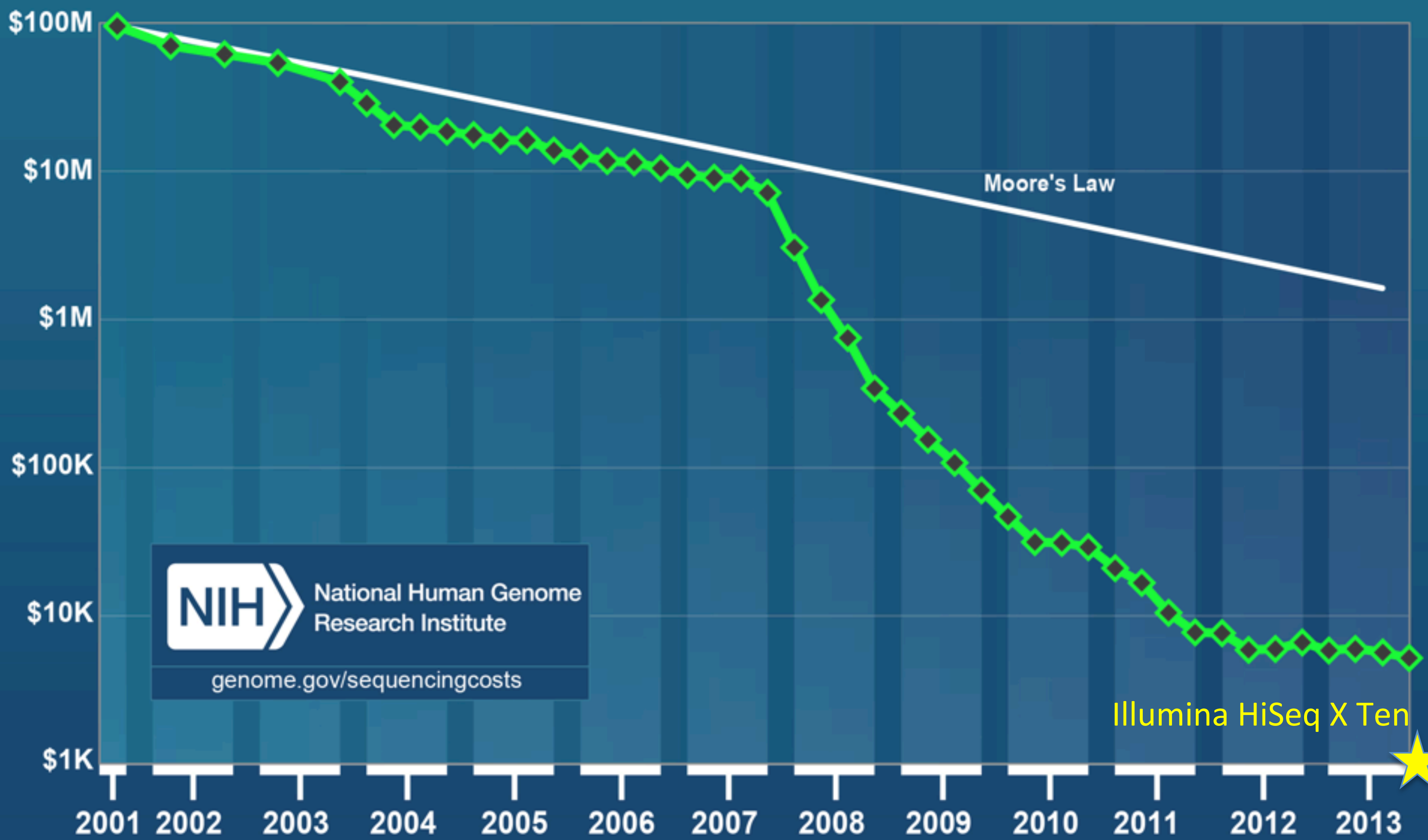
Brain Connectivity Graphs: Jesse Brown, Bill Seely (UCSF)

- **Improve understanding of brain pathology.**
- **Enable personalized treatment options.**

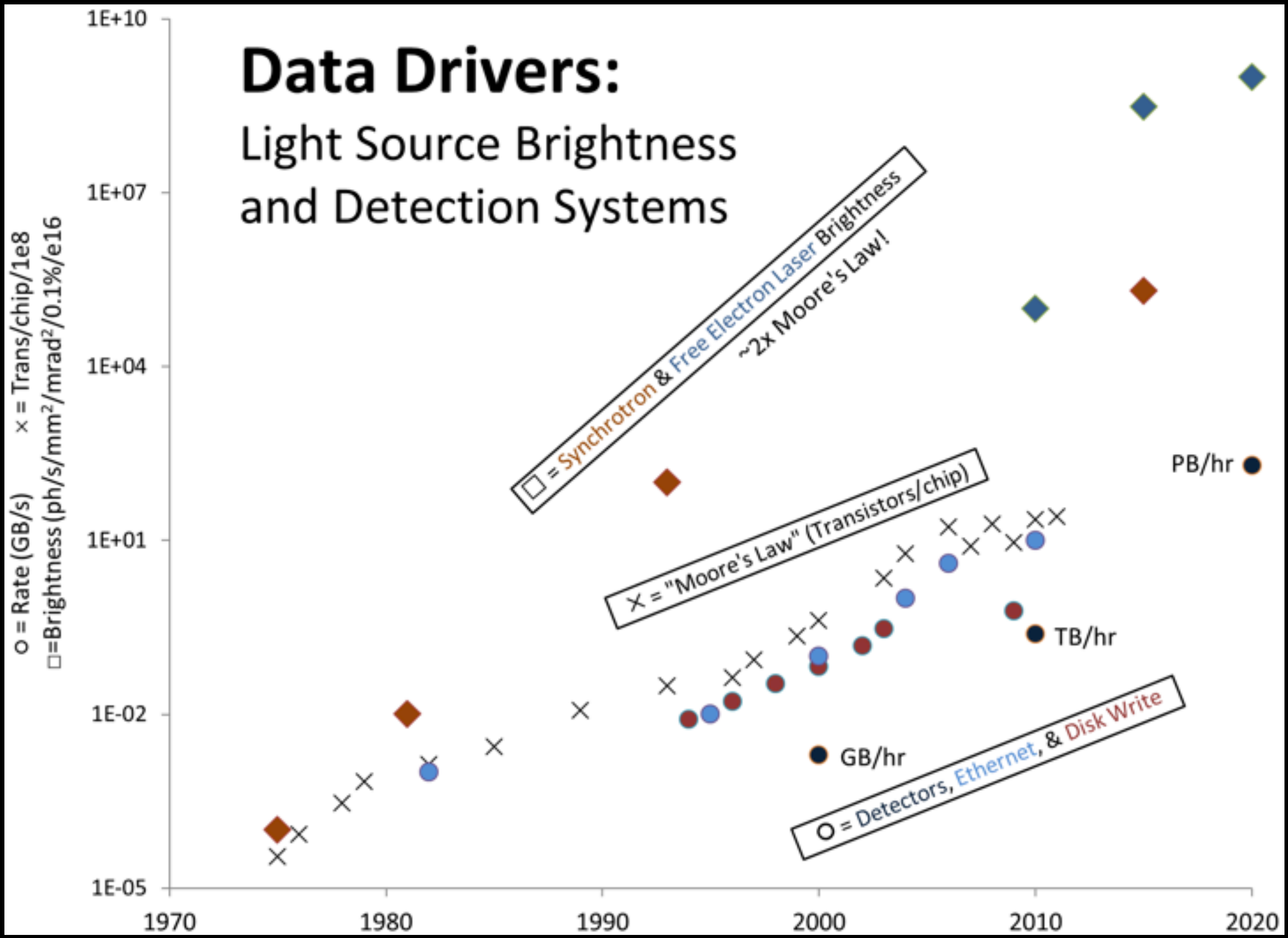


Science Data is Big (and Growing)

Cost per Genome

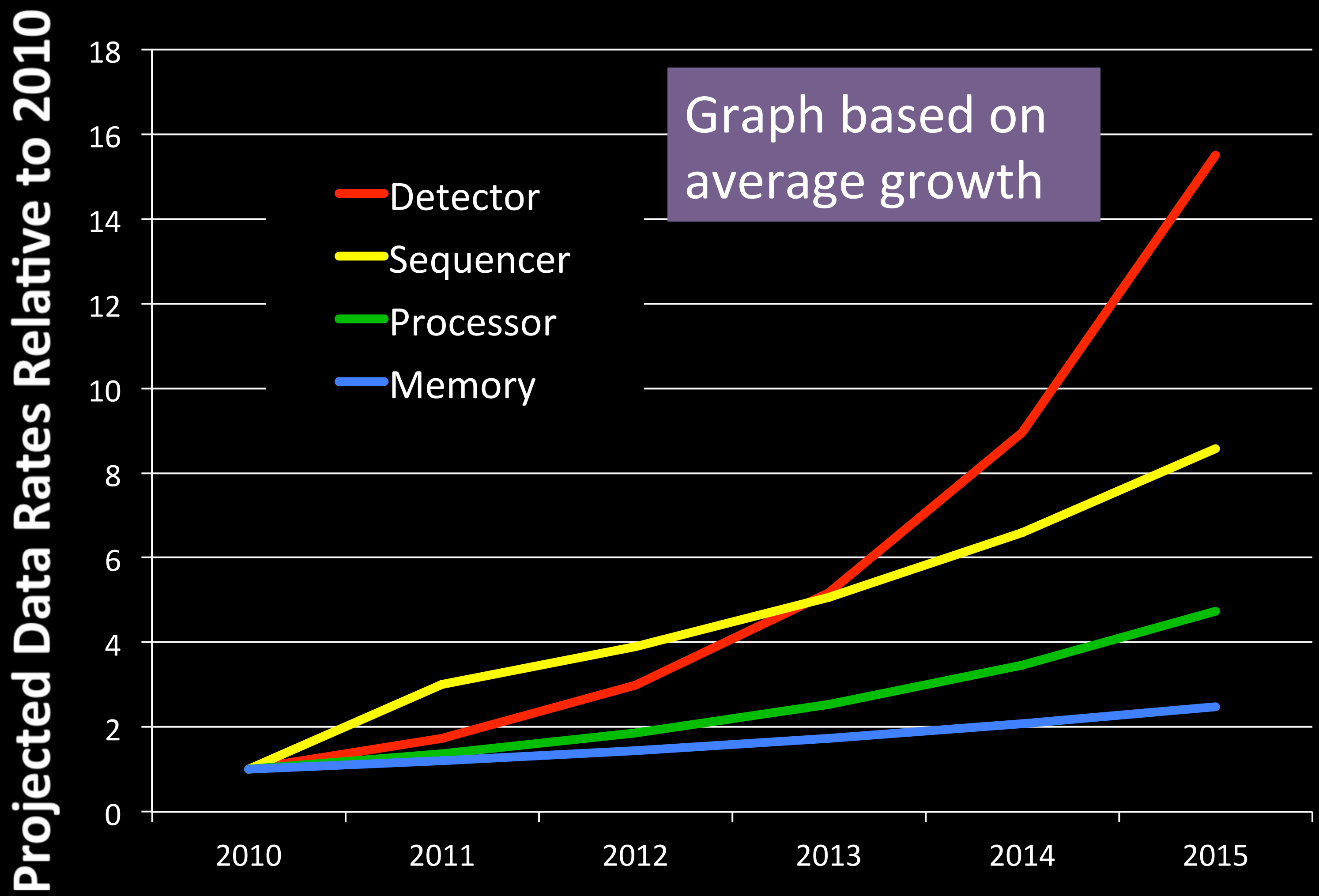


Light Sources – It's Not Just Genomics

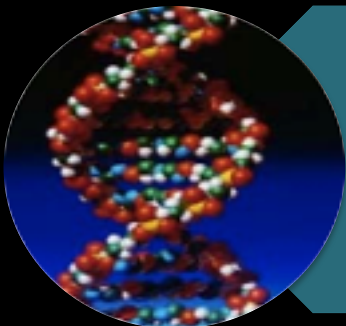


source: D. Parkinson, LBNL

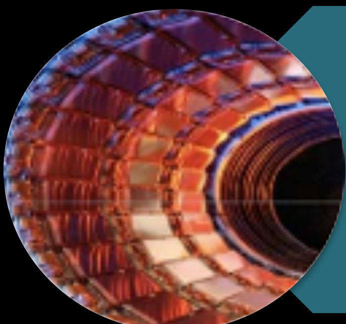
Data Growth is Outpacing Computing Growth



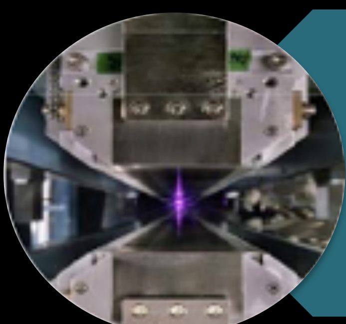
"Big Data" Challenges in Science
Volume, velocity, variety, and veracity



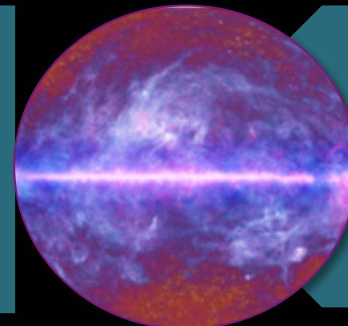
Biology
- *Volume:* Petabytes now; computation-limited
- *Variety*: multi-modal analysis on bioimages



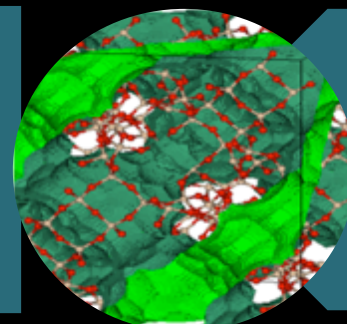
High Energy Physics
- *Volume*: 3-5x in 5 years
- *Velocity*: real-time filtering adapts to intended observation



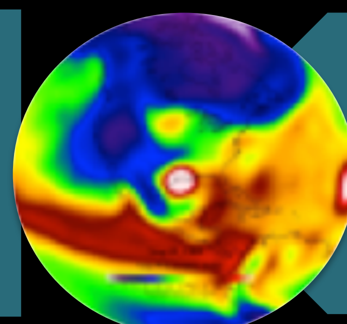
Light Sources
- *Velocity:* CCDs outpacing Moore's Law
- *Veracity:* noisy data for 3D reconstruction



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

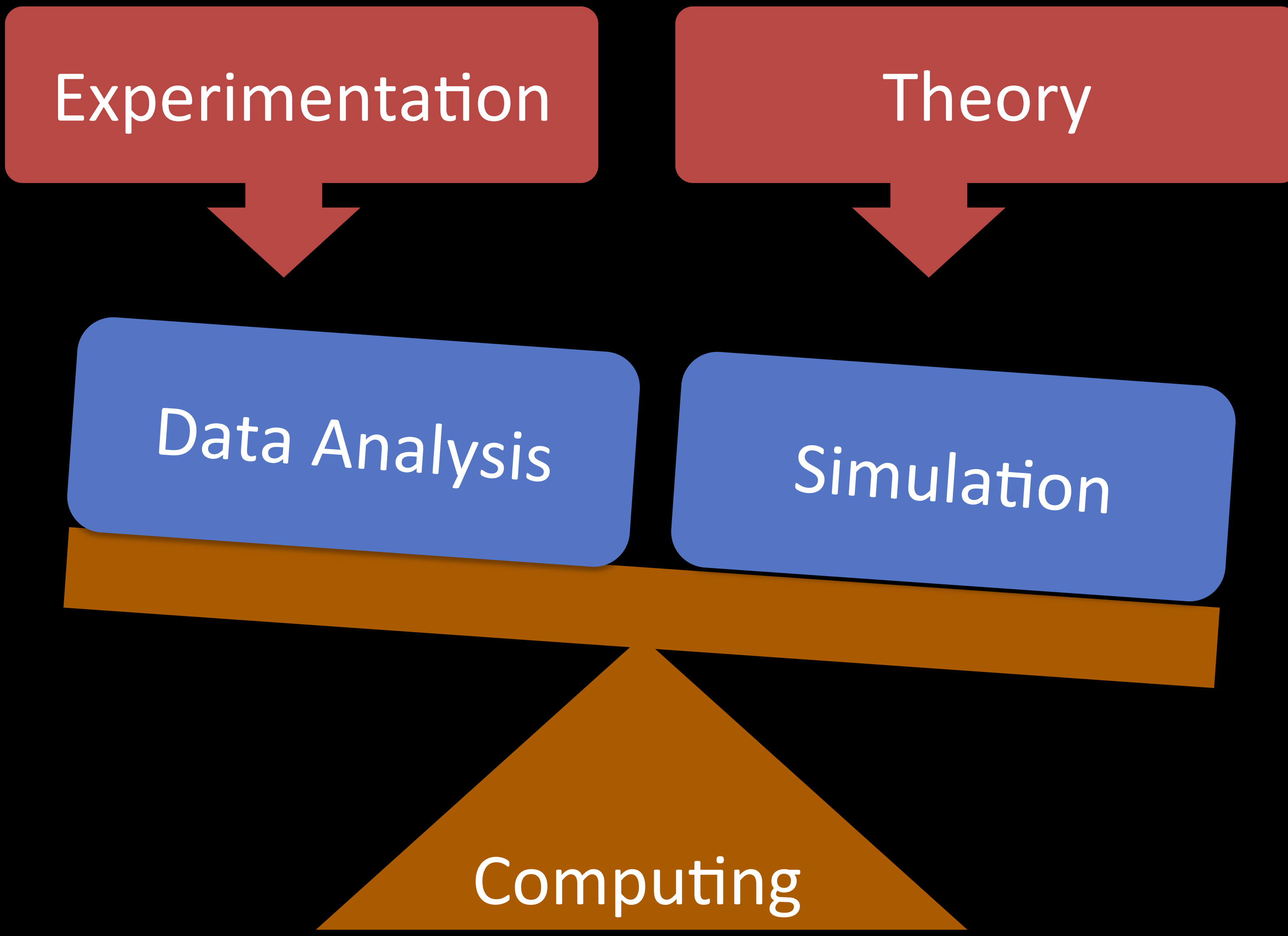


Materials:
- *Variety:* multiple models and experimental data
- *Veracity:* quality and resolution of simulations



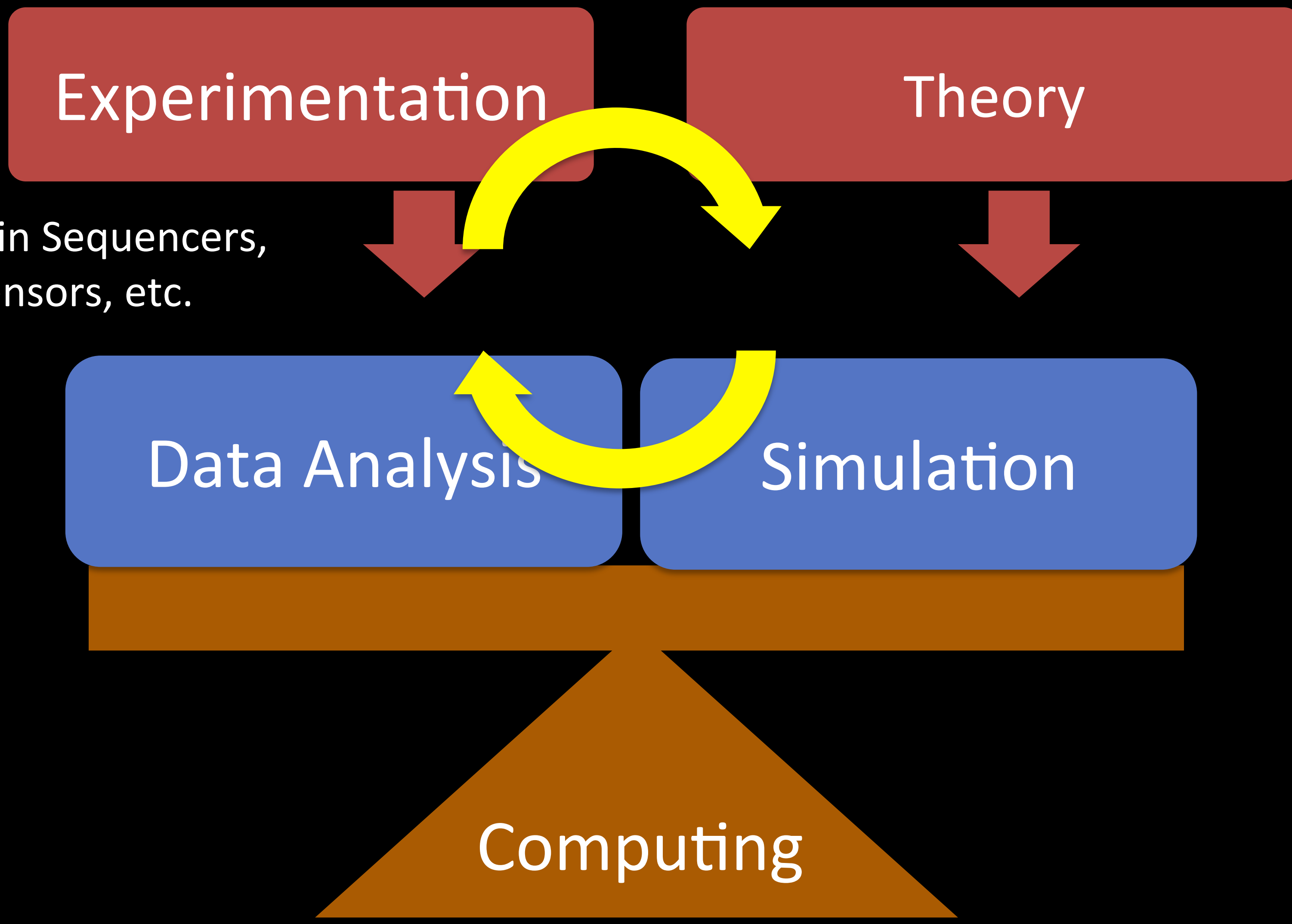
Climate
- *Volume:* Hundreds of exabytes by 2020
- *Veracity:* Reanalysis of 100-year-old sparse data

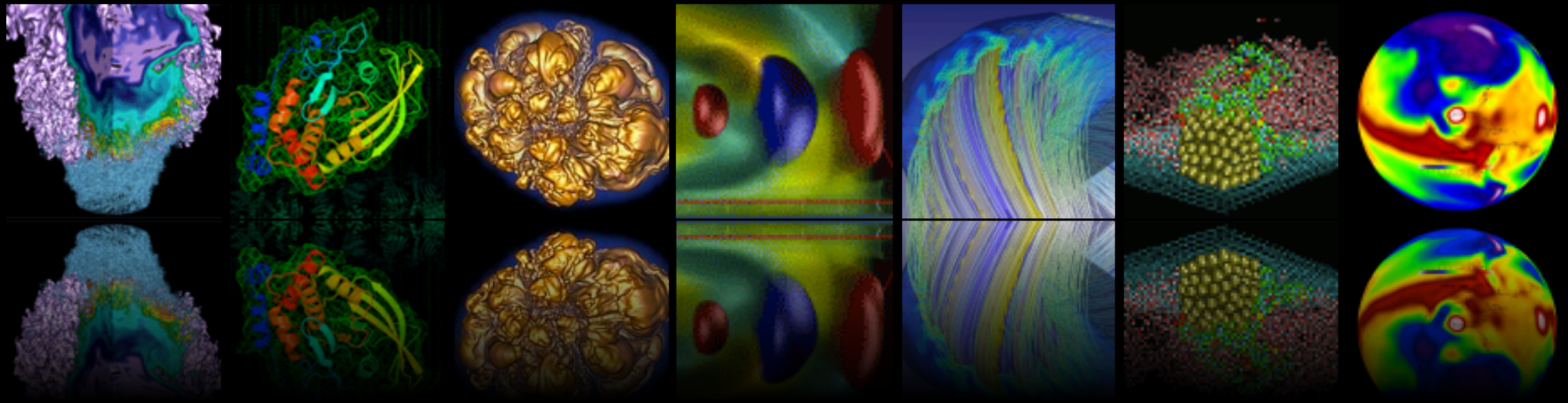
High end computing has focused on simulation



Data analysis is equally important in Science

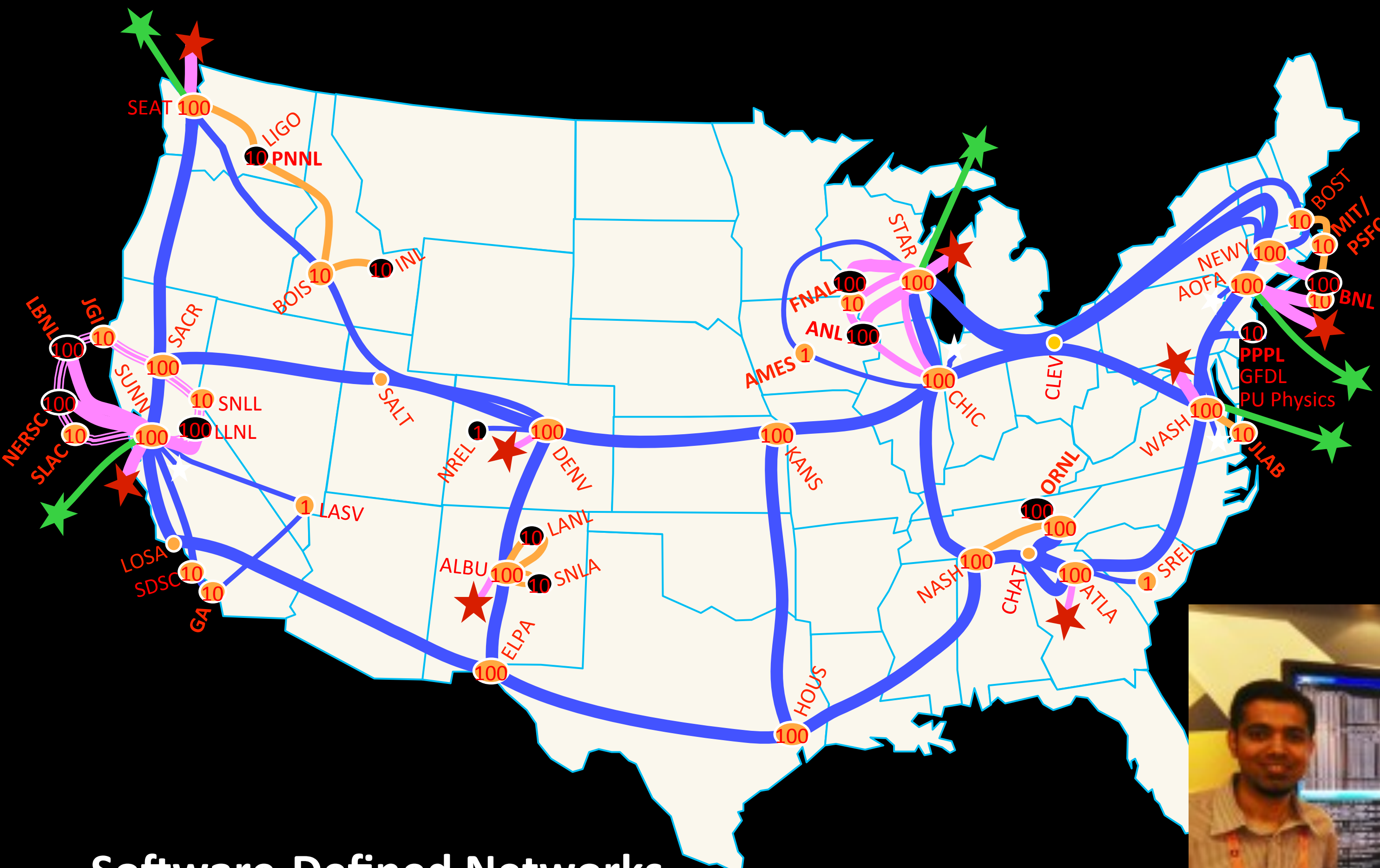
Growth in Sequencers,
CCDs, sensors, etc.





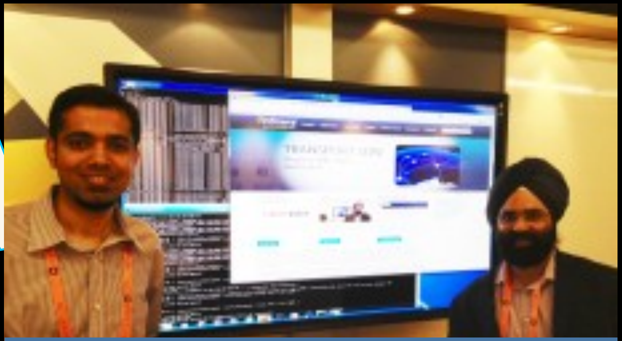
Are there Computer Science Research Challenges?

Network as ~~Infrastructure~~ *Instrument*



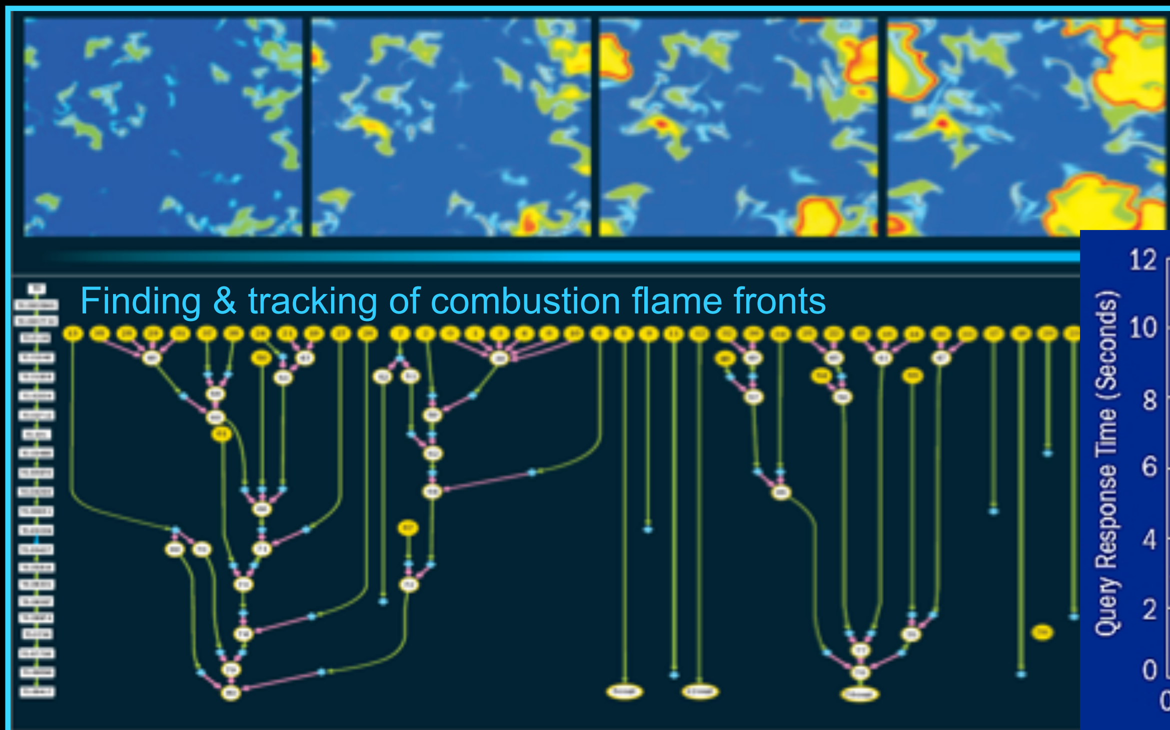
Software-Defined Networks

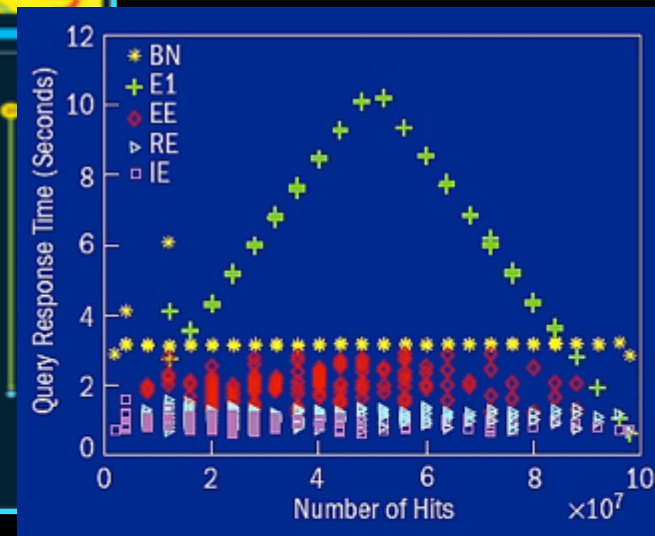
- Infrastructure: black box with complex internals
- Instrument: fast, adaptive, programmable



ESnet, Infinera, and Brocade demo transport SDN

Data structures, Algorithms, and Speed

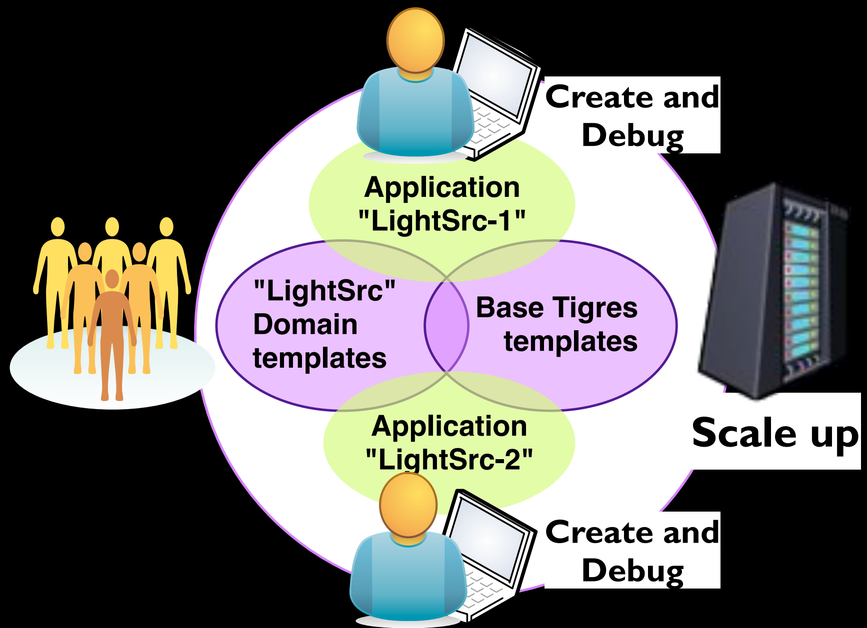




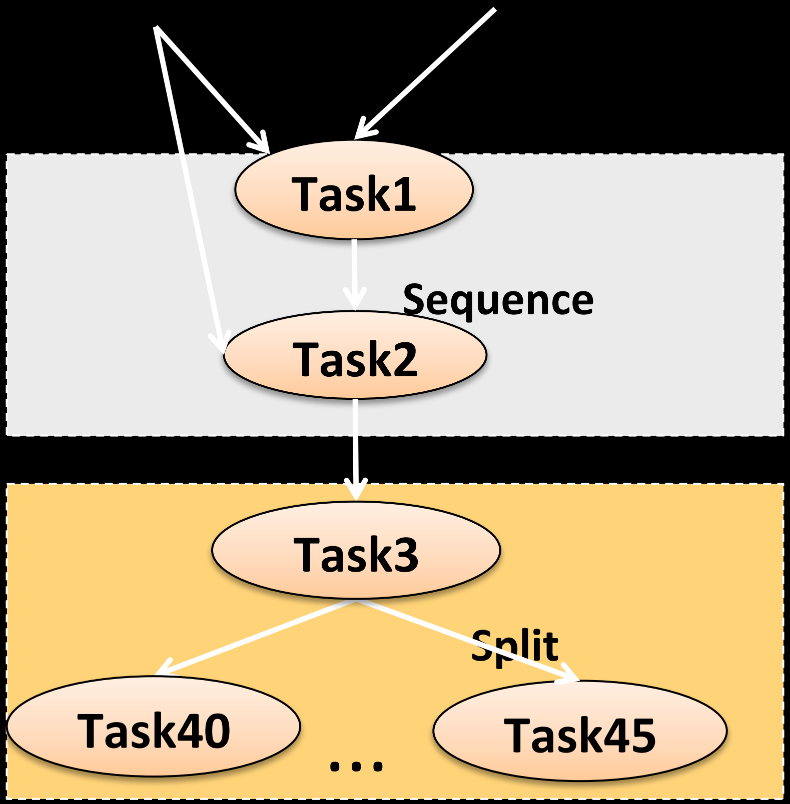
J. Wu, A. Shoshani, A. Sim, D. Rotum

- **Fastbit & Fastquery**
  – specialized compression and object-level search
  – bitmap indexing methods
  – Theoretically optimal and 10x-100x faster in practice

Scientific Workflows

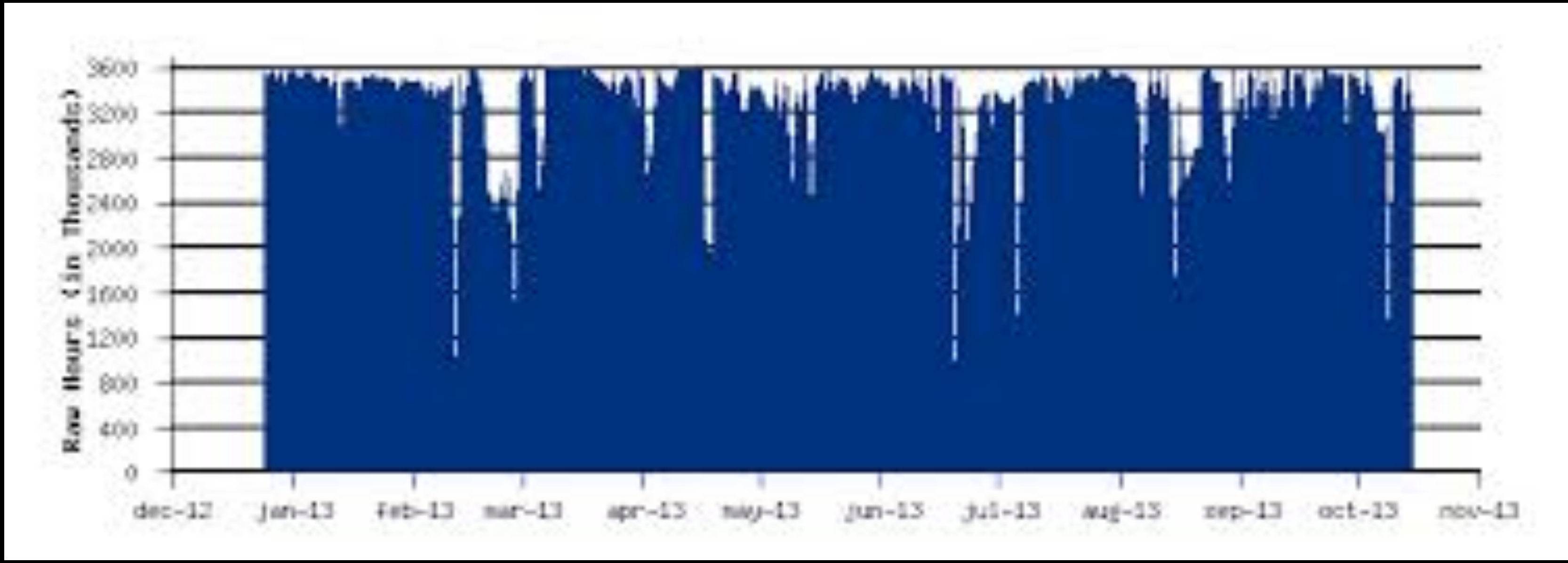


- **Tigres: Design *templates* for scientific workflows**
  - Explicitly support Sequence, Parallel, Split, Merge
- **Fireworks: High Throughput job scheduler**
  - Runs on HPC systems



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

Trade-offs in Utilization vs Response Time



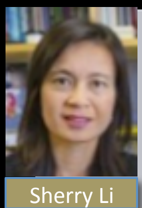
- **95% utilization, but the users wait**
- **Real-time analysis on streams**
- **Interactive access to data**

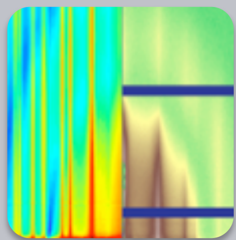
Analytics vs. Simulation Kernels:

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

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

CENTER FOR APPLIED MATHEMATICS FOR ENERGY RESEARCH APPLICATIONS

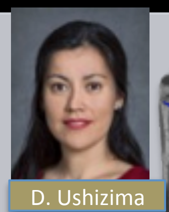


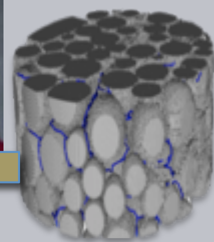


X-ray scattering data analysis

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

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

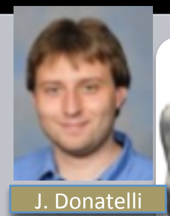




Micro-CT Sample Analysis

Quant-CT provides automated quantitative analysis

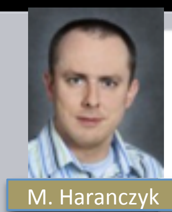
Now: 3D image segmentation, pattern recognition, classification algorithms, PDE- and graph-based analysis

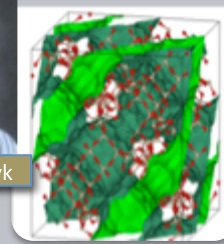


X-ray Nano-Crystallographic Reconstruction

Indexing ambiguity resolved [PNAS13]

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

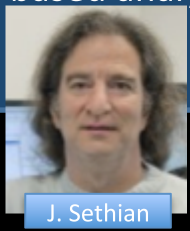




Designing New Materials

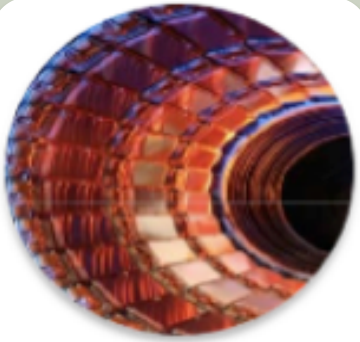
Designed record-breaking high-surface area materials

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

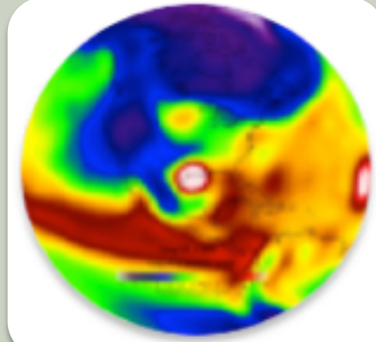


Lead PI, DOE-funded

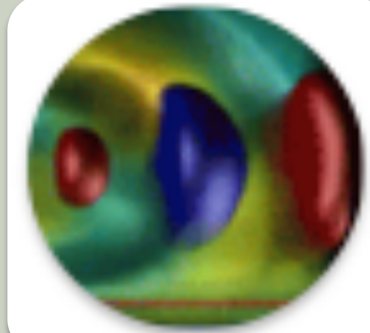
Programming Challenge? Science Problems Fit Across the "Irregularity" Spectrum



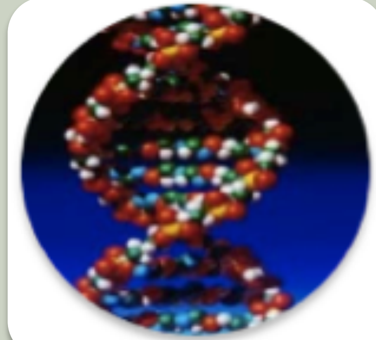
Massive Independent Jobs for Analysis and Simulations



Nearest Neighbor Simulations



All-to-All Simulations

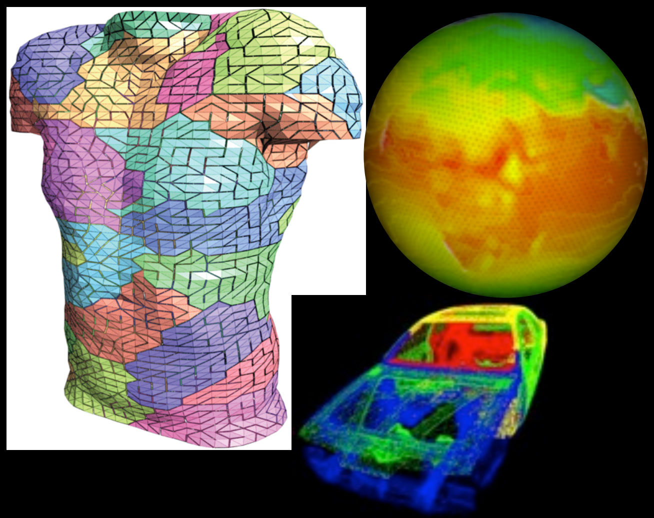


Random access, large data Analysis

… often they fit in multiple categories

The Programming Answer is Obvious…

More Regular



Message Passing Programming

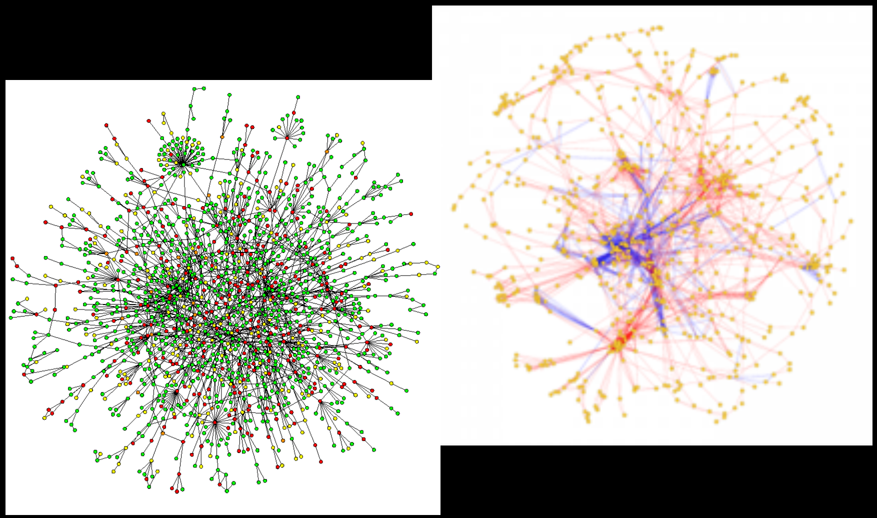
Divide up domain in pieces

Compute one piece

Send/Receive data from others

MPI, and many libraries

More Irregular



Global Address Space Programming

Each start computing

Grab whatever / whenever

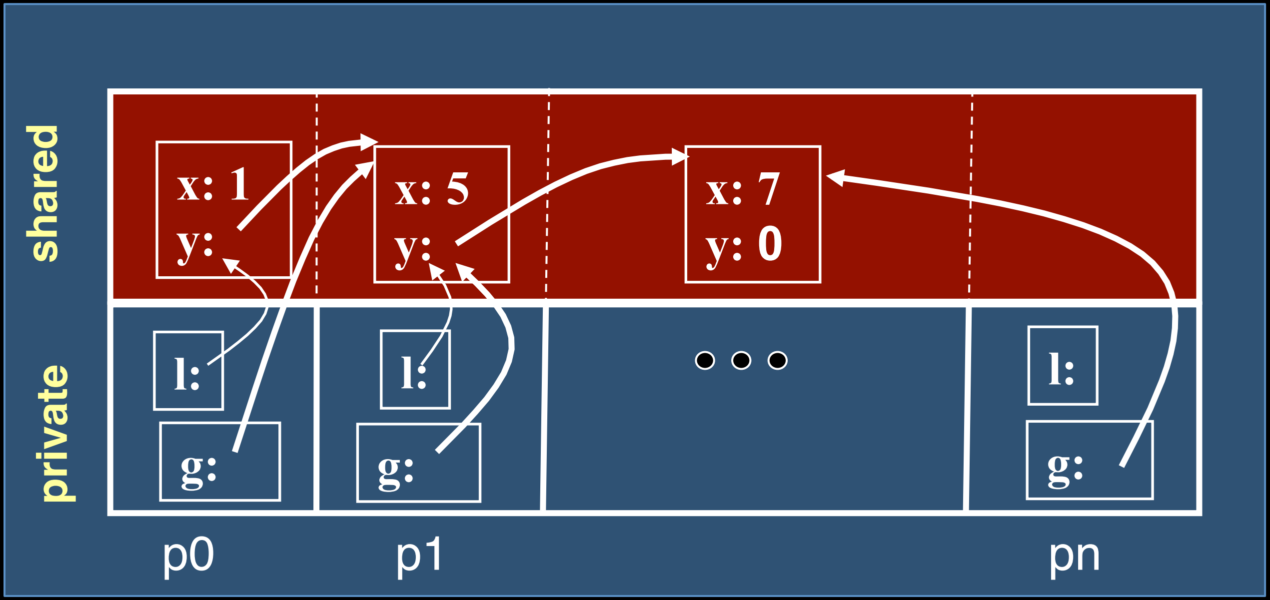
UPC, CAF, X10, Chapel, GlobalArrays

Partitioned Global Address Space for Convenience and Performance

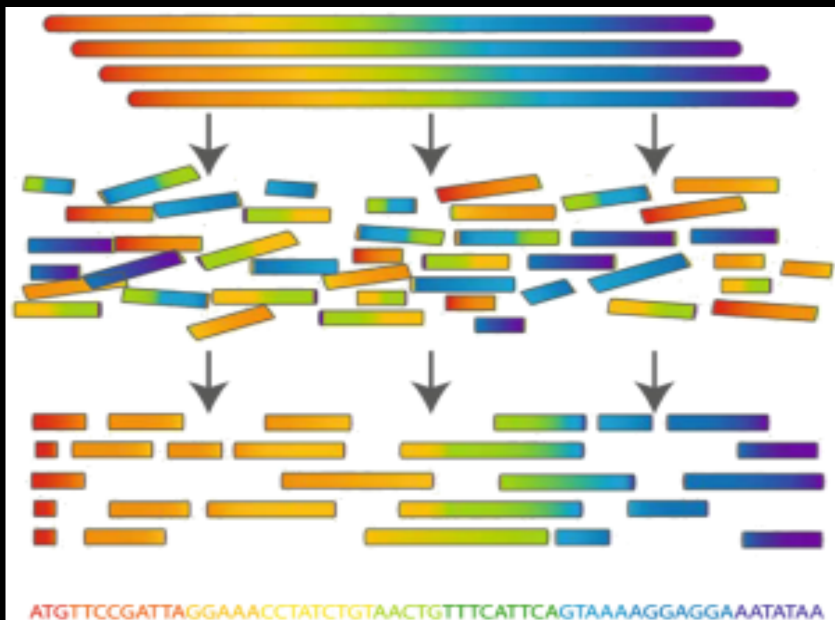
PGAS: Partitioned Global Address Space

Global address space: **directly read/write remote data**

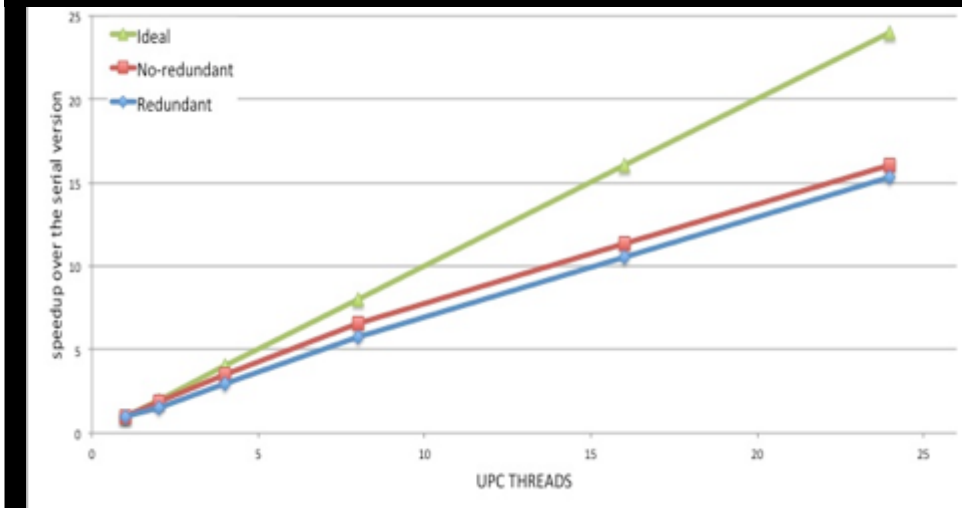
Partitioned: **data is designated as local or global**



Programming Models for Analytics



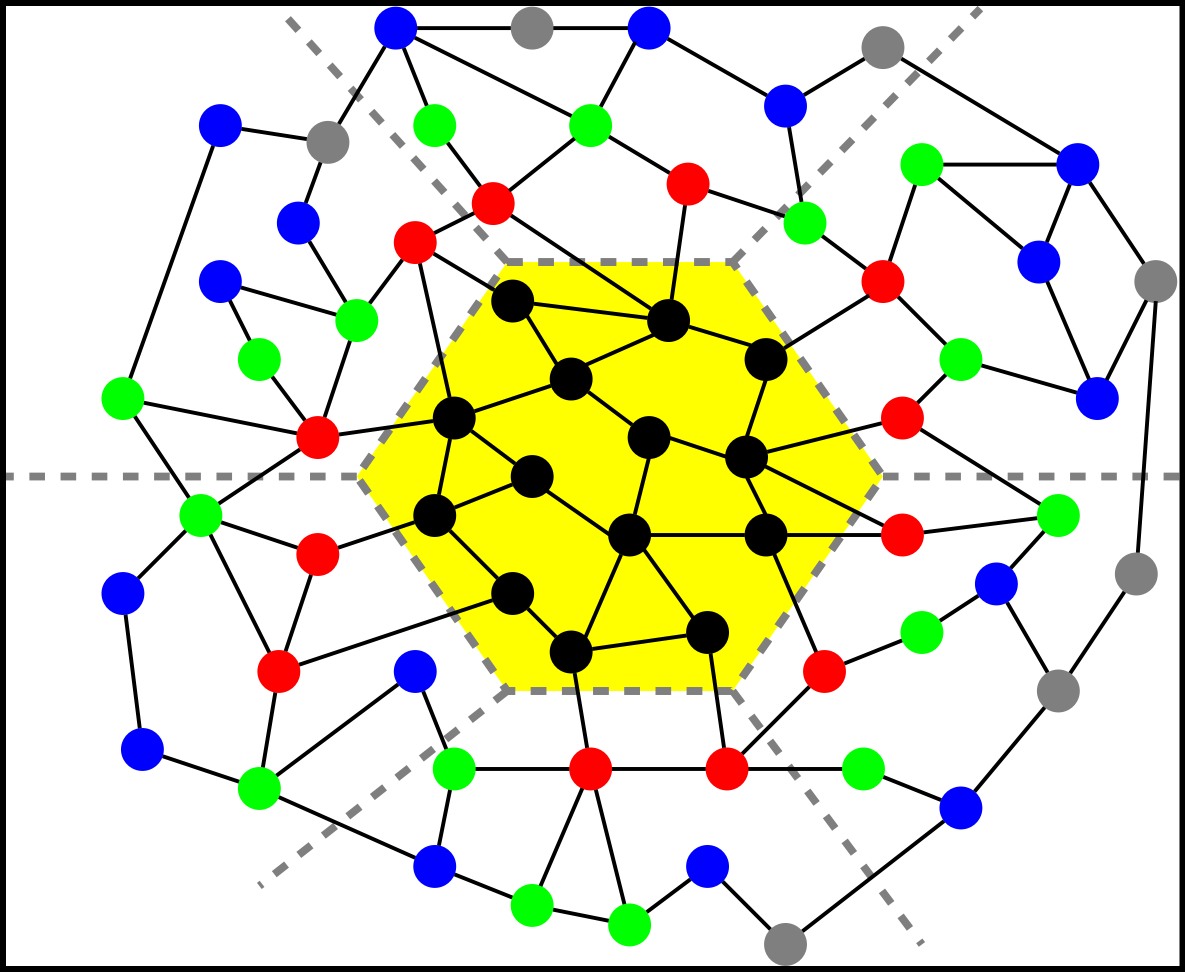
Strong Scaling of Meraculous Assember in UPC



- **Computational Biologists buy large shared memory machines to assemble genomes**
- **For many problems (including metagenomics) these are not large enough**

Work by Evangelos Georganas, Jarrod Chapmanz, Khaled Ibrahim, Daniel Rokhsar, Leonid Oliker, and Katherine Yelick

Communication Avoiding Algorithms on Sparse Matrices (aka Graphs)

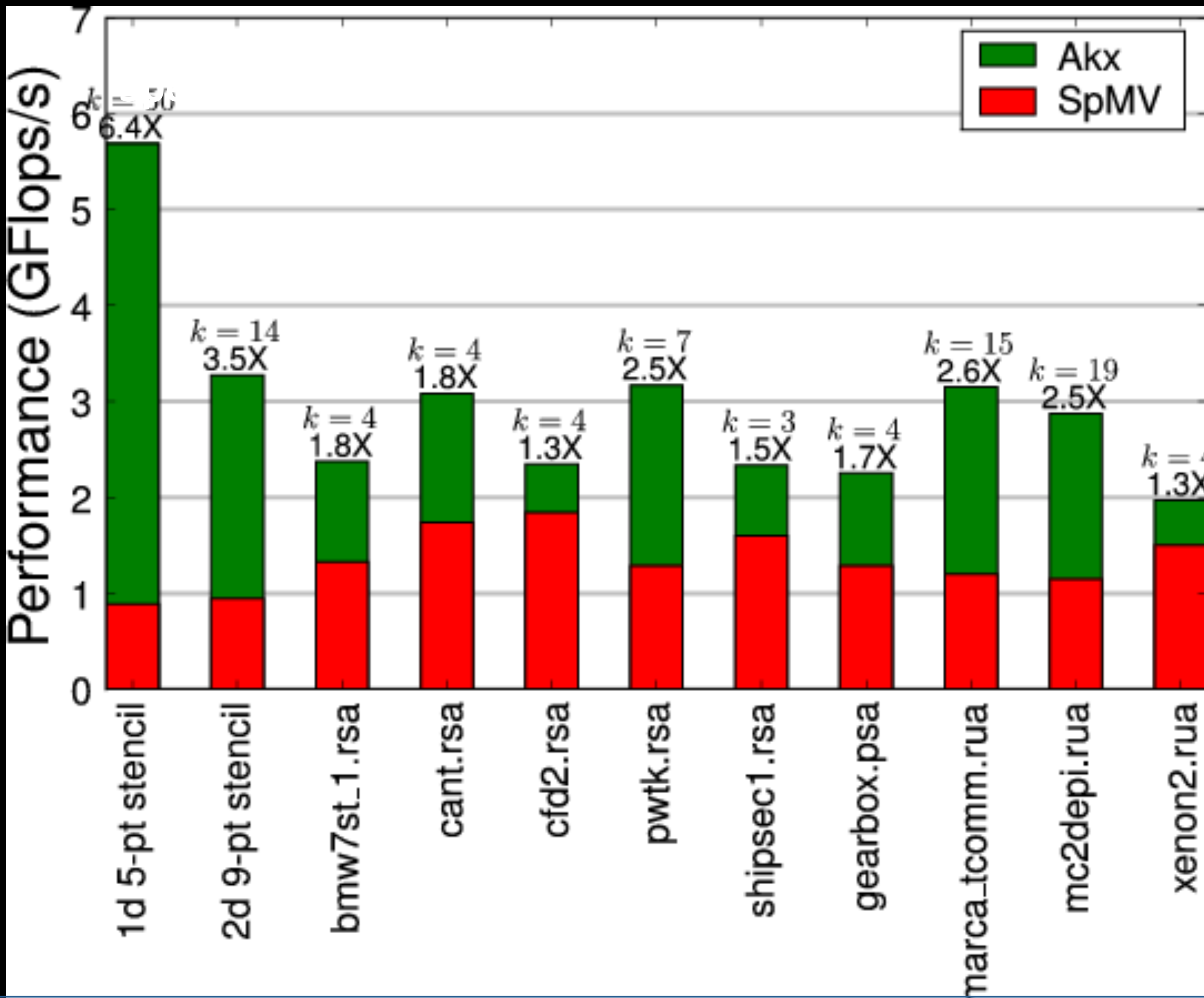


For implicit memory management (caches) uses a TSP algorithm for layout

Joint work with Jim Demmel, Mark Hoemman, Marghoob Mohiyuddin

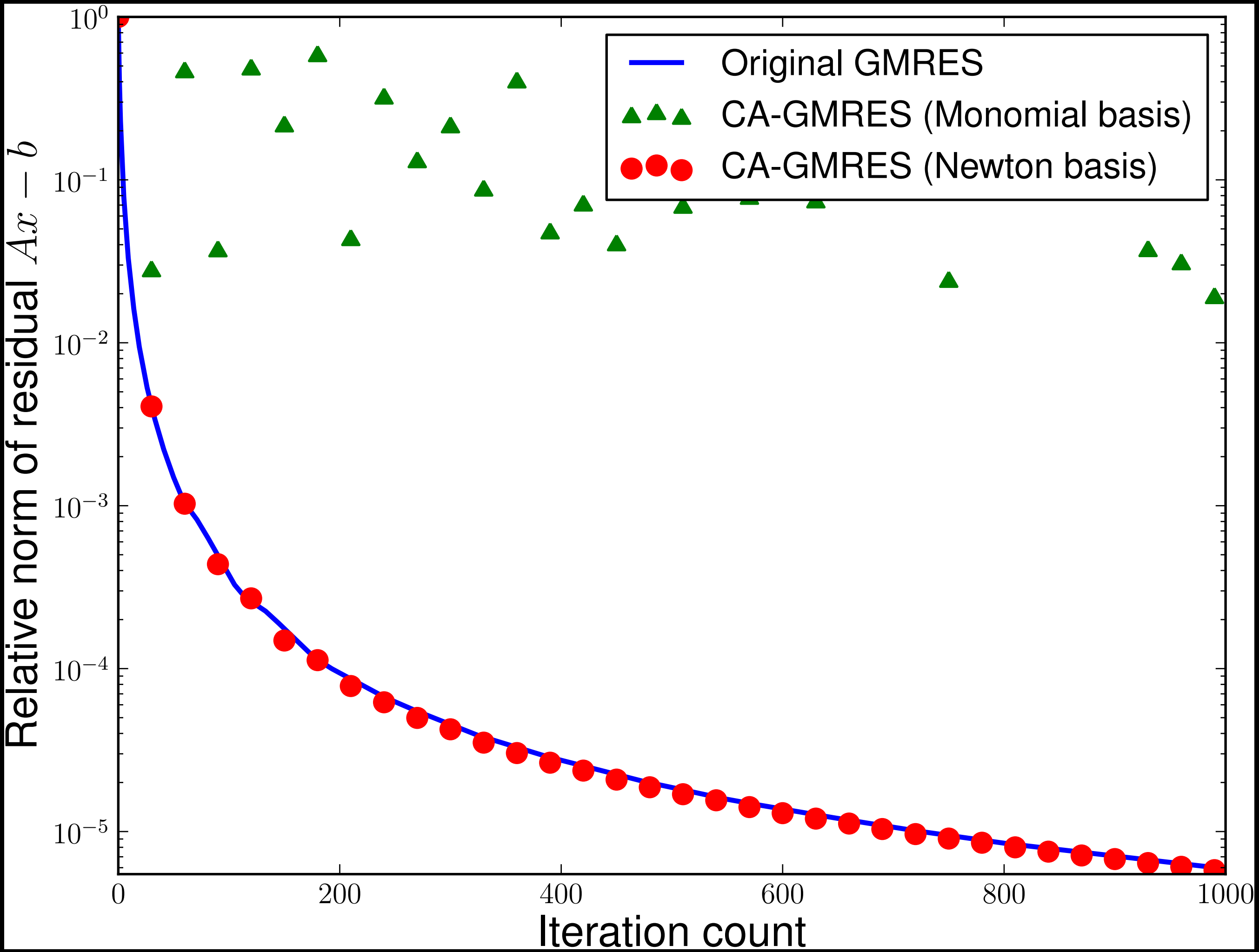
- **Can do better: 1 matrix read, multiple multiplies**
  - Serial: O(1) moves of data moves vs. O(k)
  - Parallel: O(log p) messages vs. O(k log p)

Multi-Step (A^kx) Runs at Faster Speed than Simpler (Ax)



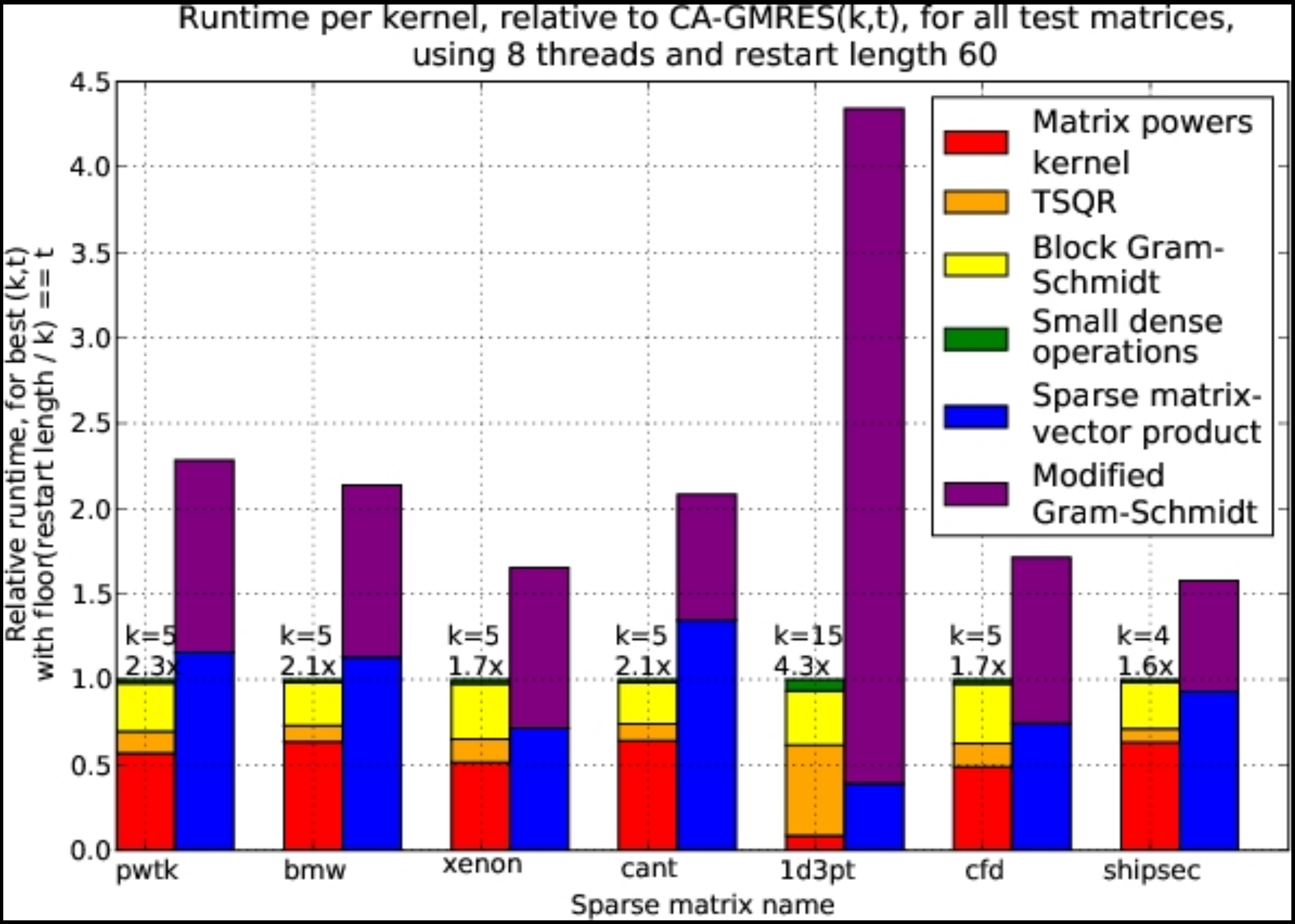
Jim Demmel, Mark Hoemmen, Marghoob Mohiyuddin, Kathy Yelick

Matrix Powers Kernel (and TSQR) in GMRES

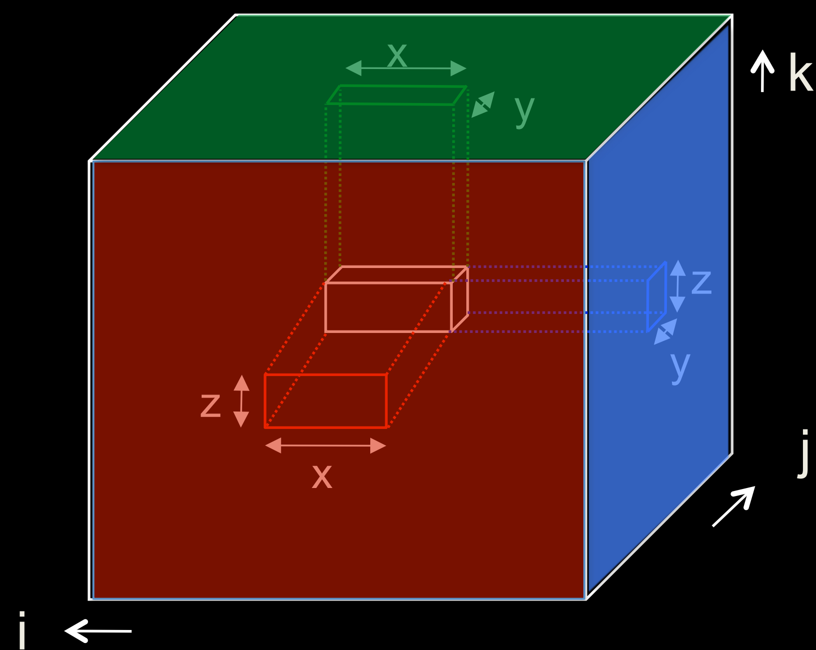


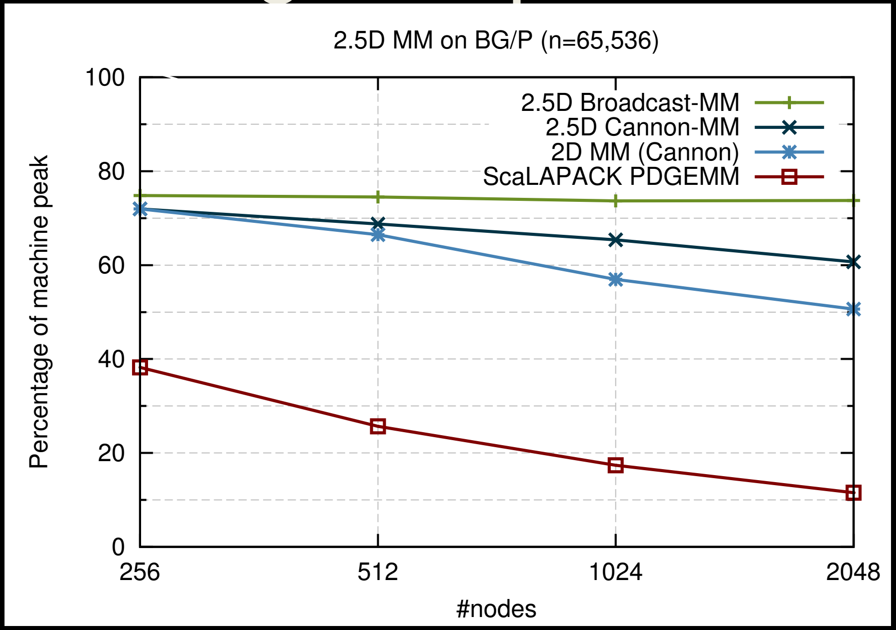
Communication-Avoiding Krylov Method (GMRES)

Performance on 8 core Clovertown



Towards Communication-Avoiding Compilers





Matrix Multiplication code has a 3D iteration space
Each point in the space is a constant computation (*/+)

for i, for j, for k C[i,j] … A[i,k] … B[k,j] …

These are not just "avoiding," they are "communication-optimal"

Technology for Scientific Data

	Compute Intensive	Data Intensive
Compute		
On-Package DRAM		
Capacity Memory		
On-node-Storage		
In-Rack Storage		
Interconnect		
Global Shared Disk		
Off-System Network		

Compute Intensive — Goal: *Maximize computational capacity and local bandwidth*

Data Intensive — Goal: *Maximize data capacity and global bandwidth*

Scientific Data Initiative at Berkeley Lab

ESnet

$$\frac{dVar(t)}{dt} =$$

$$\ldots - K^2Fg + gK^2F)ds - \int_{s_2(t)}^{s_1(t)+S} (-(g^{-1}F_s)_s - K\ldots$$

$$= -\int_{s_1(t)}^{s_2(t)} (g^{-1}F_s)_s \, ds + \int_{s_2(t)}^{s_1(t)+S} (g^{-1}F_s)_s \, ds$$

$$\ldots{}^{-1}F_s \,|_{s_2(t)} - g^{-1}F_s \,|_{s_1(t)} \Big] + \Big[g^{-1}F_s \,|_{s_1(t)+S} - g\ldots$$

$$= -2\,(g^{-1}F_K K_s)\,|_{s_2(t)} + 2\,(g^{-1}F_K K_s)\,|_{s_1(t)}$$

Leverage Berkeley Lab talent in math, computer science, interdisciplinary team science, networking, software engineering and our new infrastructure to enable new modes of inquiry and discovery from scientific data sets

Thanks to Collaborators, Students, Postdocs, Program Managers

- Jim Demmel
- Paul Hargrove
- Costin Iancu
- Bill Carlson
- Lauren Smith
- Tarek El-Ghazawi
- Khaled Ibrahim
- Erich Strohmaier
- Eric Roman
- Bob Lucas
- Yili Zheng
- Greg Balls
- Randi Thomas
- Carelton Miyamoto
- C.J. Lin
- Deborah Weisser
- Ngeci Bowman
- Arvind Krishnamurthy
- Soumen Chakrabarti
- Chih-Po Wen
- Jeff Jones
- Steve Steinberg
- Rich Vuduc
- Leonid Oliker
- Sam Williams
- Michael Driscoll
- Evangelos Georganas
- Penporn Kaonantakool
- Shoaib Kamil
- Amir Kamil
- Dan Bonachea
- Jimmy Su
- Rajesh Nishtala
- Kaushik Datta
- Brian Kazian
- Wei Chen
- Jason Duell
- Christian Bell
- Mani Narayanan
- Ed Givelberg
- Sabrina Merchant
- Etienne Deprit
- Noah Treuhaft
- Siu Man Yau
- Eun-Jin Im