

A Superfacility for Data Intensive Science

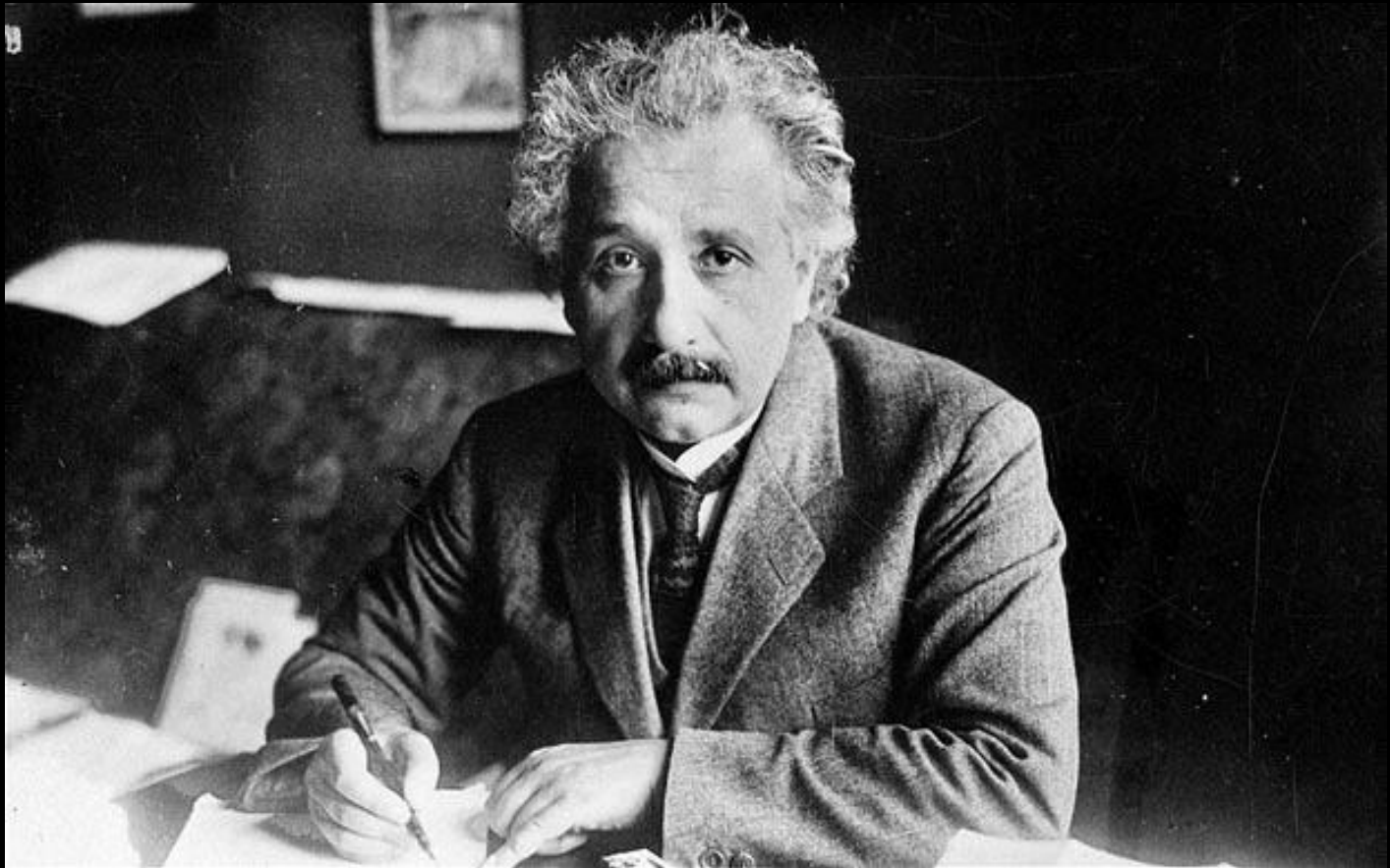
Kathy Yelick

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

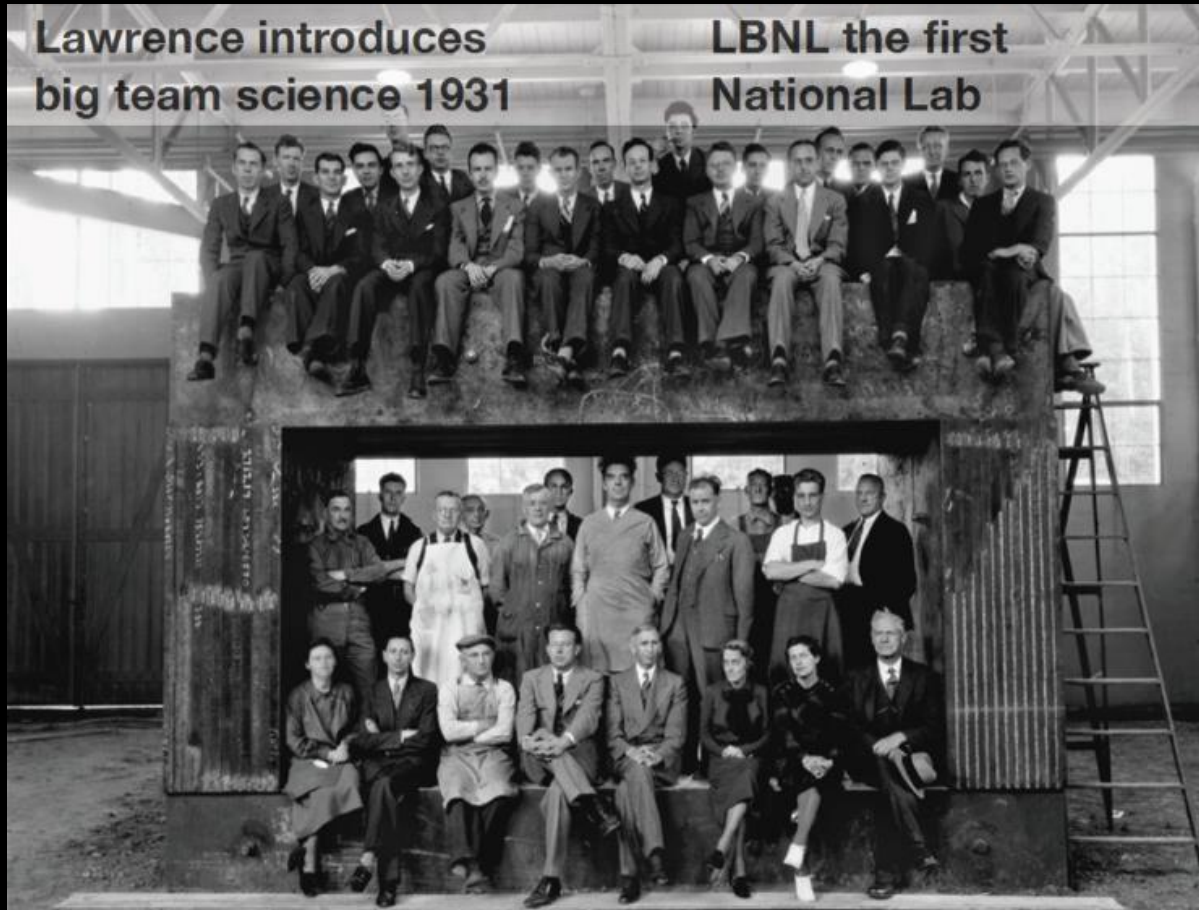
Part 1

Science is poised for transformatio

Old School Scientists: The Lone Scientist



Team Science



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

Experimental Science is Changing

By using our website you agree to our use of cookies in accordance with our cookie policy.

OK

PRIVACY POLICY +

JAX[®] MICE & SERVICES



JAX[®] Mice are the highest quality and most-published mouse models in the world. Take advantage of our large inventories of common inbred strains and the convenience of having your breeding and drug efficacy needs met by the leading experts in mouse modeling.

Search for Mice

Advanced Mice Search

Search for mice by strain, stock, gene, allele and synonyms

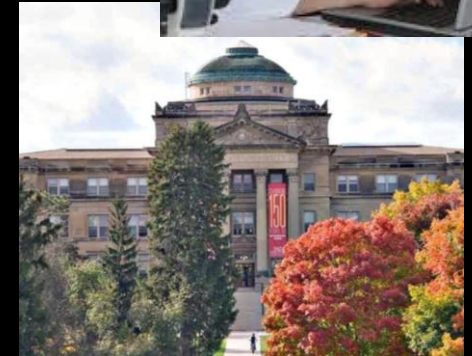
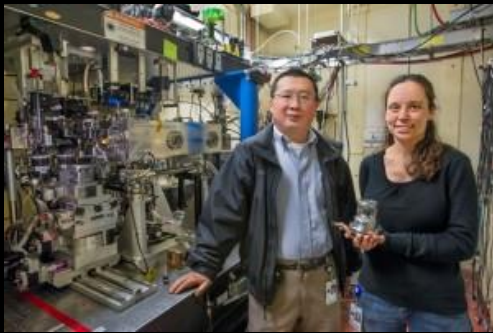
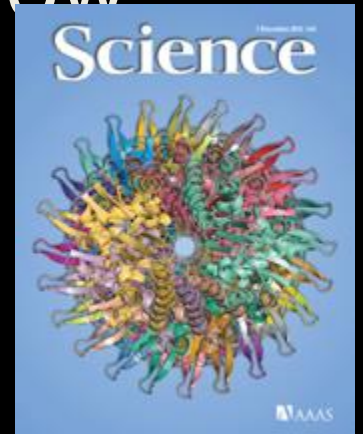
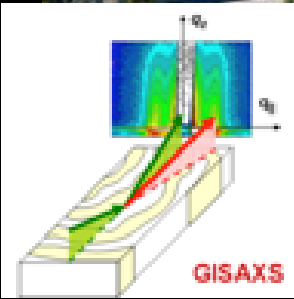


Breed Your Mouse

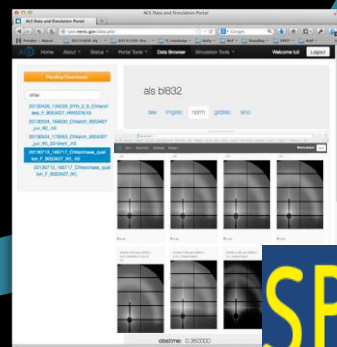
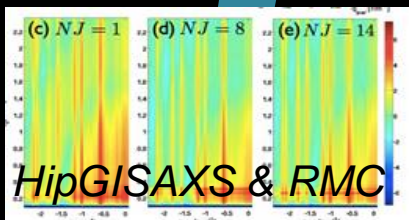
Test Your Drug

Cryopreserve Your Mouse

Old School Scientific Workflow

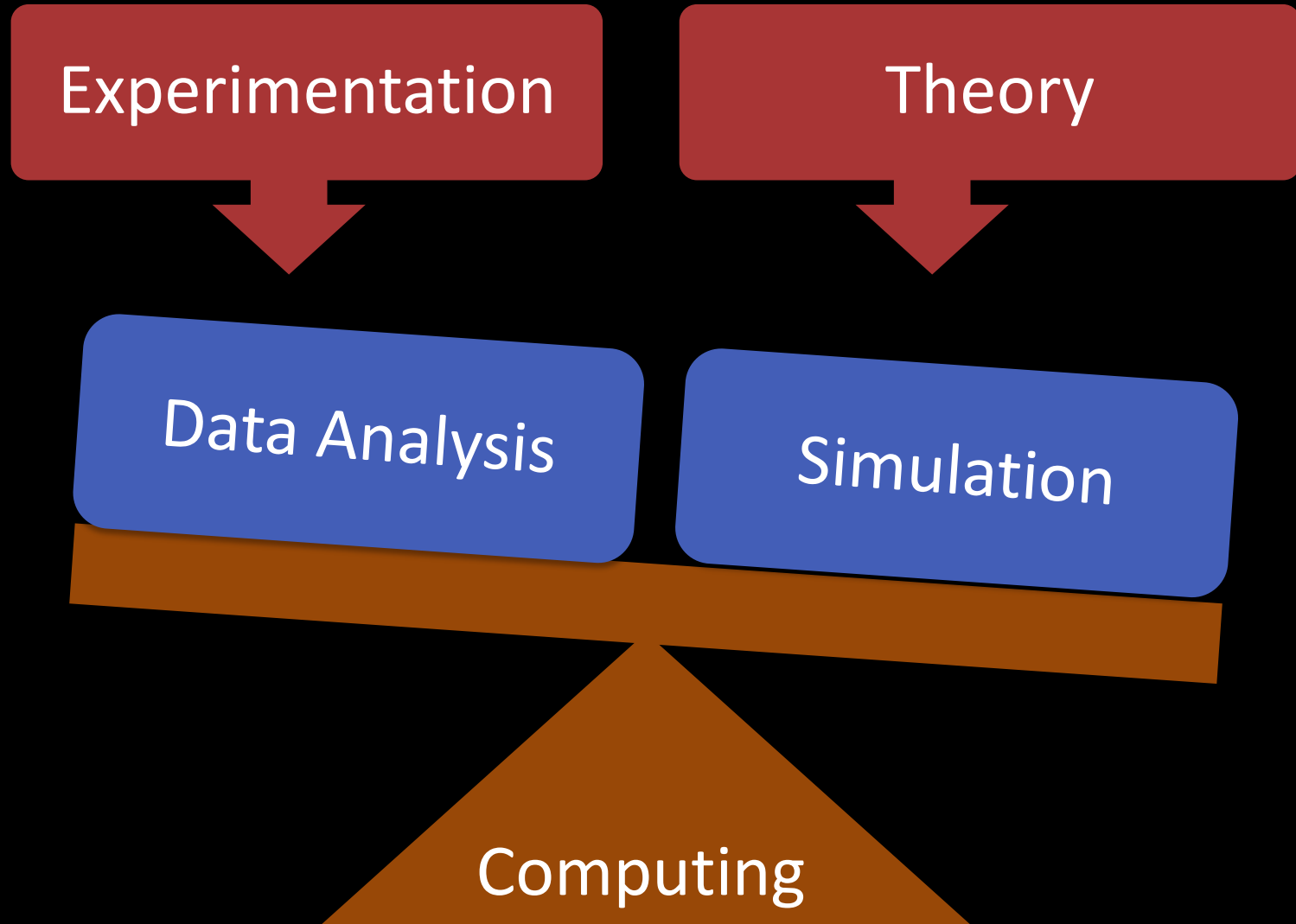


Computing, experiments, networking and expertise in a “Superfacility” for Science

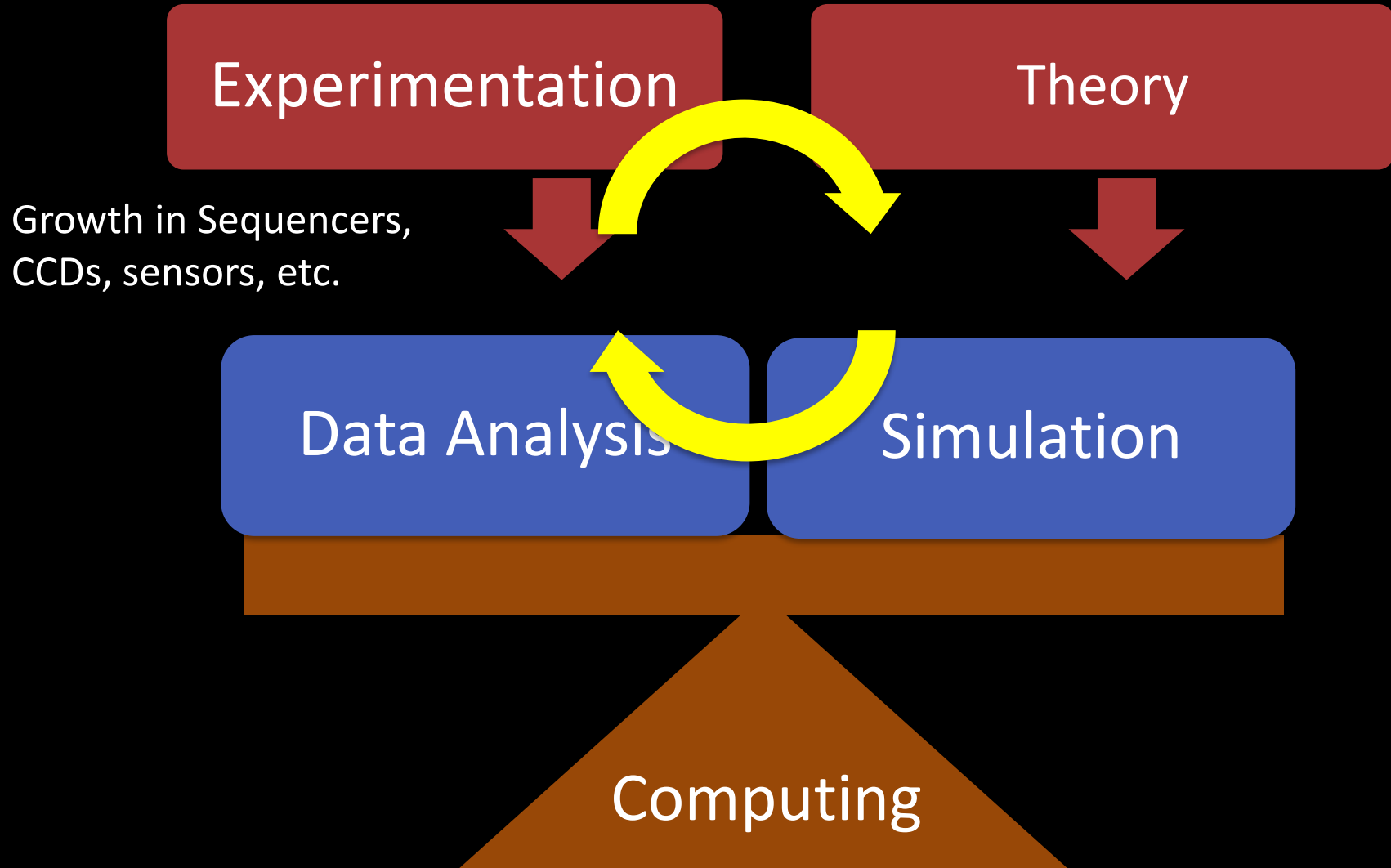


Liu et al, “Fast printing and in situ morphology ...”. Adv Mater. 2015

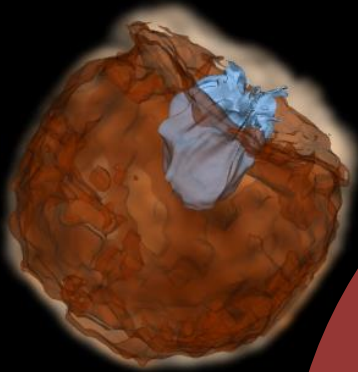
Old School HPC: only for Simulation



HPC is equally important in experimentation

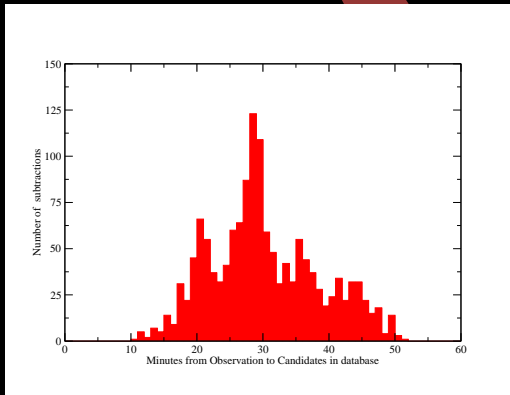


Integration of Simulation and Observational Science



Intermediate Palomar Transient Factory

- Nightly images transferred
- Subtractions, machine learning
- Candidates in database in < 5 minutes
- Simulations aid in interpreting data



Yi Cao, *et al.* (2015) *Nature*,
“A strong ultraviolet pulse from
a newborn Type Ia supernova”

Old School Scientific Data Search

Safari File Edit View History Bookmarks Window Help

www.google.com/search?tbs=sbi:AMhZZIu-Ft1o4xXIjhVjcLUV_1GtY_1M9gV_1hy

Berkeley Lab (...) TeamSnap :: M... Google CalMail - You... Search Results...

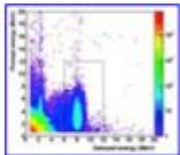
+You Search Images Mail Drive Calendar Sites Groups More -

CalMail - You must be logged in to a page.

Google Antineutrinos.jpg

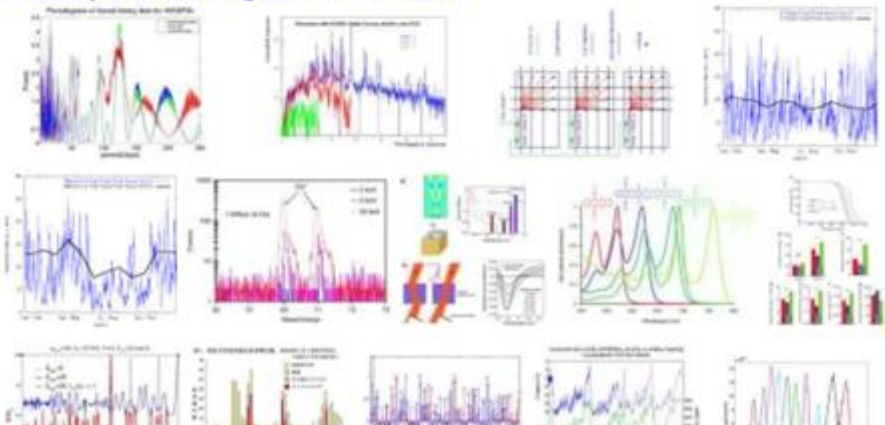
Web **Images** Maps Shopping More Search tools

Tip: Try entering a descriptive word in the search box.

 Image size:
153 × 133

No other sizes of this image found.

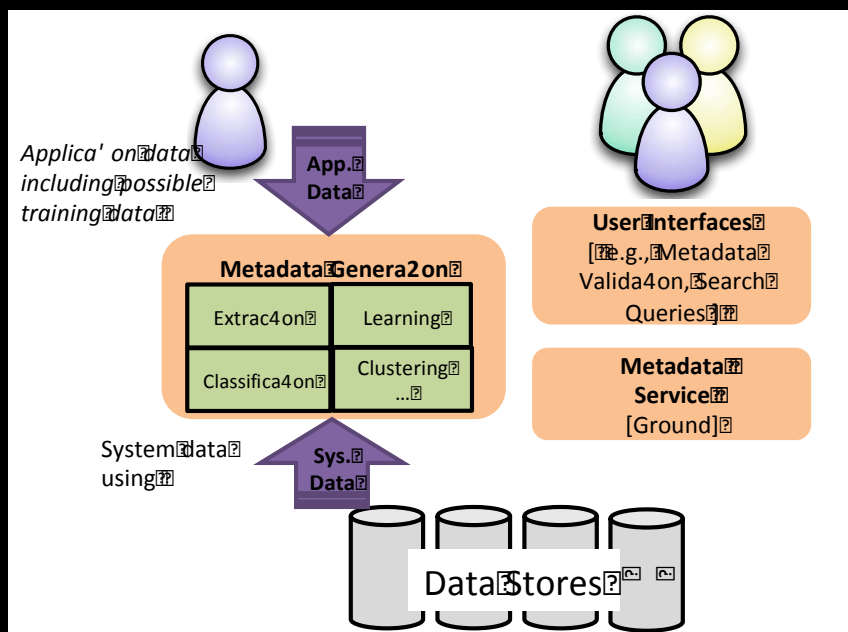
[Visually similar images](#) - Report images



Automated Search, Meta-Data Analysis, and On-Demand Simulation

The screenshot shows the Materials Project website. At the top, there are navigation links: Home, Apps, Support, About, References, and Login or Register. The main header features the Materials Project logo and the tagline "A Materials Genome Approach". Below this, a "Find Materials" section includes a "Quick Search" box with a search bar and a "Materials Explorer" section. A blue banner states: "This web site is an early release, currently containing 15433 compounds. We are continuously improving our software and database." There are two columns of options: "Register now for free, full access" (with bullet points: Unlimited access, Up to 500 search results, History of your searches and analyses) and "Or try the apps in demo mode" (with bullet points: 10 minute usage limit, Search results limited to 10 best matches, Just click an app to start). Below these are four app tiles: "Phase Diagram App", "LI-Ion Battery Explorer", "Reaction Calculator", and "Structure Predictor". At the bottom, there are "Press Highlights" and "Latest News" sections.

Automated metadata extraction using machine learning

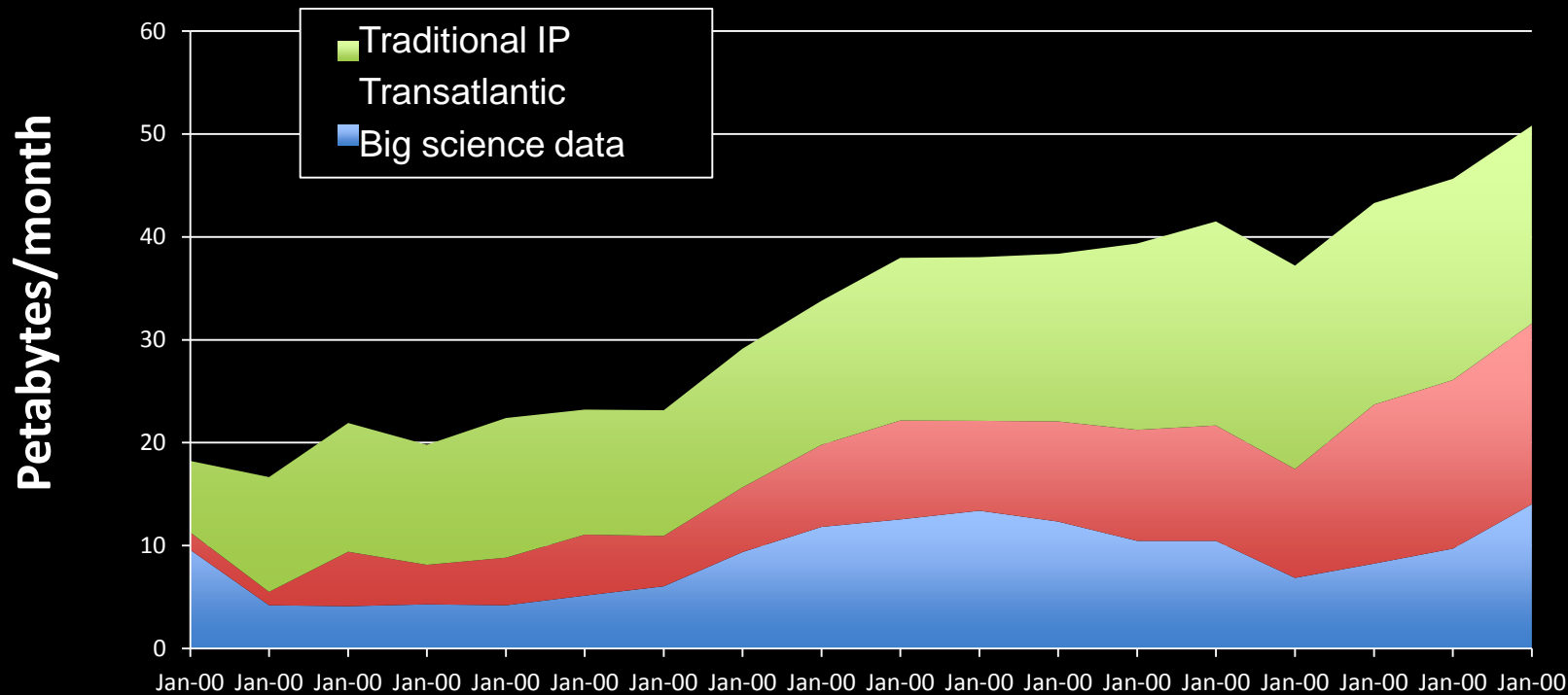


Jobs submitted by "bots" based on queries; algorithms extract informatics for design

Part 2

ASCR Facilities need to adapt

ESnet: Exponential data growth drives capacity



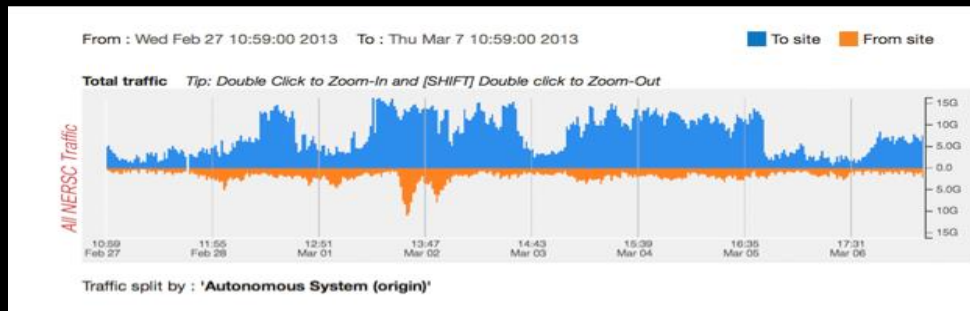
Science DMZ to deliver bandwidth to the end users

OSCARS for bandwidth reservation



100 Exabytes/year by 2024!

ESnet: Discovery Unconstrained by Geography



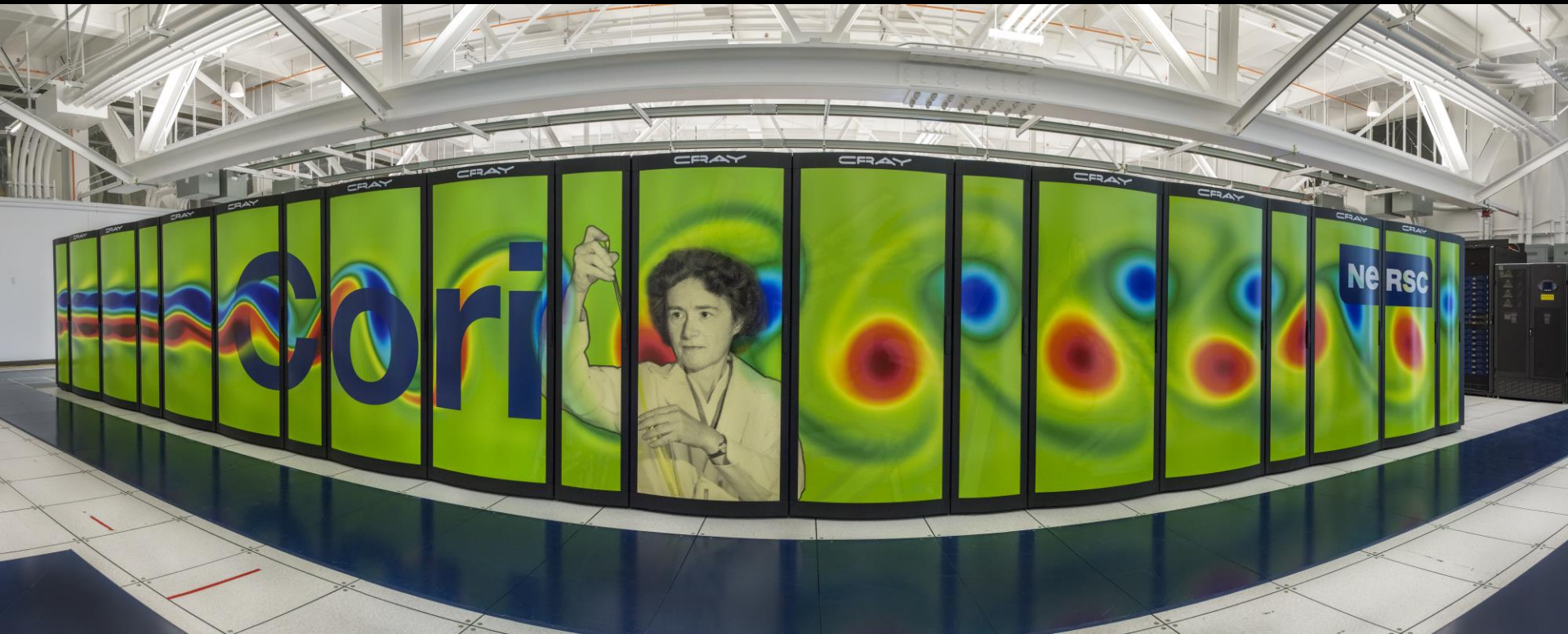
LCLS/NERSC/Esnet Superfacility demo for Photosystem II

→ 3x network traffic

ESnet-6 Upgrade Options trade off risk and capability

- | | | |
|---|--|--|
| <p>Software Defined Network</p> <ul style="list-style-type: none"> ▪ Programmable switches may improve cost and speed ▪ Adapt lower level network layers for major science flows | <p>Packet Optical</p> <ul style="list-style-type: none"> ▪ Combine hardware for packetization/routing with optical transport ▪ Lower cost | <p>Current Architecture</p> <ul style="list-style-type: none"> ▪ Keep optical separate with current fixed routing tables ▪ Known technology |
|---|--|--|
- Network performance enables efficiency of centralized computing**

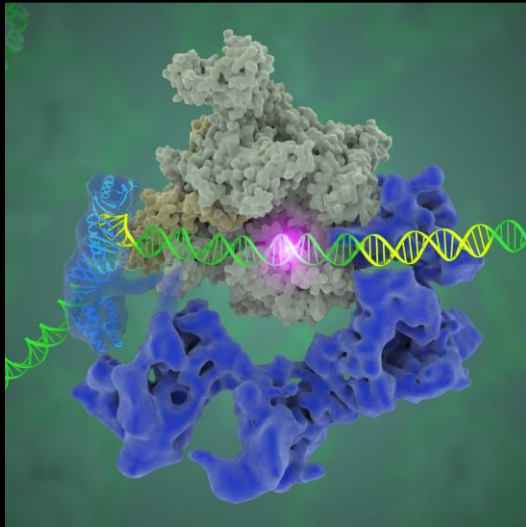
Systems configured for data-intensive science



NERSC Cori has data partition (Phase 1, Haswell) pre-exascale (Phase 2, KNL preproduction)
WAN-to-Cori optimized for streaming data: 100x faster from LCLS to Cori and Globus to CERN

Real-time queue prototyped at NERSC

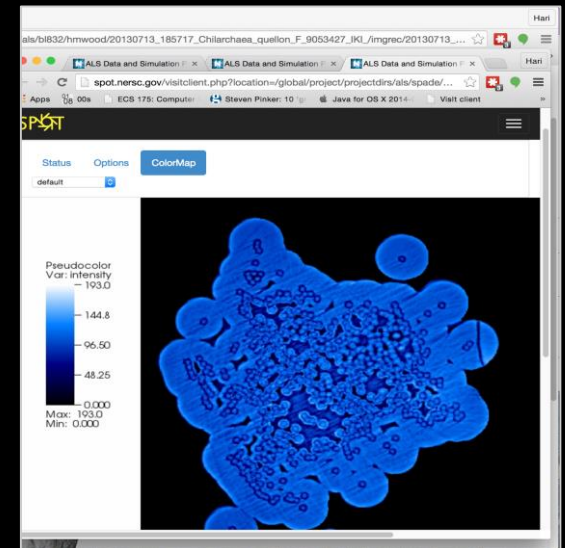
- In 1998 dedicated hardware; now prototype queue on Cori
- <1% of NERSC allocation
- Cryo-Em, Mass spec, Telescopes, Accelerator, Light sources



Cryo-EM: Image classification
Nogales Lab



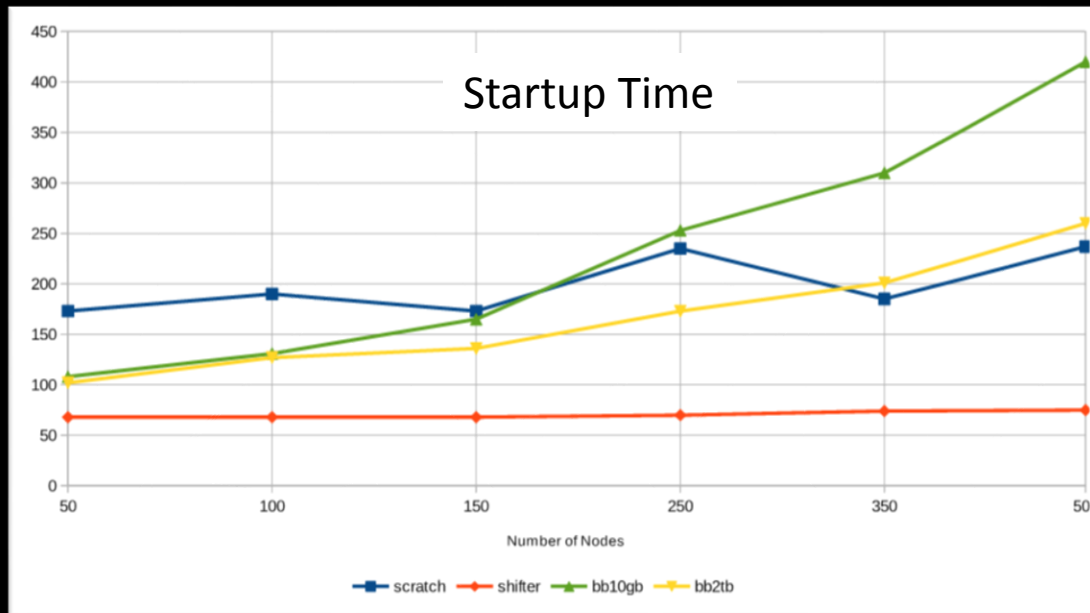
PTF: Image subtraction pipeline



ALS: 3D Reconstruction,
rendered on SPOT web portal

Containers for HPC Systems

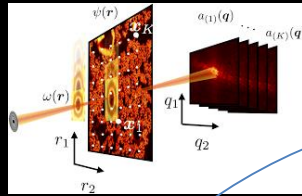
- Data analysis pipelines are often large, complex software stacks
- NERSC Shifter (with Cray), supports containers for HPC systems
- Used in HEP and NP projects
(ATLAS, ALICE, STAR, LSST, DESI)



Part 3

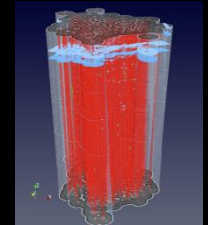
ASCR Research challenges
are substantial

CAMERA: Math for the Facilities

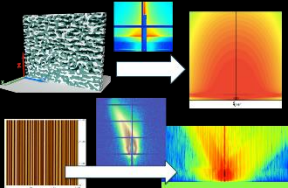


Designing mathematical algorithms to allow real-time analysis next to the equipment

Real-time streaming ptychography—ALS, delivered to NLSLS2, LANL, BESSY,



Automatic image processing for the ALS/GE



SFM/TEM + GISAXS

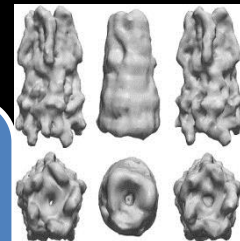
Multi-modal: Building the math that fuses information from multiple experiments

New algorithms to transform manual into automatic analysis



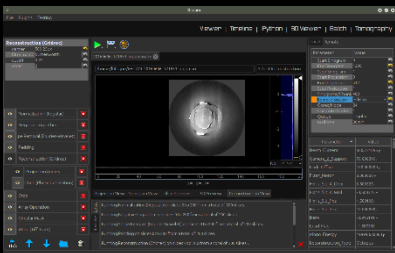
Compare and integrate multiple analysis tools

Inventing new math and models to match new acquisition technologies



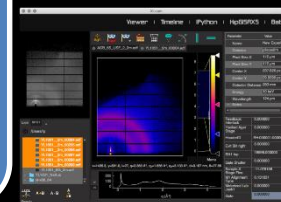
Fluctuation scattering and single particle imaging for the LCLS

CAMERA workshop on Tomography: Joint with APS, ESRF, DIAMOND, LNLS, LLNL, SSRL,....,



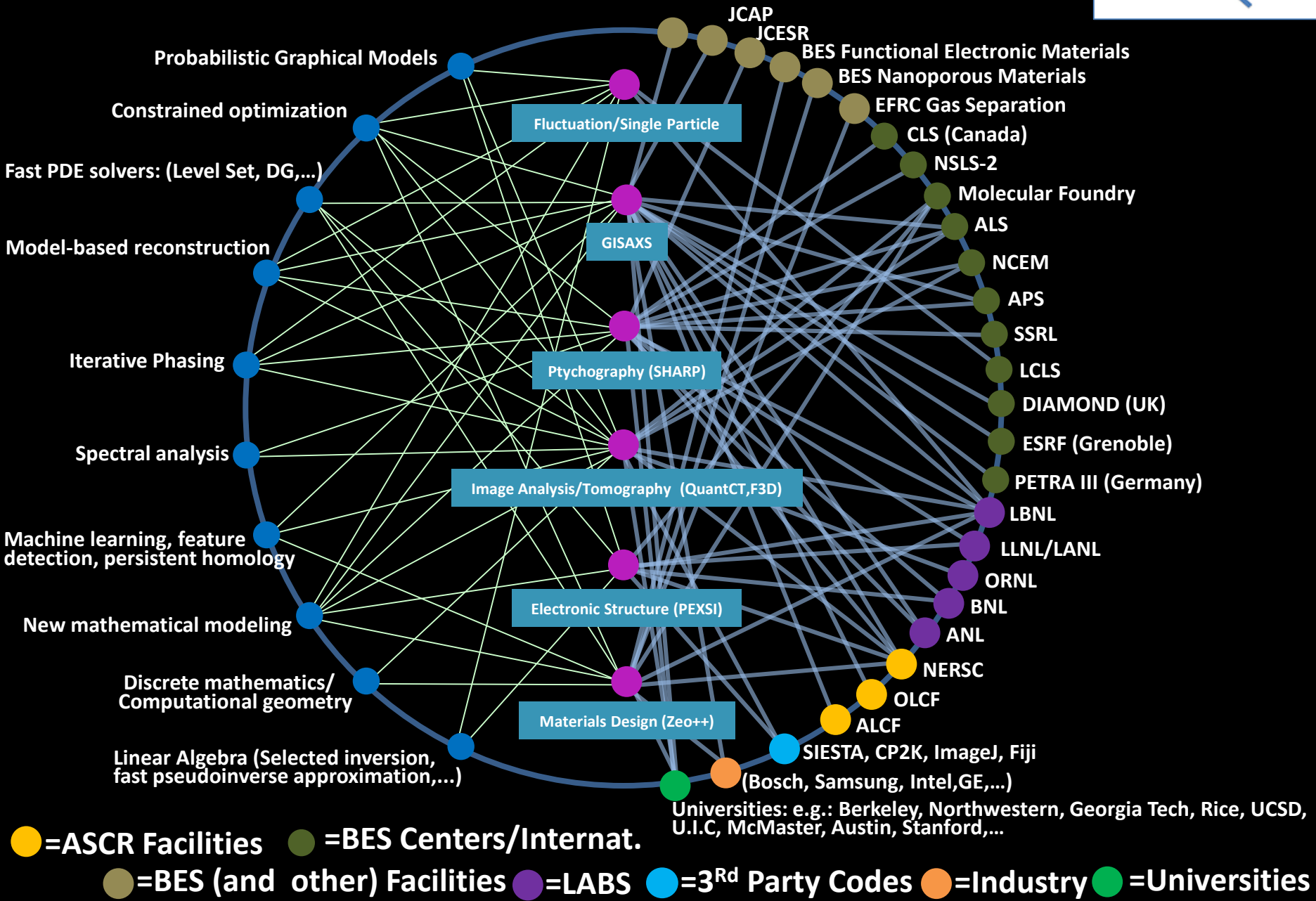
Cultural and Sociological Challenges

Robust and reliable codes and data flow: workflow environments



Workflow and access to remote supercomputers: XiCAM for ALS, SSRL, APS, NLSLS2

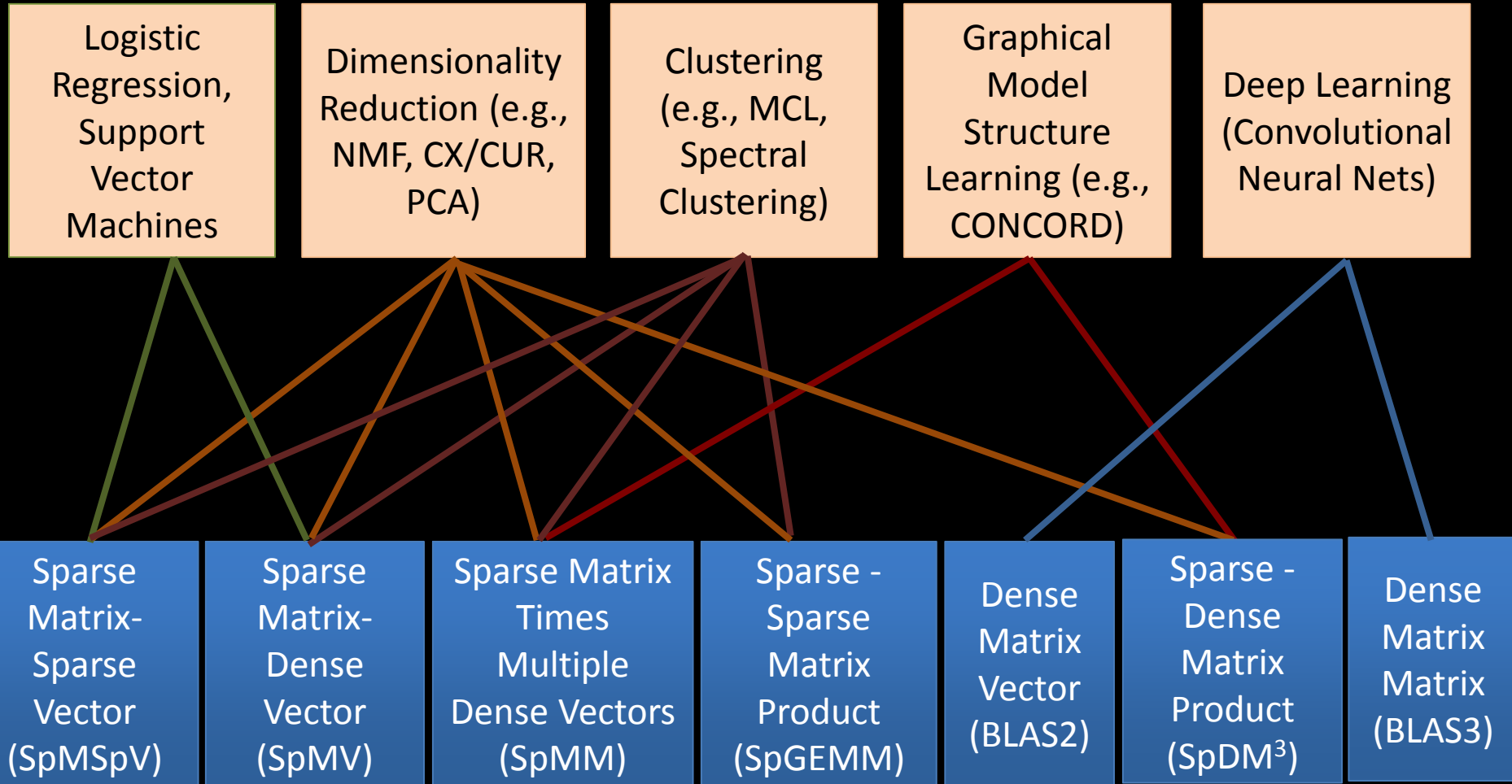
CAMERA: Making the connections



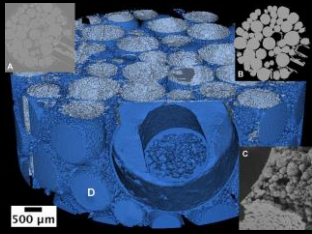
Analytics vs. Simulation Kernels:

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	
Integrations	Spectral methods
Alignment	Structured Meshes

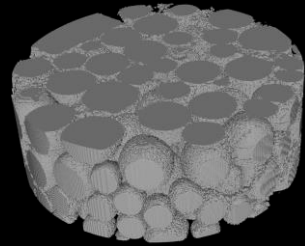
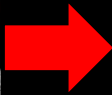
Machine Learning Mapping to Linear Algebra



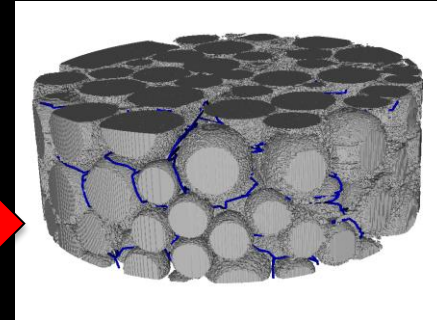
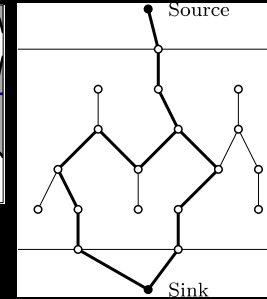
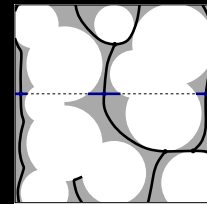
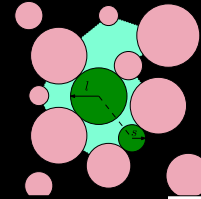
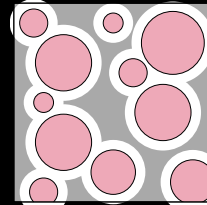
Software implementations at scale in pipeline



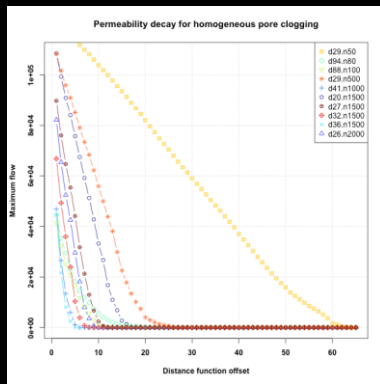
MicroCT
imaging



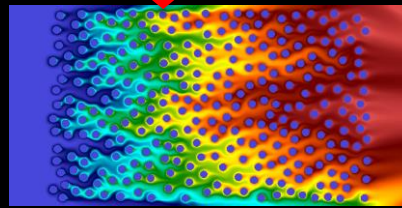
Segmentation



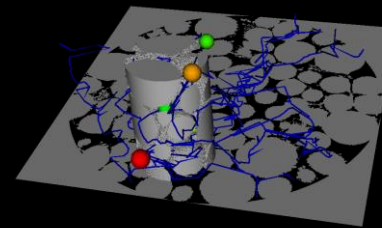
Topological
Analysis



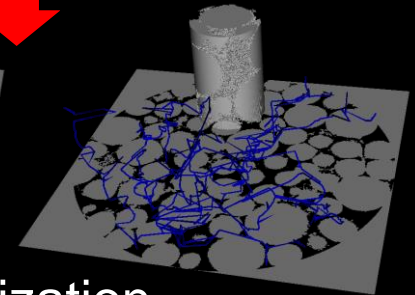
Analysis



Simulation



Visualization



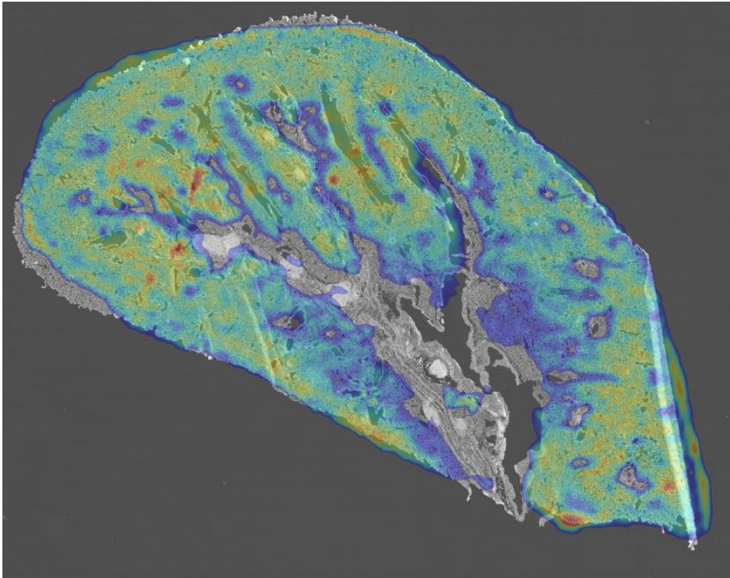
Interactive Analytics using Jupyter

```
In [10]: # overlaying the small H&E and MS images

registered_ms_image = ird.transform_img_dict(my_images[2], result)
big_registered_ms_image = imresize(registered_ms_image, optical_image.shape, interp='bicubic')

# cut out low intensity region of MS image for easy viewing of underlying H&E
masked_big_ms_image = np.ma.masked_where(big_registered_ms_image < 100, big_registered_ms_image)

# plot the two images overlaid
f = plt.figure(1, figsize=(20, 20))
plt.imshow(optical_monochrome, alpha=0.7, cmap=cm.Greys_r)
plt.imshow(masked_big_ms_image, alpha=0.3, cmap=cm.jet)
plt.axes().set_axis_off()
```



Science notebooks through Jupyter (iPython)

- Widely used in science
- Interactive HPC LDRD

Deployed at NERSC:

- >100 users pre-production

Random Access Analytics

- Genome assembly “needs shared memory”

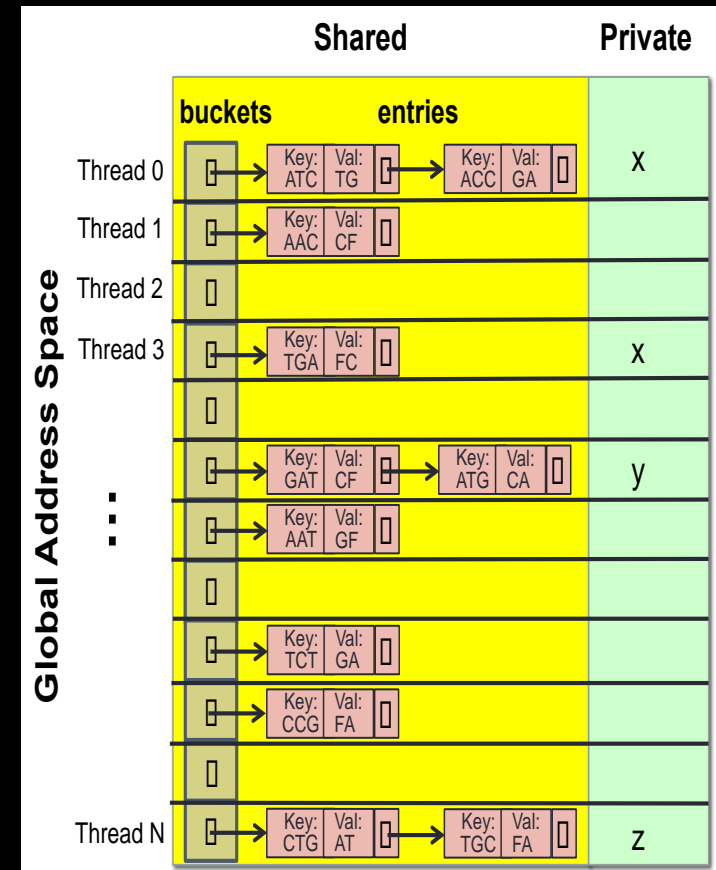
Global Address Space

- Low overhead communication
- Remote atomics
- Partitions for any structure

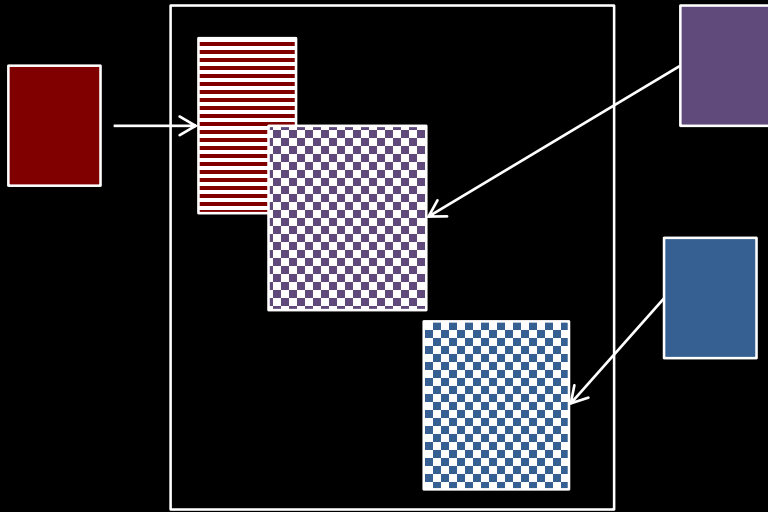
Scales to 15K+ cores

Under 10 minutes for human

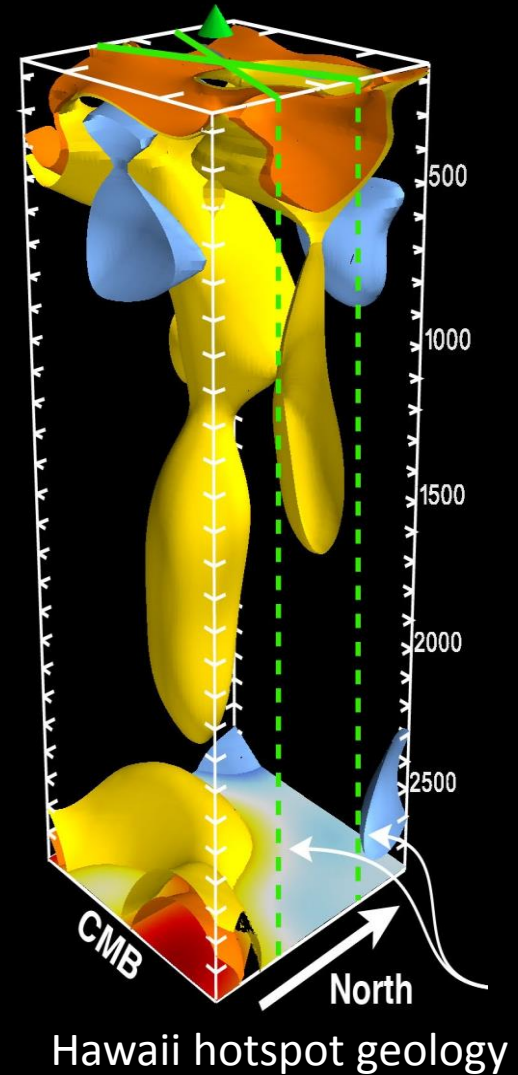
First ever solution



Data Fusion for Observation with Simulation



- **Unaligned data from observation**
- **One-sided strided updates**



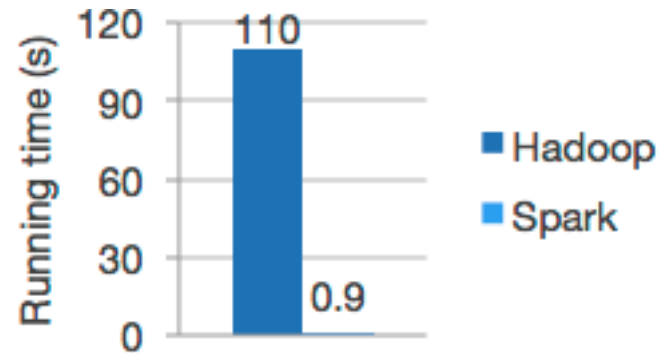
Scott French, Y. Zheng, B. Romanowicz, K. Yelick

Productive Programming



Speed

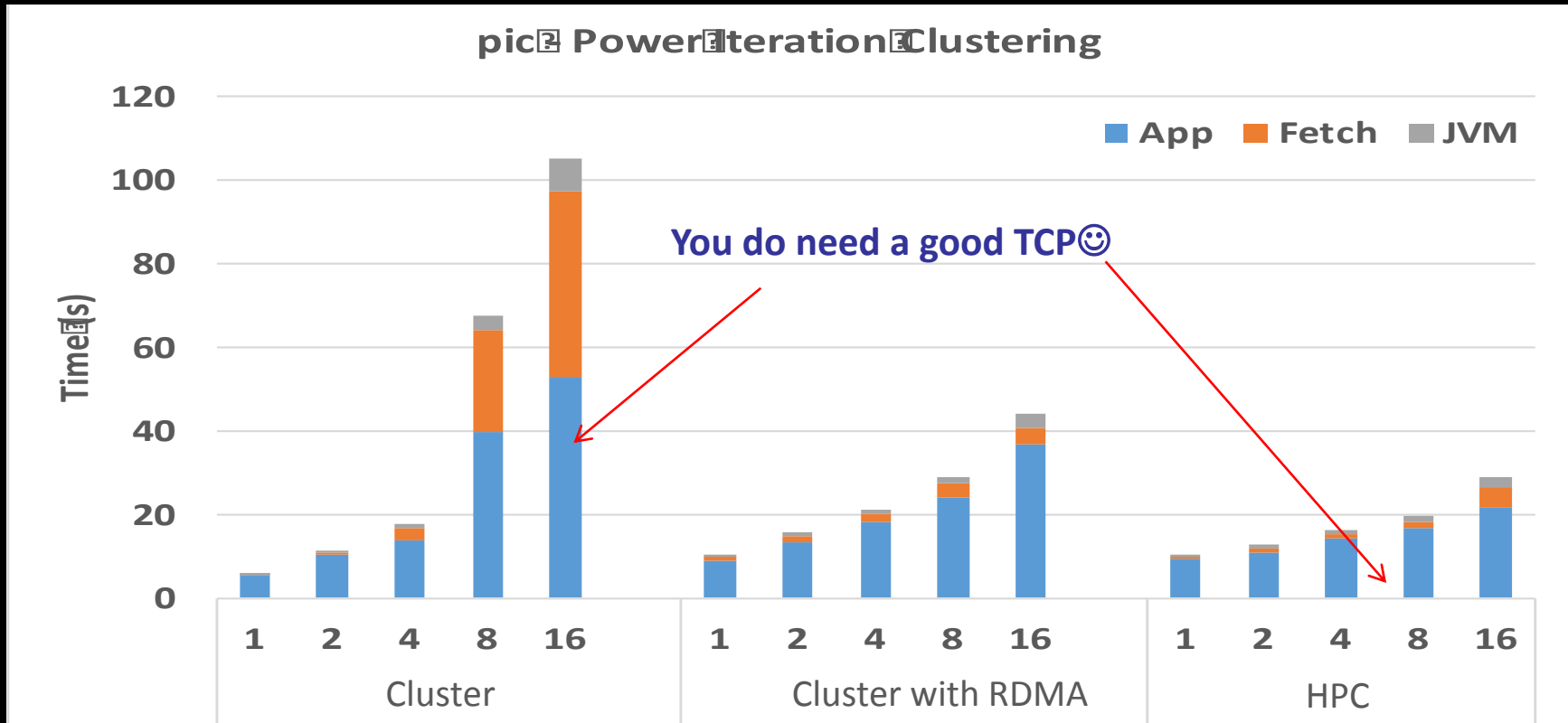
Run programs up to 100x faster than Hadoop MapReduce in memory, or 10x faster on disk.



- High failure rate
- Slow network
- Fast (local) disk

And Spark is still 10x+ slower than MPI

SPARK Analytics on HPC



SPARK on HPC vs. clusters

Network, I/O, and virtualization all key to performance

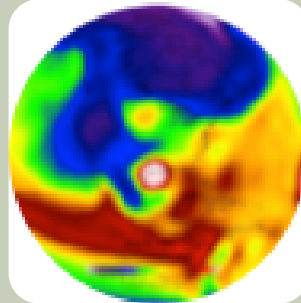
Architectures for Data vs. Simulation



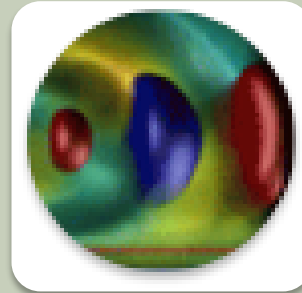
**Massive
Independent Jobs
for Analysis
and
Simulation**



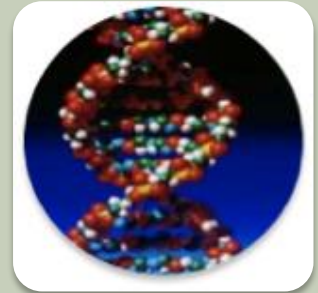
**Compute-
Intensive
Dense LA
for Deep
Learning
and
Simulation**



**Nearest
Neighbor
Simulation**



**All-to-All
Simulation
(3D FFTs)
and
analysis**

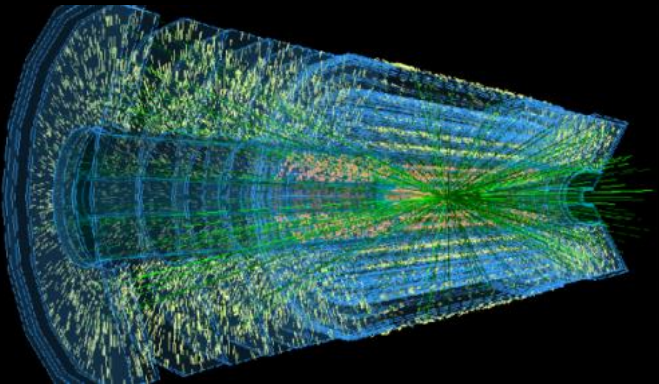


**Random
access,
large data
Analysis**

**Different architectures for simulation? Can
simulation use data architectures?**

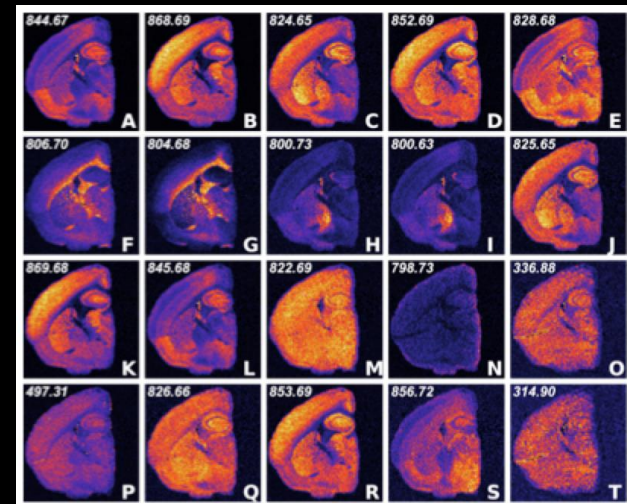
Data processing with special purpose hardware

- General trend towards specialization for continued performance growth
- Data processing (on raw data) will be first in DOE



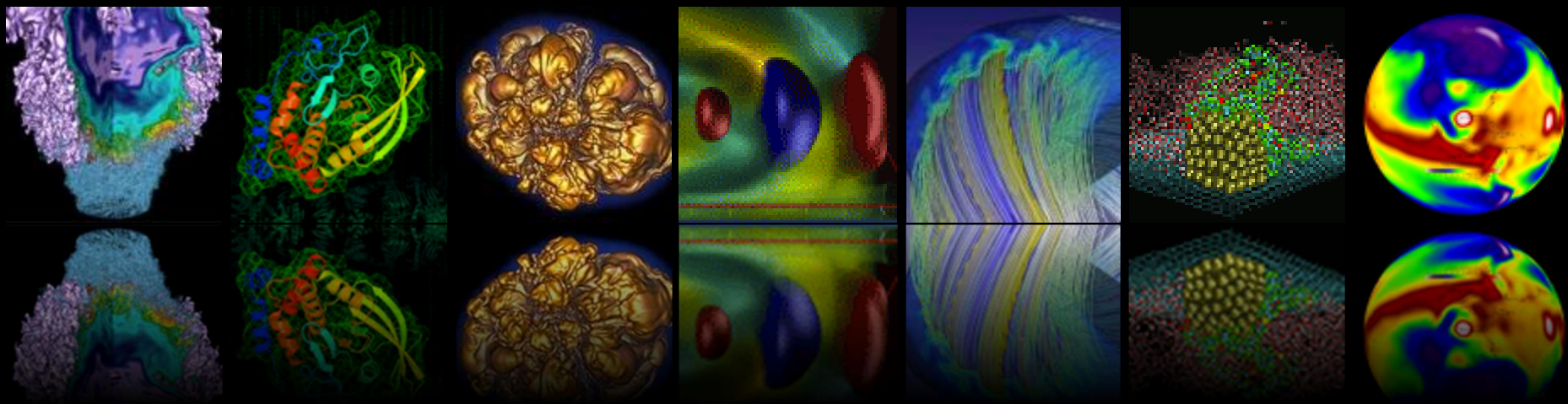
Particle Tracking with Neuromorphic chips

Computing in Detectors



Deep learning processors for image analysis

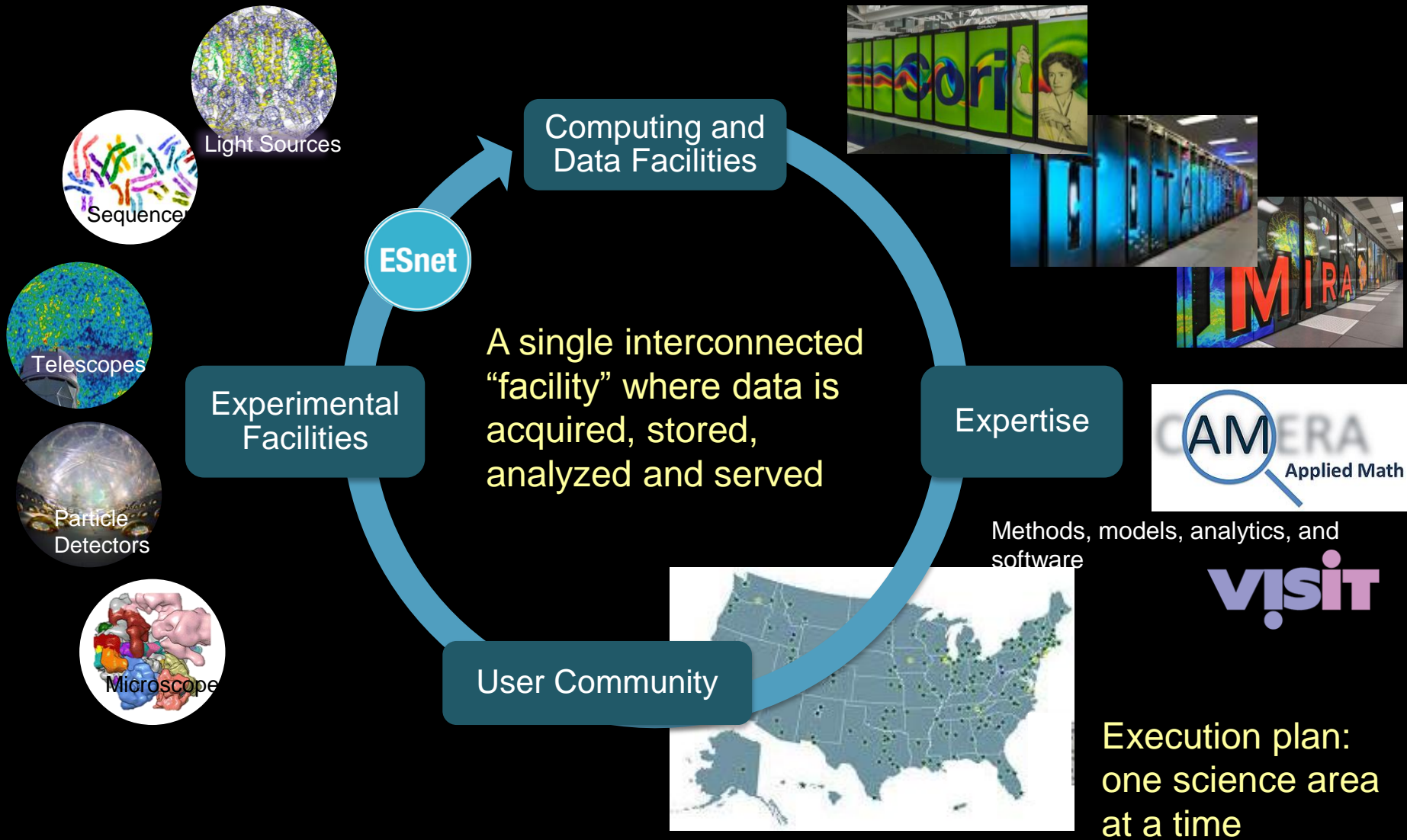
FPGAS for genome analysis



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.

Superfacility: Integrated network of experimental and computational facilities and expertise



Execution plan:
one science area
at a time