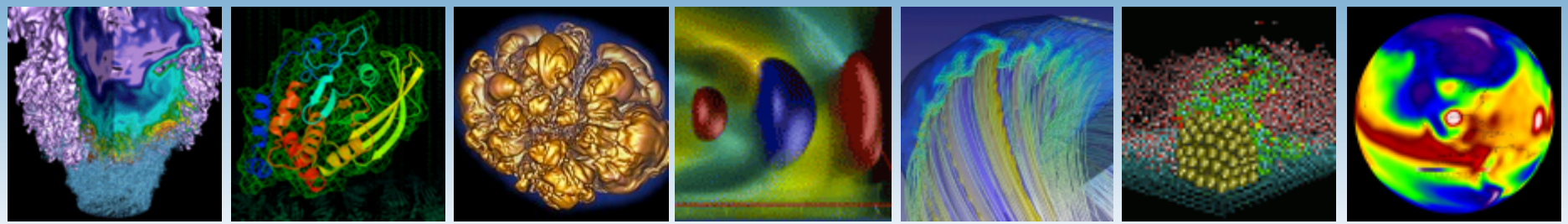


Open Problems, Solved Problems and Non-Problems in DOE's Big Data

Kathy Yelick

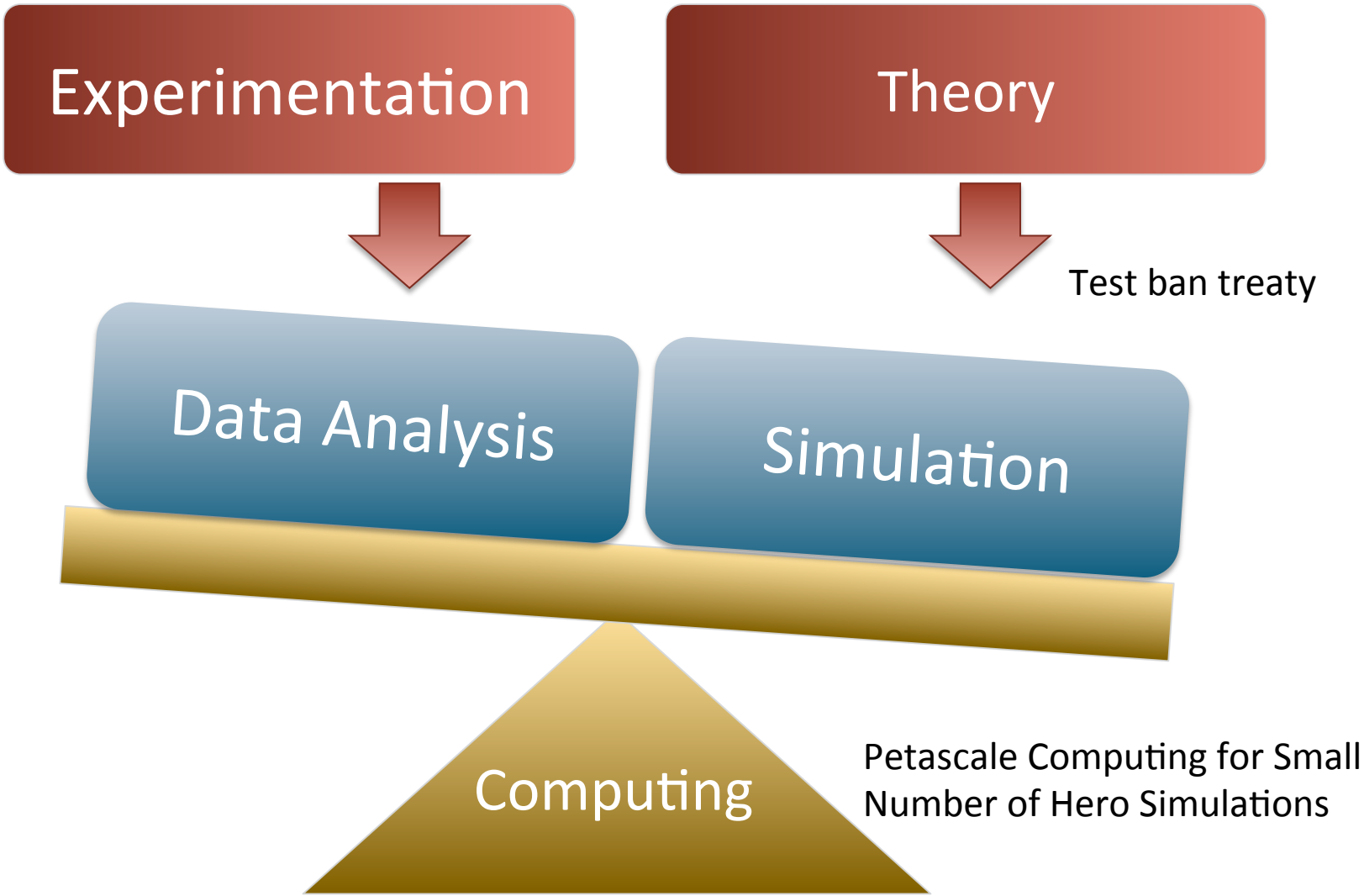
Professor of Electrical Engineering and Computer Sciences
University of California at Berkeley

Associate Lab Director for Computing Sciences
Lawrence Berkeley National Laboratory

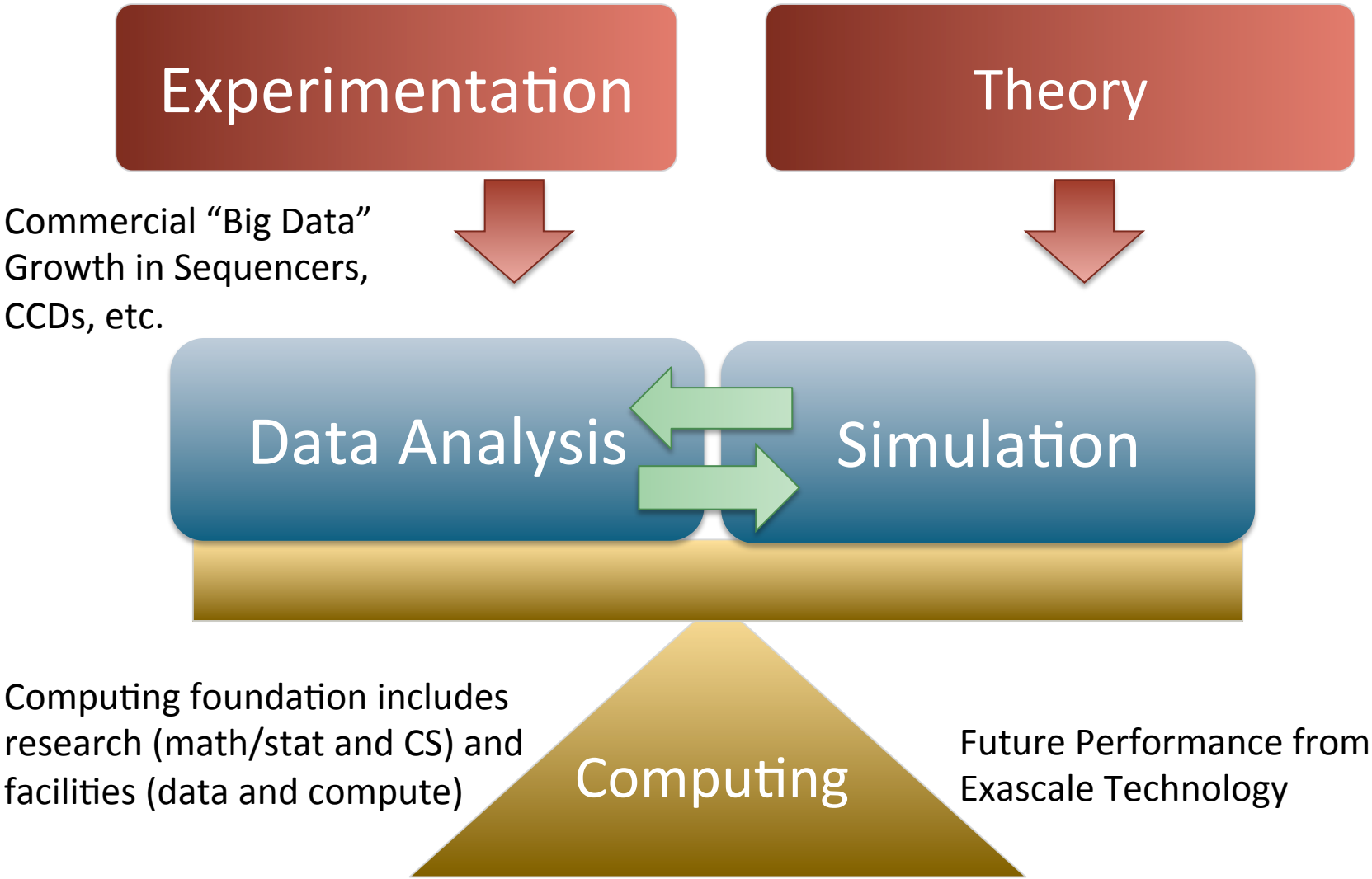
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

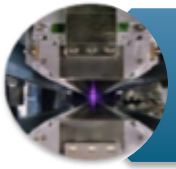
Myth #1: DOE Computing = Modeling and Simulation



DOE Should Have a Balanced Strategy of Experiment (Data) and Theory (Modeling and Simulation)

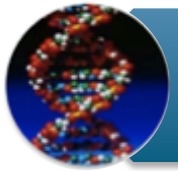


DOE has Big Data Needs

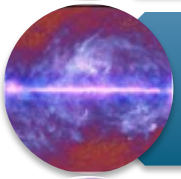


Light Sources

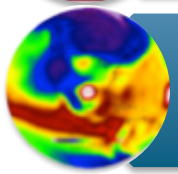
Particle Physics



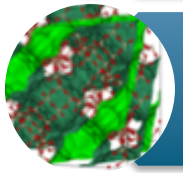
Genomics



Astronomy

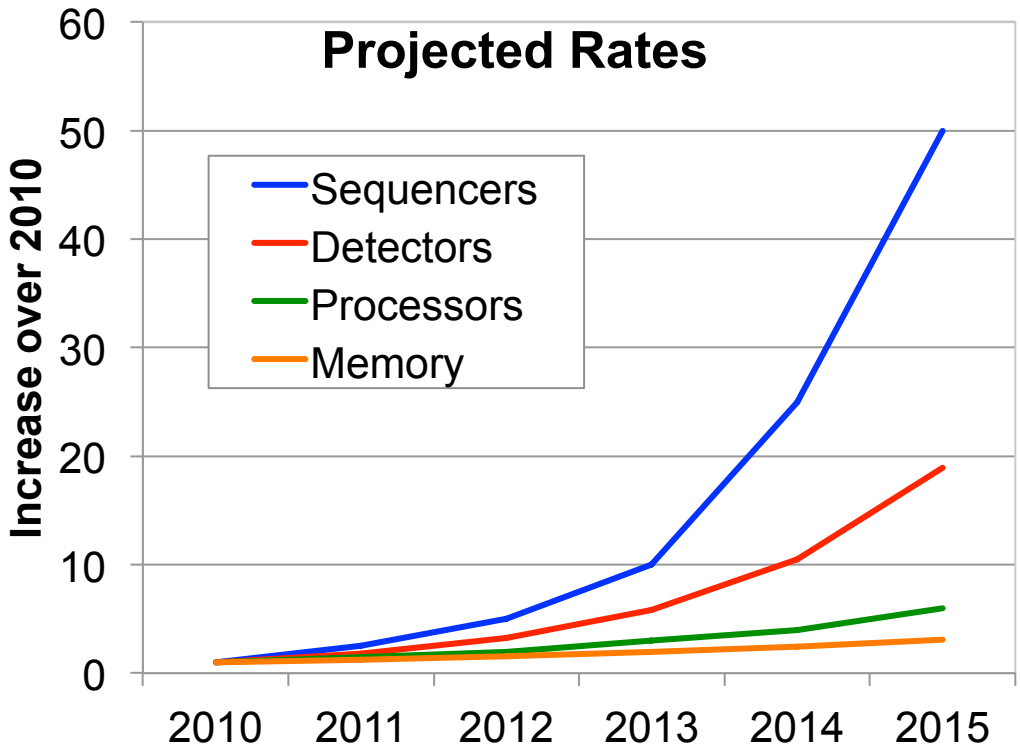


Climate

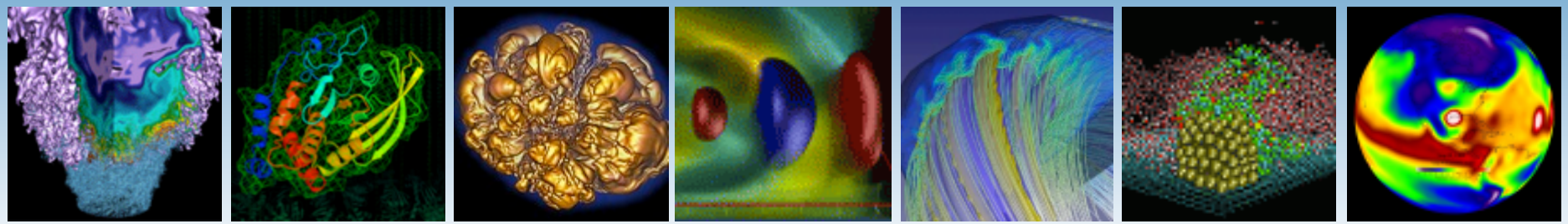


Materials Omics

Data Growth is Outpacing Computing Growth



- **Sequences and detectors (used in light sources, particle accelerators) are growing faster than processor speeds**
- **And all are growing faster than memory density**

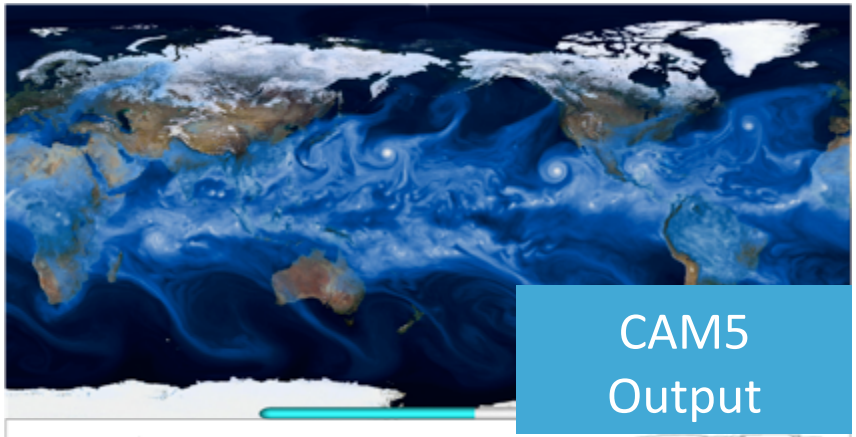


Solved Problems

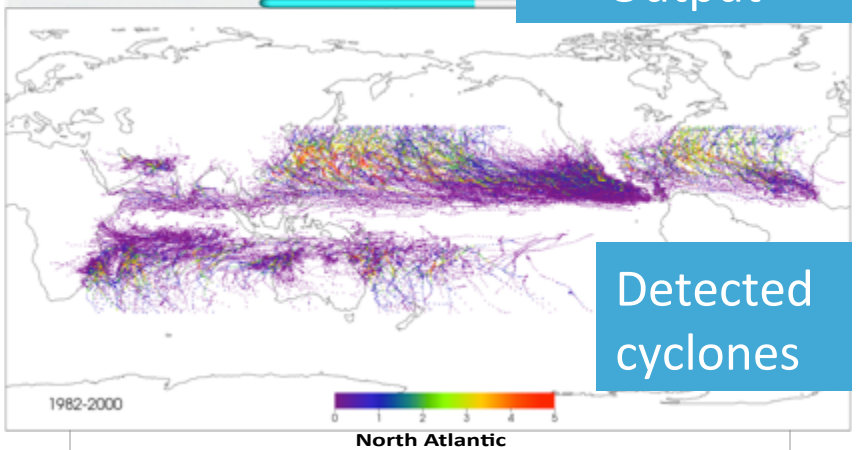
Example results-to-date (and case for more Big Data research)

DOE Data Requires Sophisticated Algorithms and Software: Extreme Climate Events

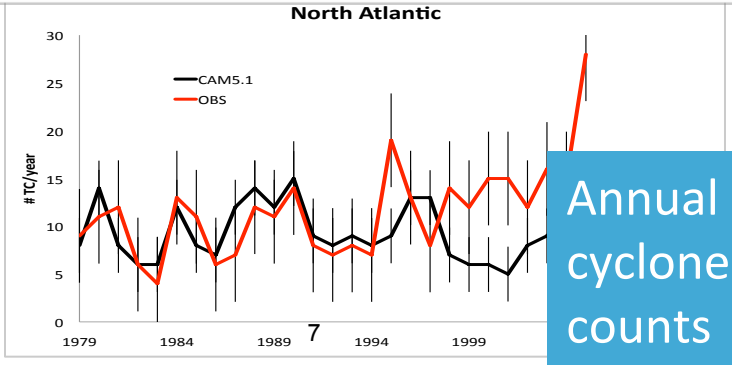
- "Extreme weather events" are features hidden in large collections of data.
- Research on feature detection, tracking, analysis:
  - Reduces processing time from years/months to hours/minutes.
  - New software: Toolkit for Extreme Climate Analysis (TECA)
  - Standalone tool (initial), deploying in VisIt for broader applicability.
- New data mining, image processing, and topological analysis techniques



CAM5 Output



Detected cyclones



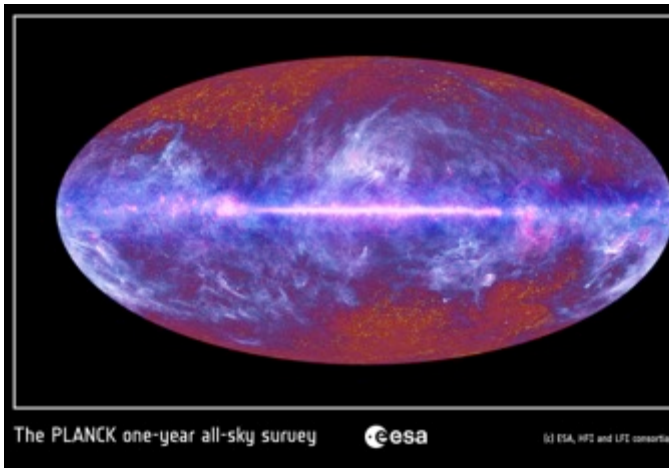
Annual cyclone counts

Michael Wehner and Prabhat, LBNL

U.S. DEPARTMENT OF ENERGY | Office of Science

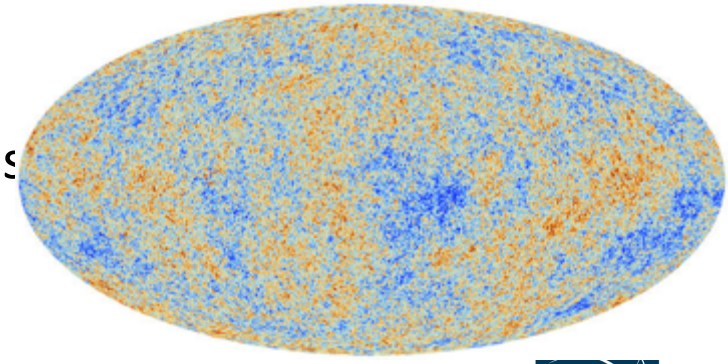
Simulation is Key to Experimental Setup and Data Analysis

- **CMB data analysis pre-mission**
  - Quantify/correct biases and uncertainties
  - Validate and verify methods
- **Each realization involves simulating and mapping the entire mission**
  - 10^4 realizations, $10^{9\text{-}15}$ samples => 10^8 pixels
  - Algorithms & implementations evolve
  - Largest simulation: 250,000 maps, 10M CPU-hours, 50TB disk
- **Planck ESA/NASA satellite mission**
  - Simulations handle survey inefficiencies the intervening Milky Way galaxy, and imperfect detectors.



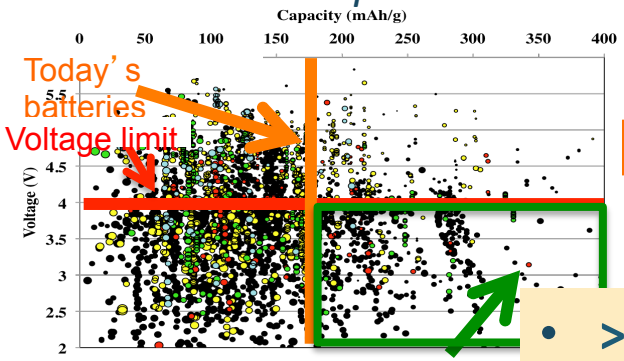
The New York Times

Universe as an Infant: Fatter Than Expected and Kind of Lumpy

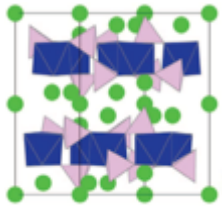


Materials Genome Turns Massive Simulations into Data Analysis, Storage and Delivery

Materials Properties



Atomic positions

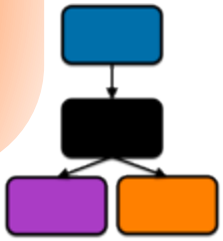


- > 3500 users since 10/2011 start
- > 10 million hours in < 6 months
- > 30,000+ materials
- **Used in industry, teaching, universities, & labs (e.g., JCESR)**
- **Machine Learning (MLBase from AMPLab): 85% accurate prediction of runtime**

Materials Data

Submitted Materials

Supercomputers



Calculation workflow

NSF AMPLab at UCB

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

Biology in DOE is Data Intensive

- **Biology is important / growing area in DOE**
- **DOE's focus on Metagenomes**
  - IMG Community resource for comparative analysis
  - About 10x growth in genes in 2012 (now at 16B)
    - *Data streams over Esnet, stored & computed at NERSC*
    - *Weeks to days for 100M gene sets*
- **BER's KBASE Project**
  - ***Knowledgebase*** enabling ***predictive*** systems biology, e.g., microbes, plants, communities
- **Leveraging UCB, UCSF and others**
  - LDRD on graph algorithms for biology
  - New joint postdoc in Simons Institute

Radial Phylogenetic Tree

Simon's Institute at UCB

KBASE

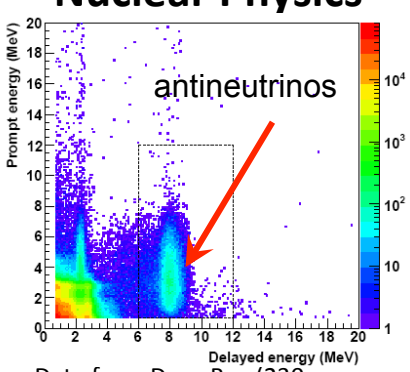
DOE Systems Biology Knowledgebase

Data Science in DOE has Huge Impact

- **Big:** NERSC imports 1 PB per month; ESnet transfers 10 PB per month and this rate grows exponentially (2x commercial nets)
- **Useful:** Over 9,000 data users access "Science Gateways" at LBNL & NERSC via ESnet
- **Hard:** Data is noisy, incomplete, unstructured, and heterogeneous

length, time →

High Energy and Nuclear Physics



Data from Daya Bay (230 users), LHC (1000), STAR (500) take ESnet to NERSC

Materials

Materials Project stores 30K simulations at NERSC in web-accessible database (3500 users)

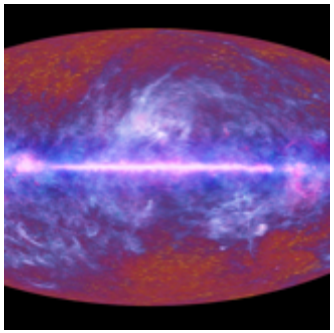
Biology

LBNL scientists use NERSC, ESnet and CRD software to define healthy microbiome; IMG variants have 2500 users

Environment

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

Cosmology

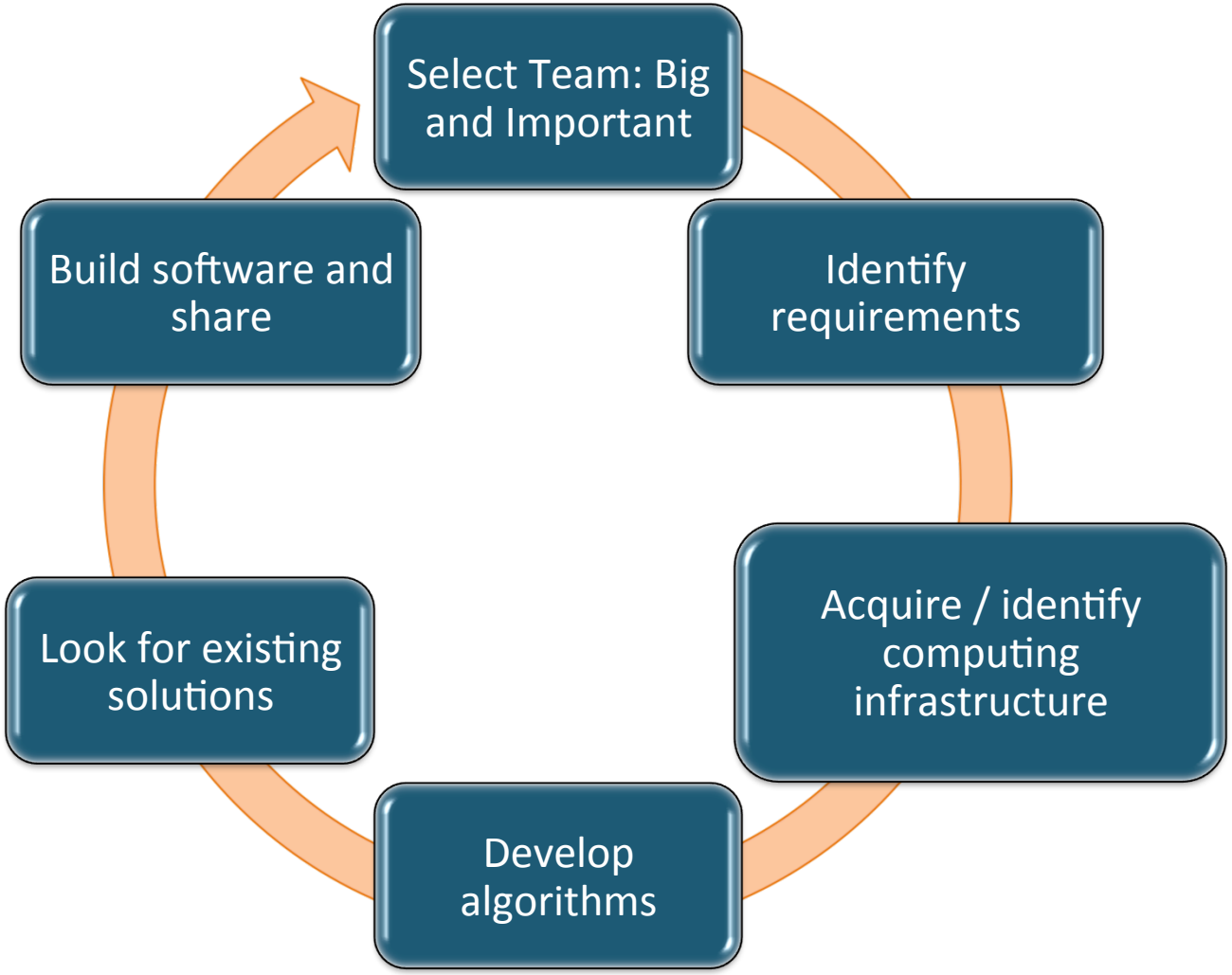


Planck data pipeline (100 users) and PTF (50) at NERSC

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

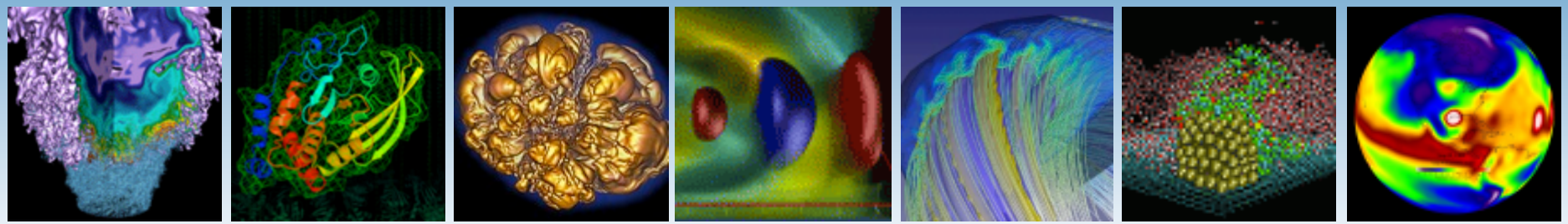
Successful Partnership Model for Data Projects



Data & Simulation are Intertwined
And Both need a Growth in Computing Performance

- **DOE facilities have big data needs**
  - Growth from detectors, sequencers, sensors, etc.
- **Facilities use simulation in design**
  - Risk mitigation and engineering
- **DOE computing generates big data**
  - Massive scale and massive throughput simulations
- **Data access improves science**
  - Community data sets democratize and improve quality
- **Data analytics uses big computing**
  - Both capacity and novel architectures
- **Data requires simulation**
  - Missing, noisy data; models to interpret

Take-home message



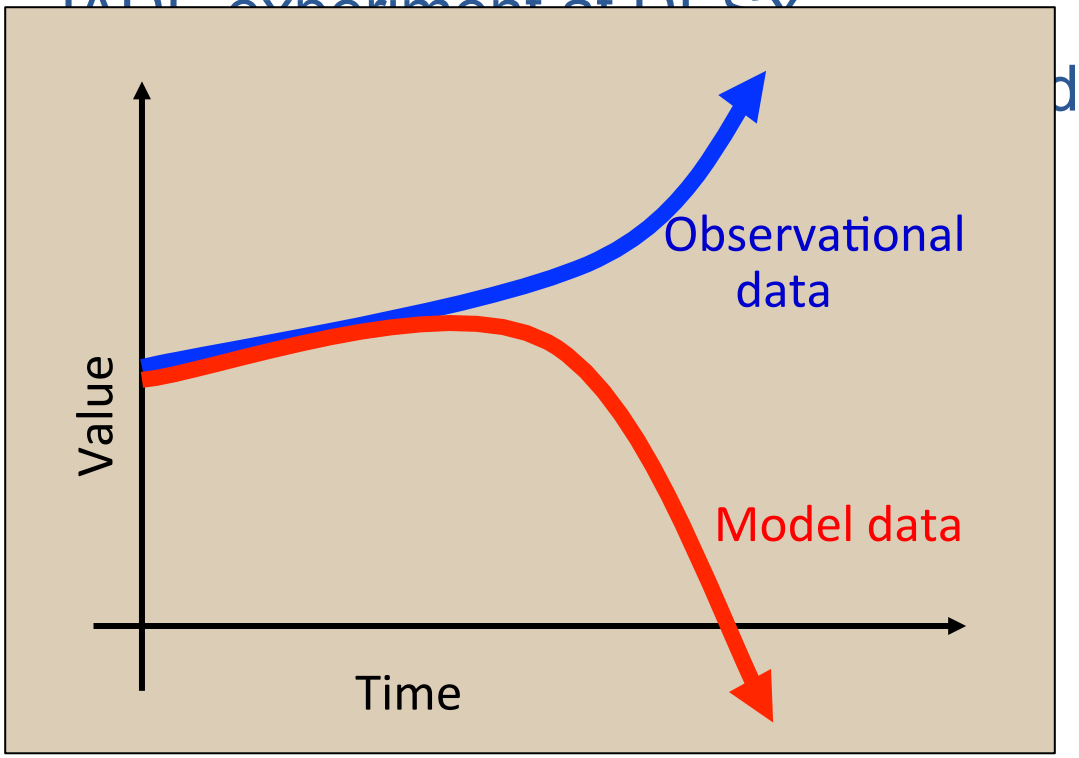
Open Problems

DOE Data is not just Big

Tales of a Particle Physicist

Sigfried Bethke working on the JADE detector in 1984.

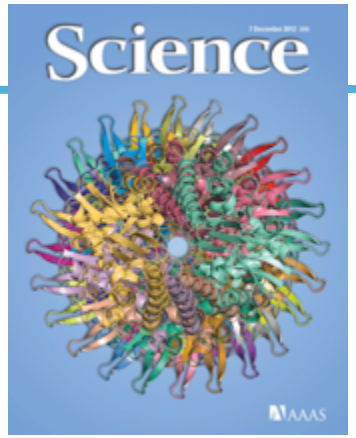
- Siegfried Bethke was part of the JADE experiment at DESY

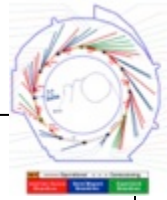


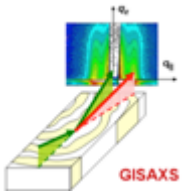
Source: L. Buja

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

ALS Scientific Workflow today



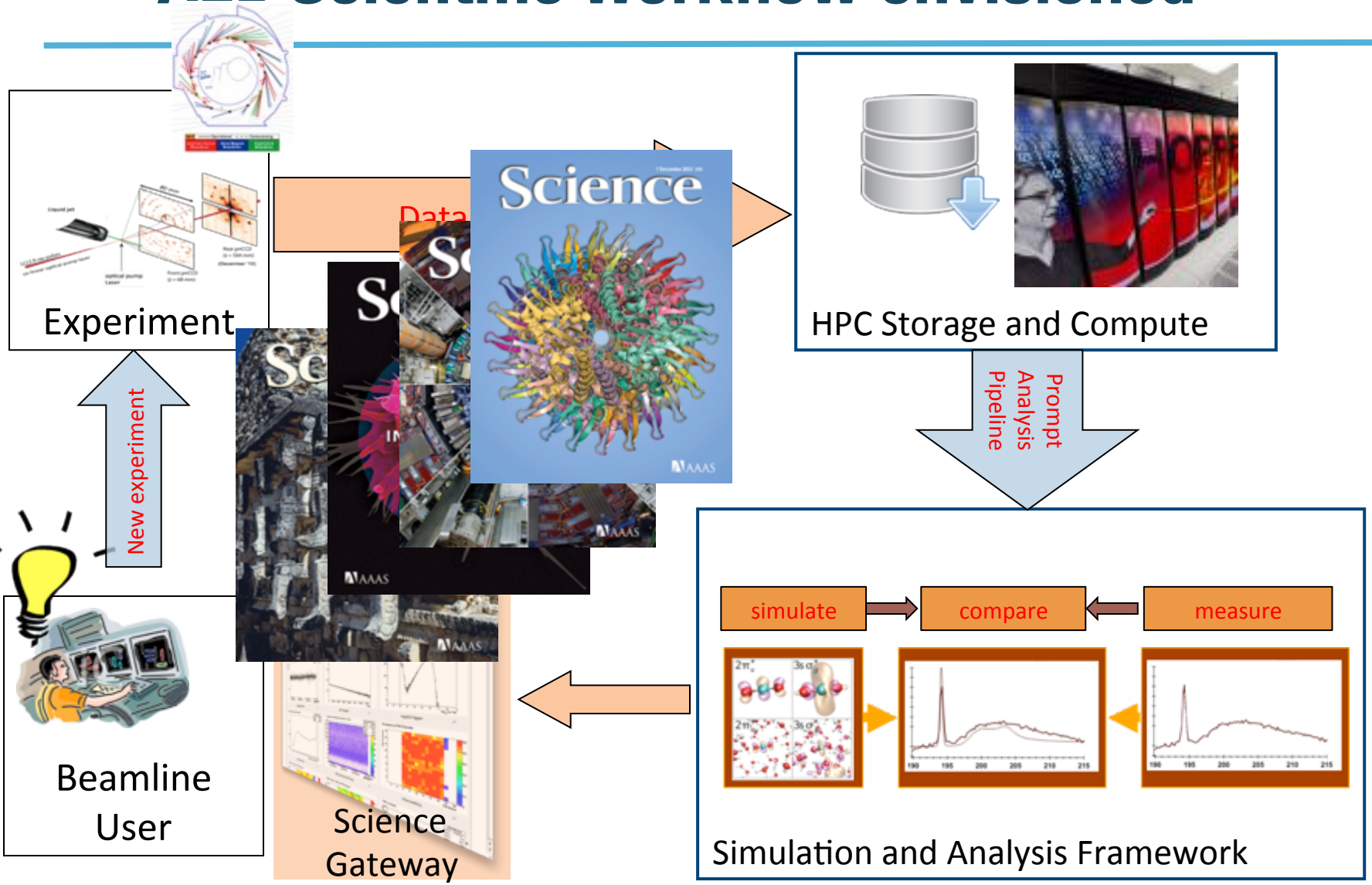




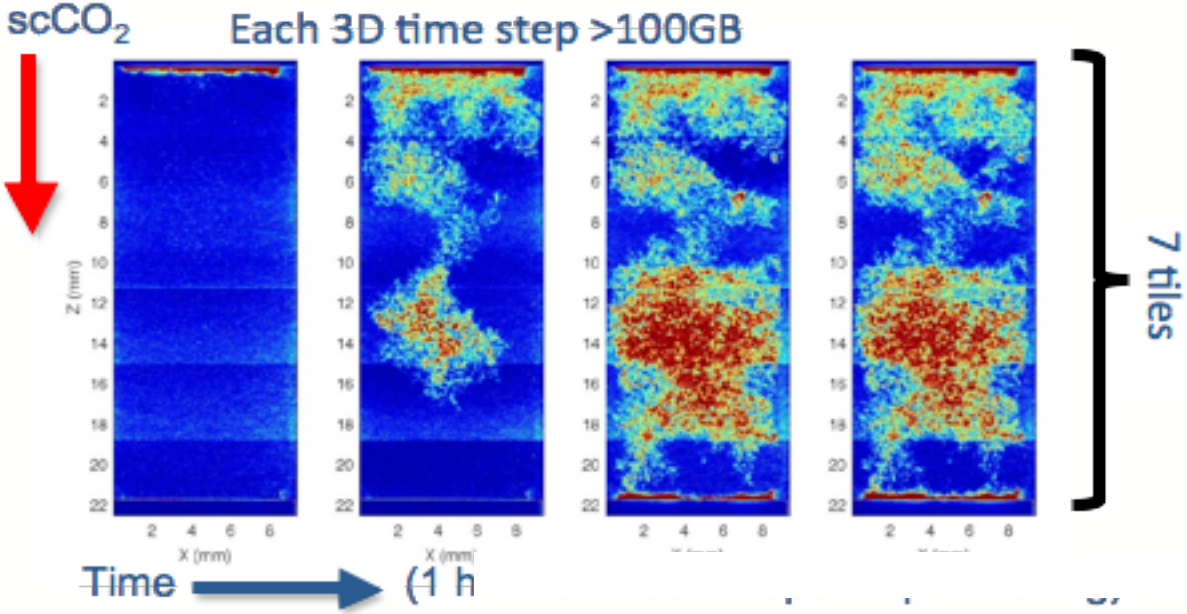
Experiment

Beamline User

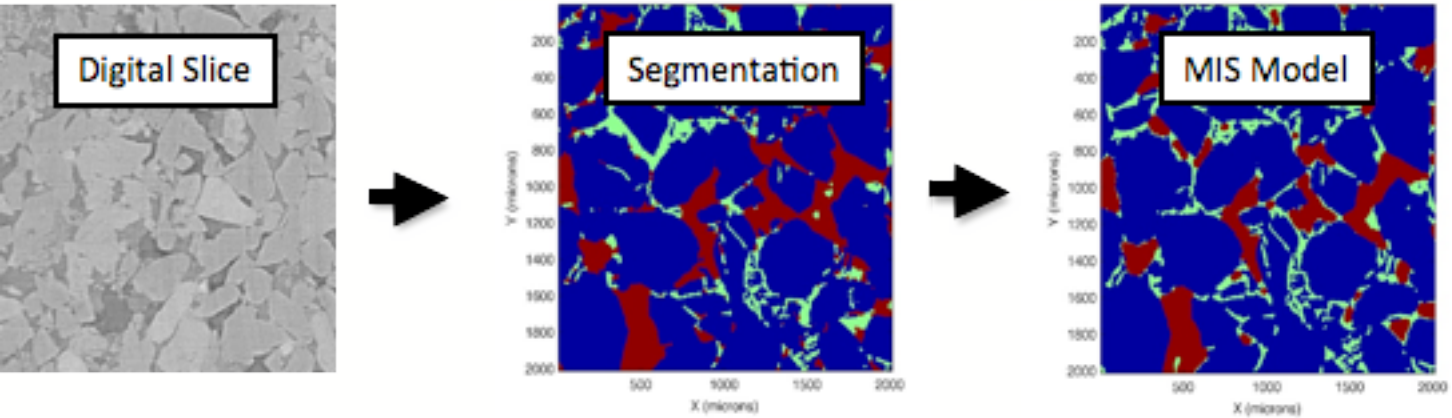
ALS Scientific Workflow envisioned



Real-time Image Analysis will have high-impact.

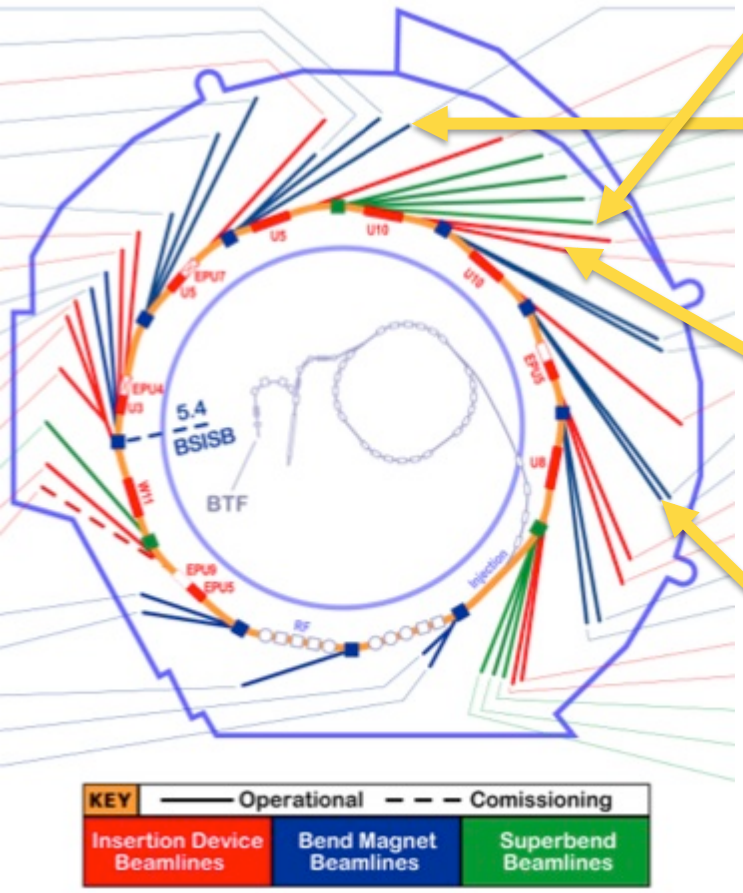


- **3-D beam-time feedback requires stopping data taking for > 1 hour for data processing.**
- **Many users analyze a small fraction.**



Geologic CO2 Sequestration
Jonathan Ajo-Franklin[1/23/13]

ALS Beamlines and Codes



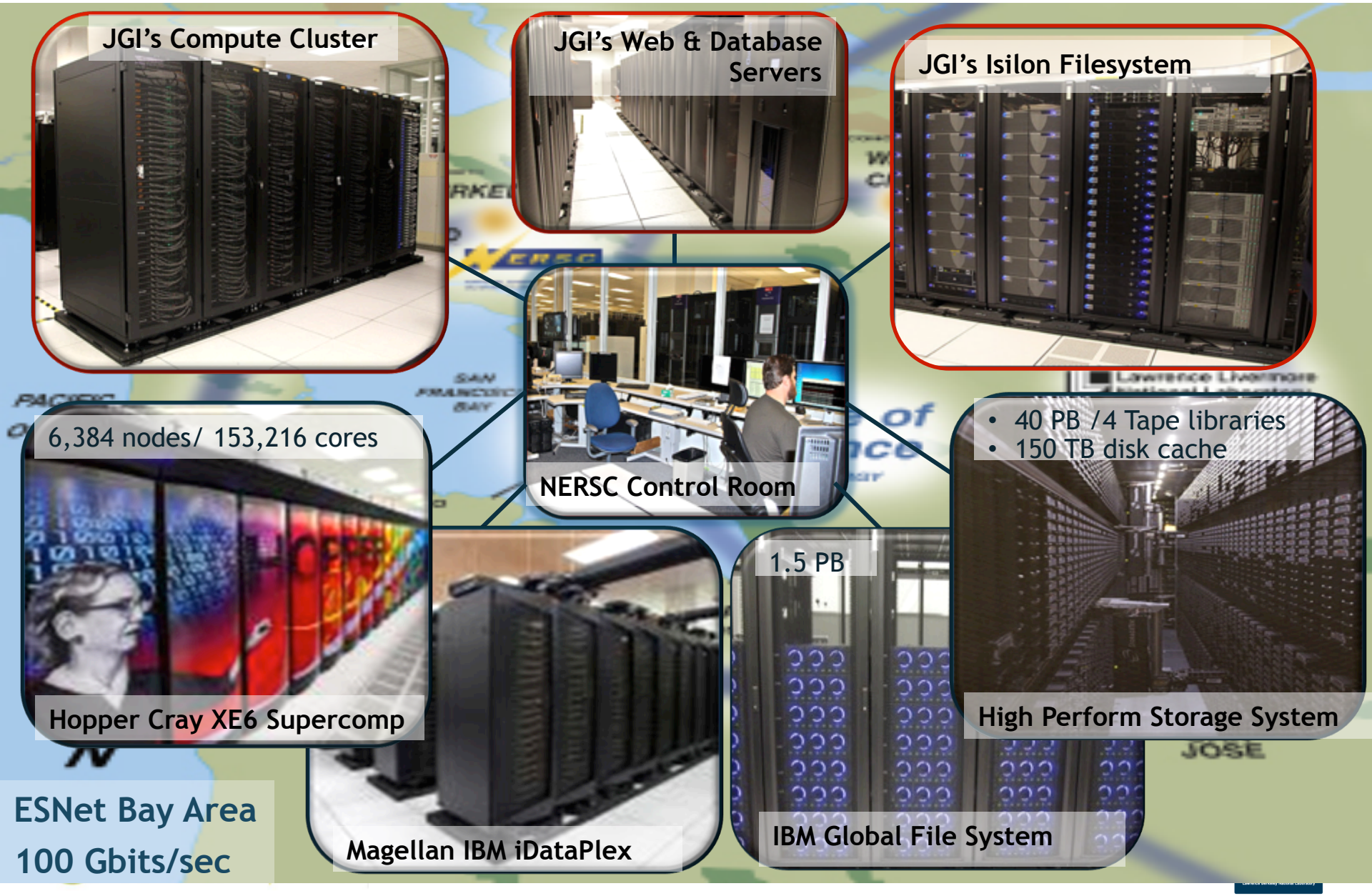
Real Space

Reciprocal Space

Spectroscopy

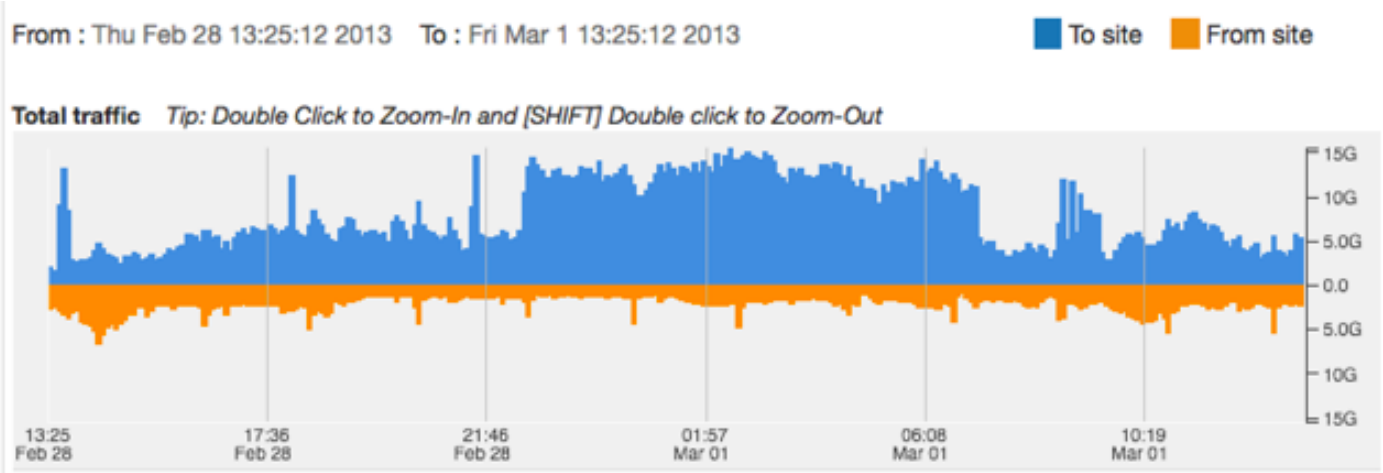
U.S. DEPARTMENT OF ENERGY | Office of Science

Facility Impact: Moving Data to Computing for Efficiency

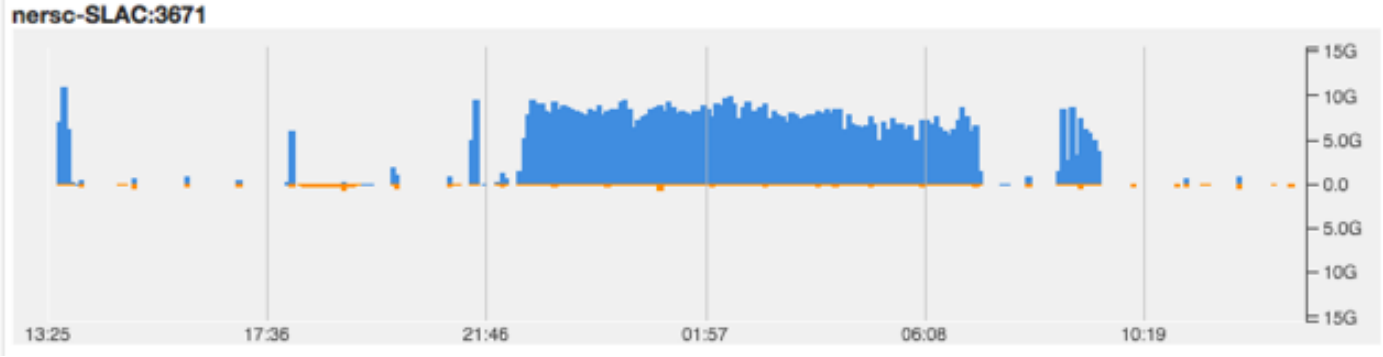


LCLS/NERSC… a model for grouping facilities?

All NERSC Traffic



Photosystem II X-Ray Study

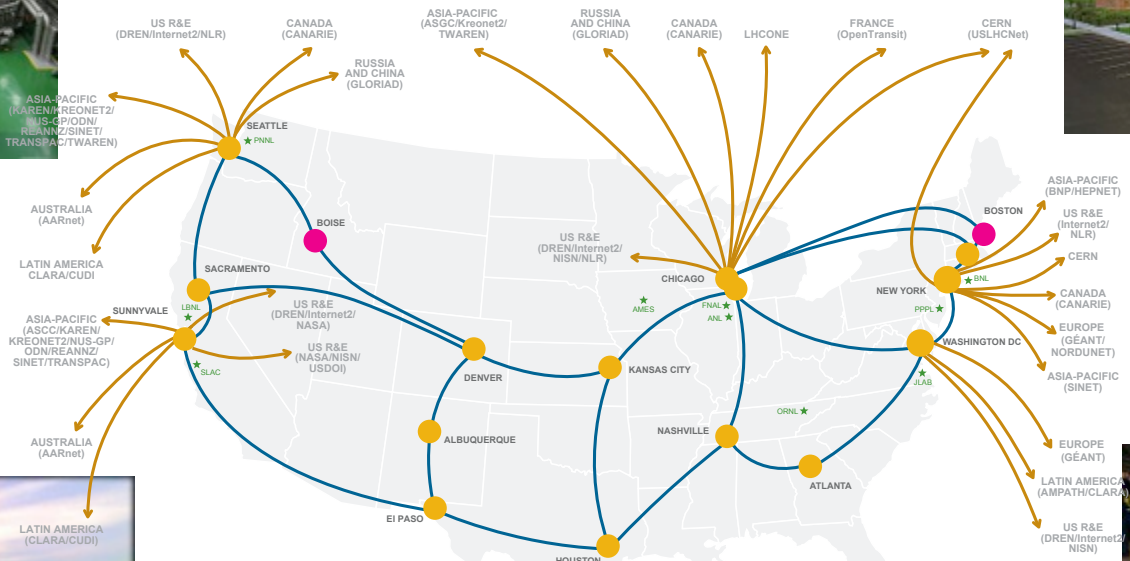


There may be an "opportunity" for NERSC to be the computing facility for LCLS

Unique data-centric resources will be needed

	Compute Intensive Arch	Data Intensive Arch
Compute	**Goal:** *Maximum Computational Density and local bandwidth for given power/cost constraint.* Maximizes bandwidth density near compute	**Goal:** *Maximum Data Capacity and global bandwidth for given power/cost constraint.* Bring more storage capacity near compute (or conversely embed more compute into the storage). *Requires software and programming environment support for such a paradigm shift*
On-Package DRAM		
Capacity Memory		
On-node-Storage		
In-Rack Storage		
Interconnect		
Global Shared Disk		
Off-System Network		Direct from each node

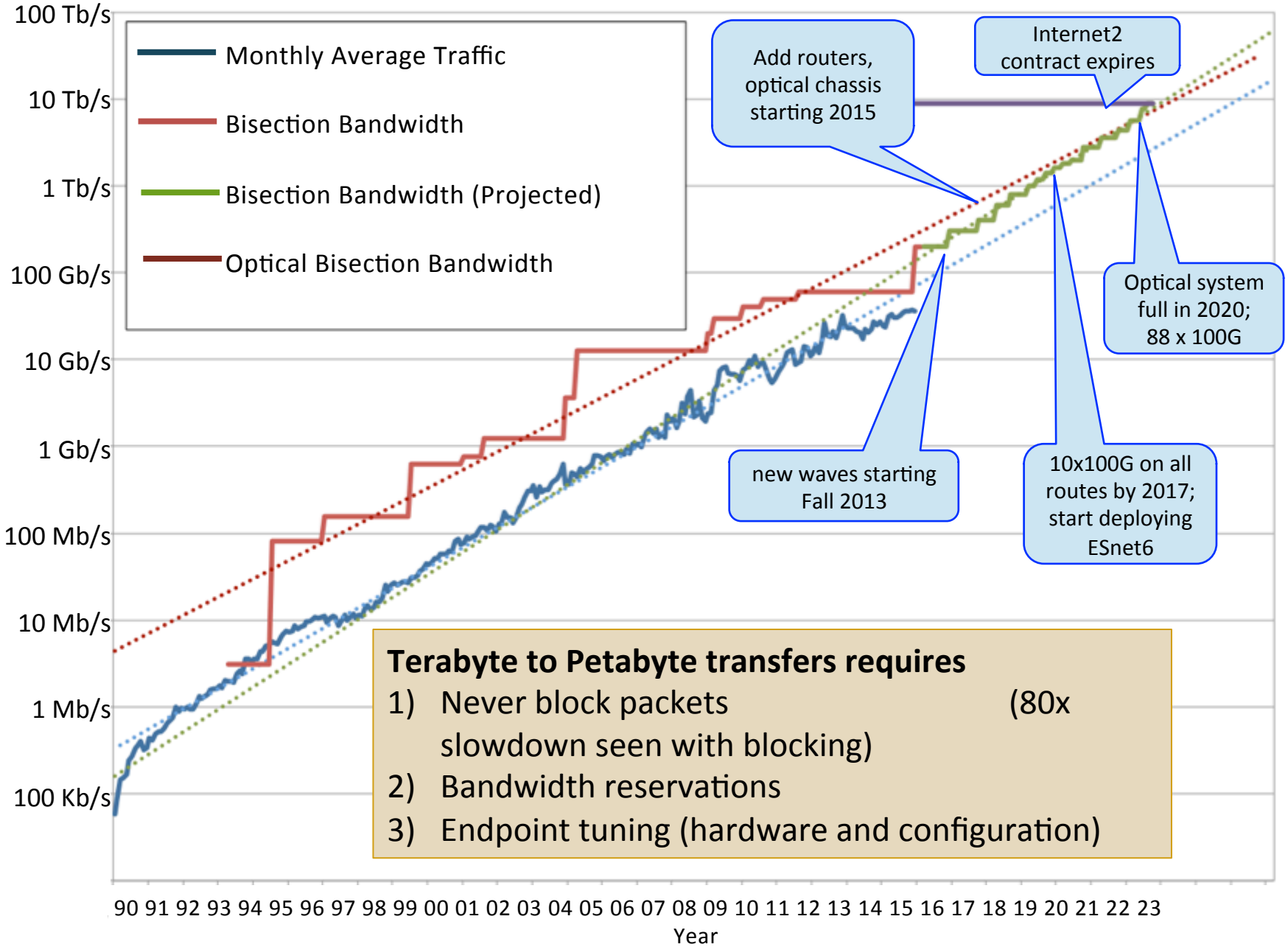
ESnet vision: discovery *unconstrained* by geography.



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

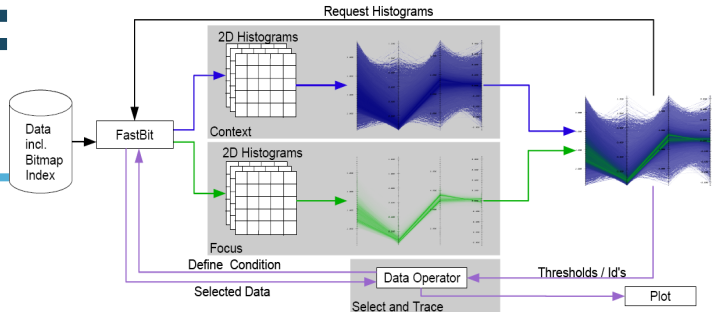
ESnet Needs to Continue Supporting Extreme Data Flows



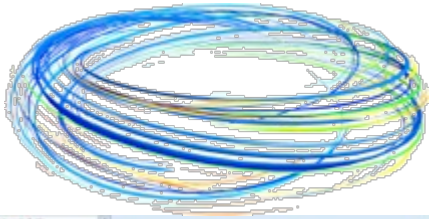
Terabyte to Petabyte transfers requires

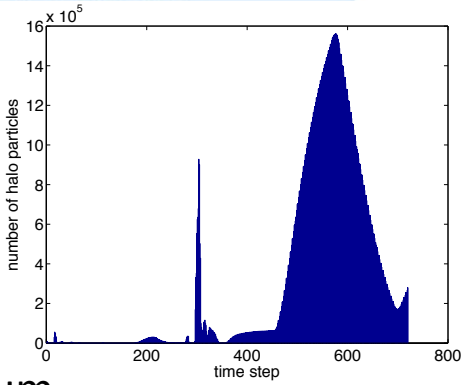
1) Never block packets (80x slowdown seen with blocking)
2) Bandwidth reservations
3) Endpoint tuning (hardware and configuration)

Computer Science Research Impact: Search/indexing Results



- **Fastbit**
  - specialized compression methods
  - Search speed by 10x – 100x than best known bitmap indexing methods
  - Theoretically optimal (order n)
- **Fastquery**
  - HDF5 and other structured files indexed and index embedded in file
  - Support indexing / query at object level
  - Queried 50TB dataset in ~10 seconds on 3,000 hopper cores





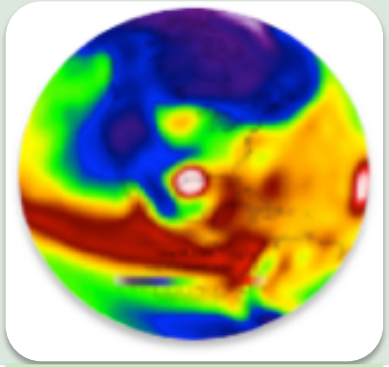
John Wu, Arie Shoshani, Alex Sim, Doron Rotum

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

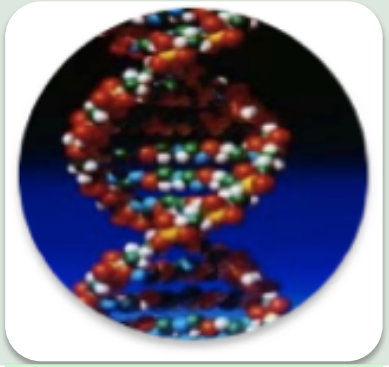
Programming Challenge? Science Problems Fit Across the "Irregularity" Spectrum

Massive Independent Jobs for Analysis and Simulations



Nearest Neighbor Simulations

All-to-All Simulations

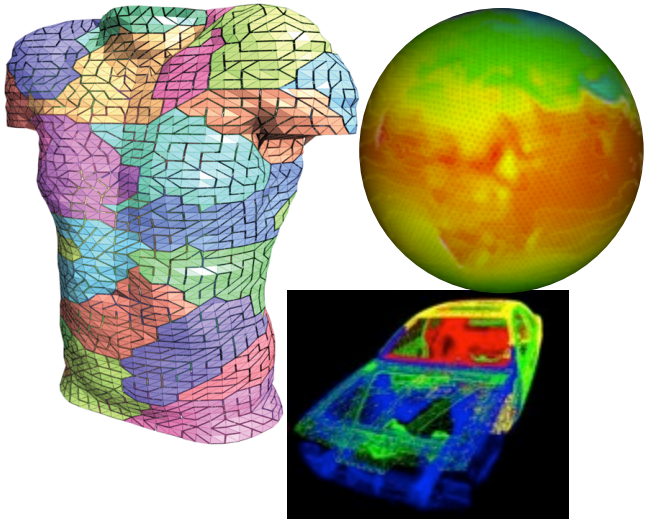


Random access, large data Analysis

… often they fit in multiple categories

The Programming Answer is Obvious…

More Regular



<u>Message Passing Programming</u>

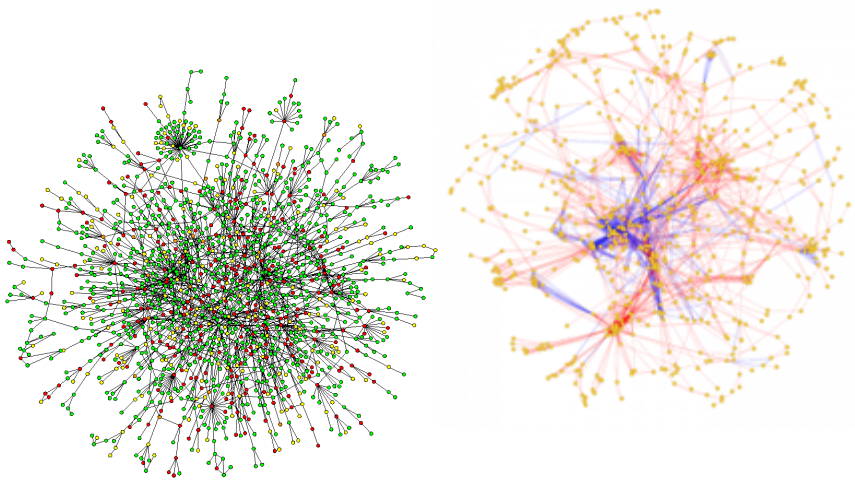
Divide up domain in pieces

Compute one piece

Send/Receive data from others

MPI, and many libraries

More Irregular



<u>Global Address Space Programming</u>

Each start computing

Grab whatever / whenever

UPC, CAF, X10, Chapel, GlobalArrays

Big Data needs advanced mathematics: statistics / machine learning / mining

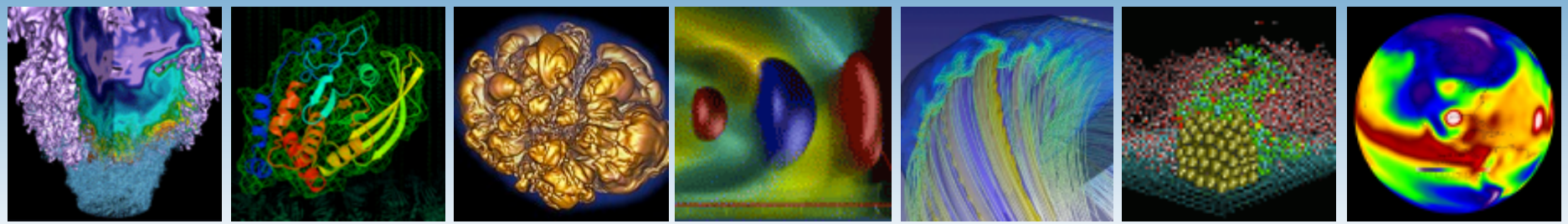
Querying	spherical range-search O(N), orthogonal range-search O(N), spatial join $O(N^2)$, nearest-neighbor O(N), all-nearest-neighbors $O(N^2)$
Density estimation	mixture of Gaussians, kernel density estimation $O(N^2)$, kernel conditional density estimation $O(N^3)$
Regression	linear regression, kernel regression $O(N^2)$, Gaussian process regression $O(N^3)$
Classification	decision tree, nearest-neighbor classifier $O(N^2)$, nonparametric Bayes classifier $O(N^2)$, support vector machine $O(N^3)$
Dimension reduction	principal component analysis, non-negative matrix factorization, kernel PCA $O(N^3)$, maximum variance unfolding $O(N^3)$
Outlier detection	by density estimation or dimension reduction
Clustering	by density estimation or dimension reduction, k-means, mean-shift segmentation $O(N^2)$, hierarchical clustering $O(N^3)$
Time series analysis	Kalman filter, hidden Markov model, trajectory tracking $O(N^n)$
Feature selection & causality	LASSO, L_1 SVM, Gaussian graphical models, discrete graphical models
Fusion and matching	sequence alignment, bipartite matching $O(N^3)$, n-point correlation 2-sample testing $O(N^n)$

Source: Alexander Gray

The Missing Middle: Software Engineering

- **2008 survey**
  - Most scientists are self-taught in programming
  - Only ⅓ think formal training in SW Eng is important
  - < ½ have a good understanding of SW testing
- **For example, bug in SW supplied by another research lab forced UCSD Scripps Prof to retract 5 papers**
  - *Science, Journal of Molecular Biology,* and *Proceedings of the National Academy of Sciences*

"Computational science: ...Error…why scientific programming does not compute," by Zeeya Merali, 13 October 2010, *Nature* 467, 775-777



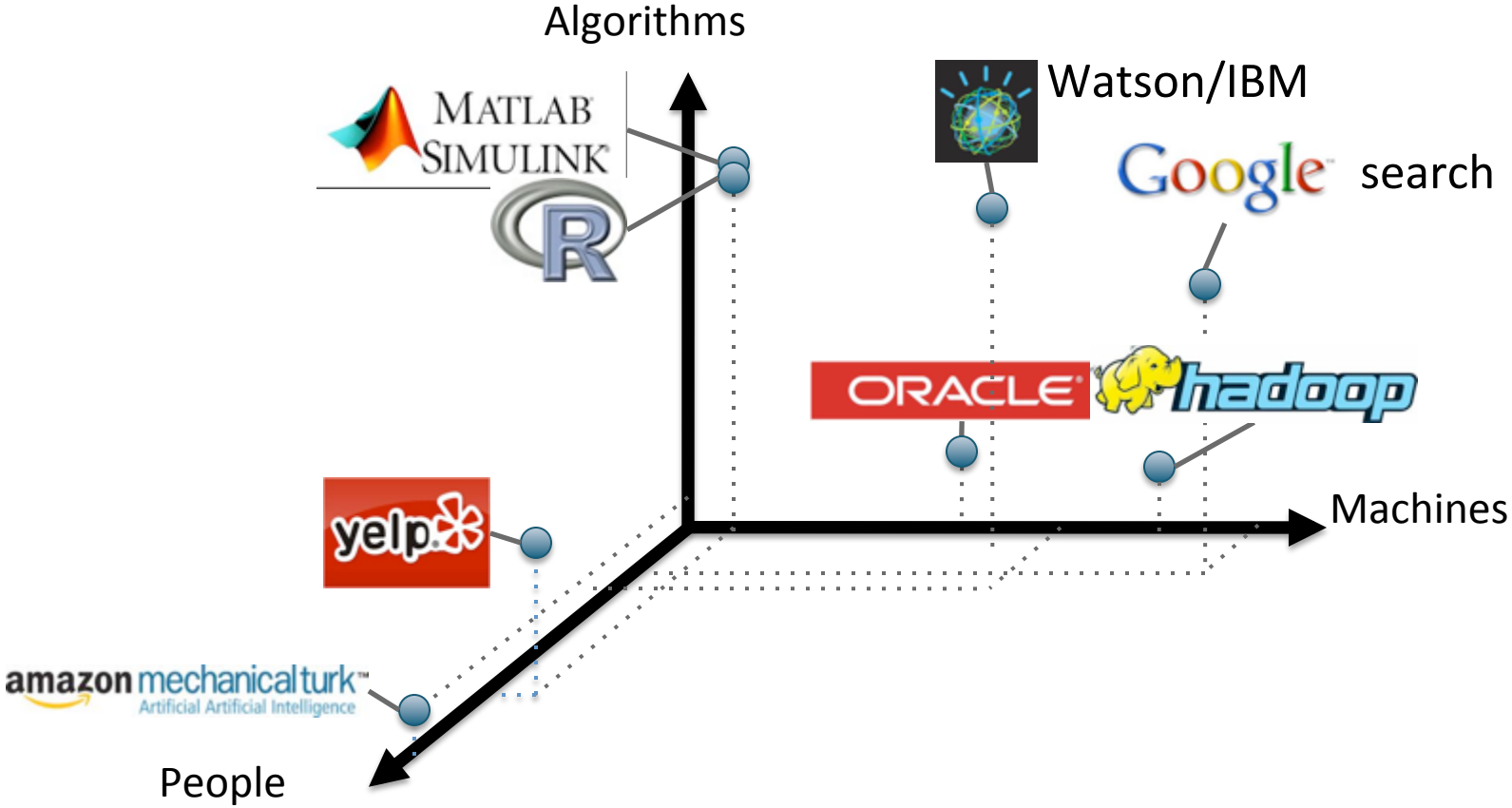
Non-Problems

Big Data problems that are not DOE's

UC Berkeley AMPLab: Algorithms, Machines, People

Franklin, Jordan, Patterson, Stoica

- **Today's apps: fixed point in solution space**



Need techniques to dynamically pick best operating point

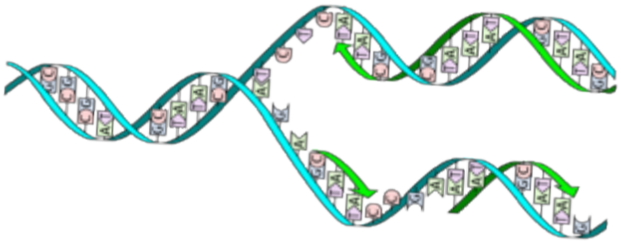
AMPLab Applications involve Privacy Issues

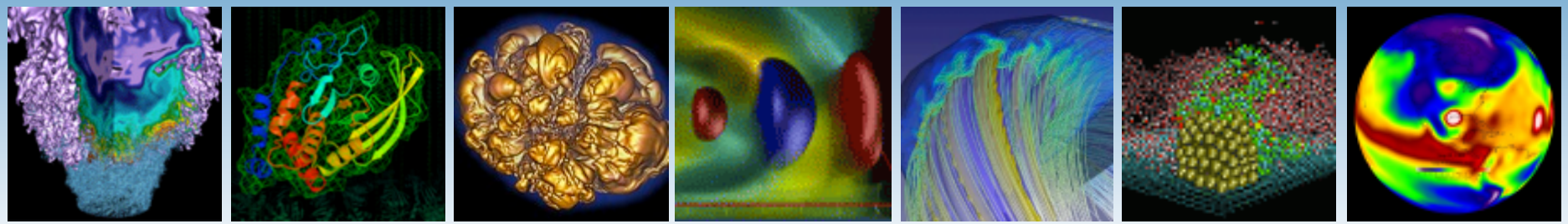
- **Mobile Millennium Project**
  - *Alex Bayen*, Civil and Environment Engineering, UC Berkeley

- **Microsimulation of urban development**
  - *Paul Waddell*, College of Environment Design, UC Berkeley

- **Crowd based opinion formation**
  - *Ken Goldberg*, Industrial Engineering and Operations Research, UC Berkeley

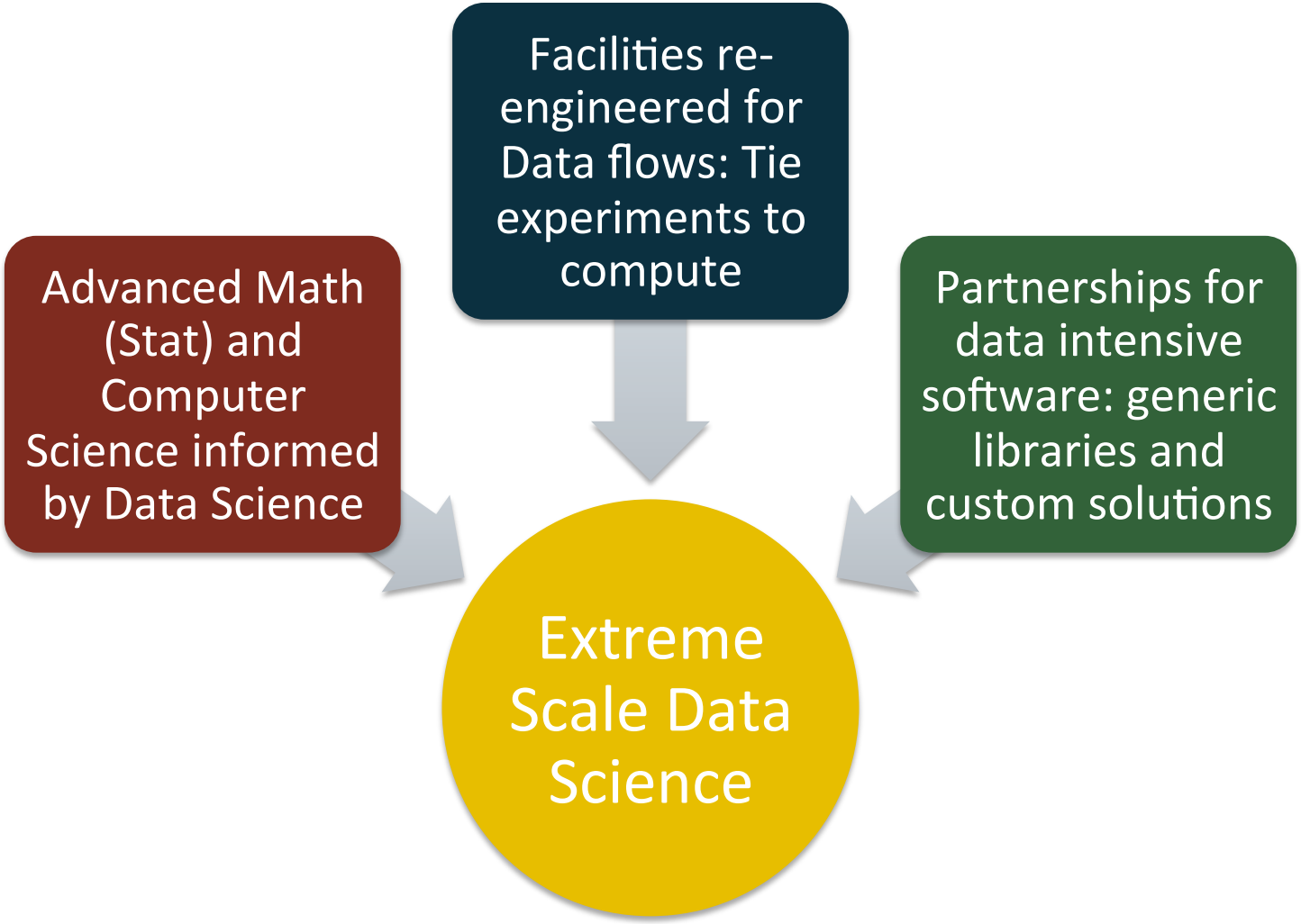
- **Personalized Sequencing**
  - *Taylor Sittler*, UCSF





Path Forward

Model of Success



A Data Strategy

- **Facilities: Drive and deploy the best technology for data science in collaboration with industry**
  - **Networking:** for bandwidth, flexibility, and to federate facilities
  - **Systems:** for data storage, analysis
- **Research: Perform basic research in data science**
  - **Math and Algorithms:** Statistics & Machine Learning, Image analysis, Graph analytics in collaboration with universities
  - **Computing Systems**: Develop and evaluate new hardware, programming, and software techniques
- **Software**
  - **Partnerships:** Leverage Lab staff and culture to develop usable robust tools for data movement, provenance, analysis..
  - **Cross-lab organization:** to encourage re-use and sharing