Lawrence Berkeley National Laboratory

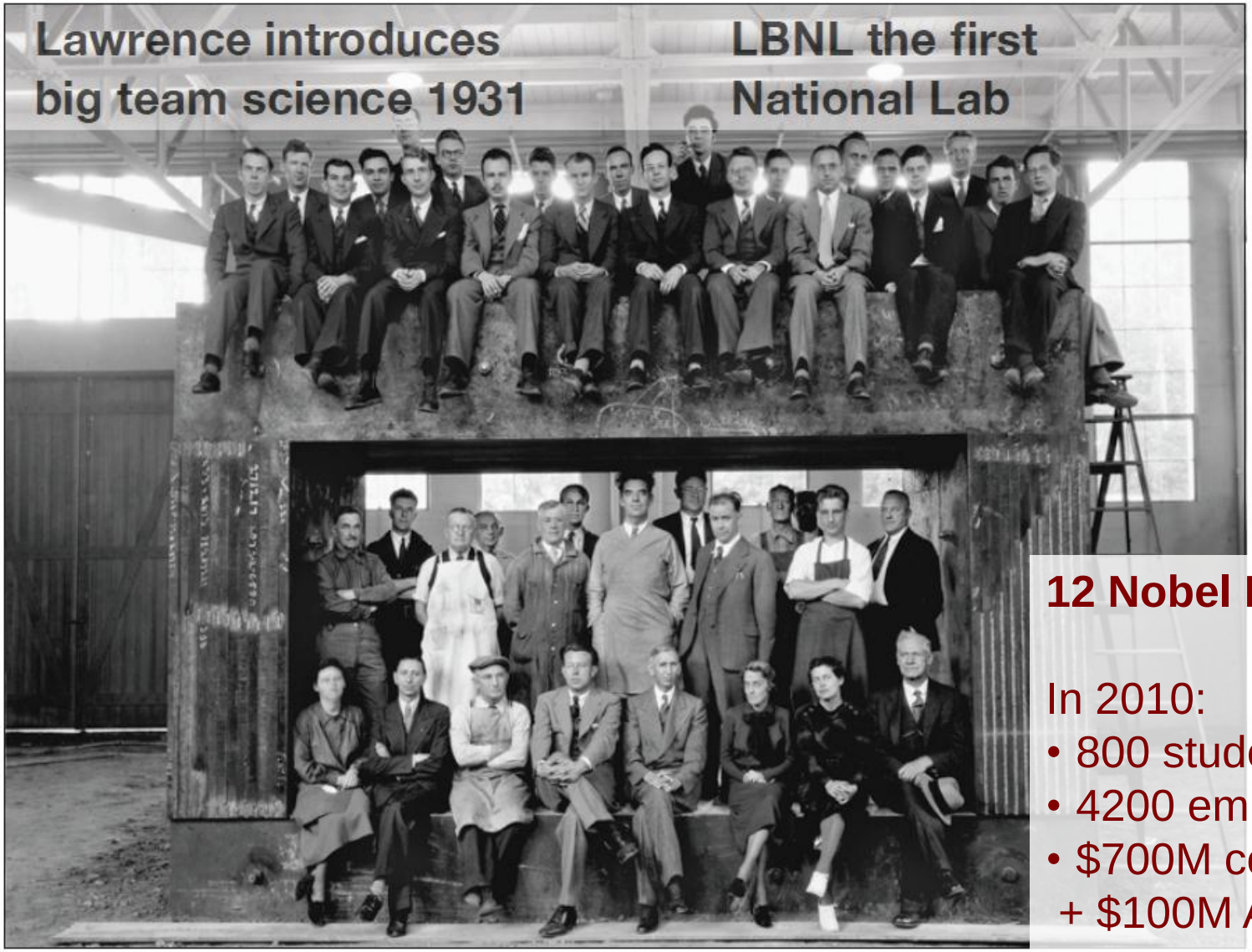
Kathy Yelick

Associate Laboratory Director for Computing Sciences

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Berkeley Lab Changes Science



12 Nobel Laureates

In 2010:

- 800 students
- 4200 employees
- $700M core
  + $100M ARRA

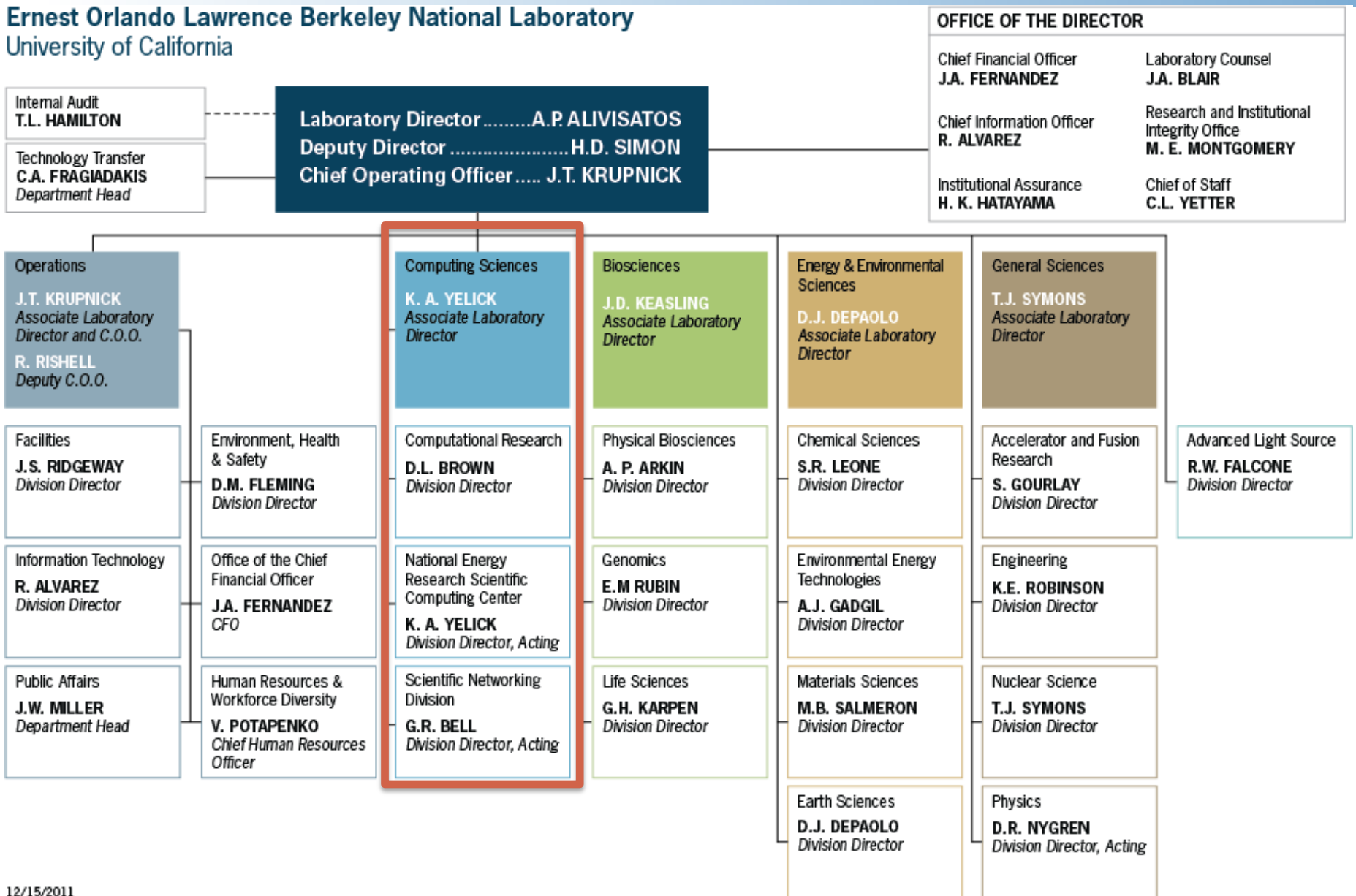
Berkeley Lab Mission

- **Solve most pressing and profound scientific problems facing humankind**
  - Basic science for a secure energy future
  - Understand living systems to improve the environment and energy supply
  - Understand matter and energy in the universe
- **Build and safely operate world-class scientific facilities**
- **Train the next generation of scientists & engineers**

U.S. DEPARTMENT OF ENERGY | Office of Science

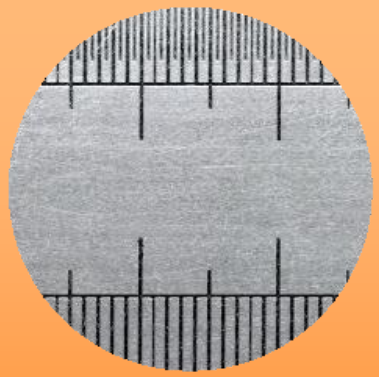
BERKELEY LAB
Lawrence Berkeley National Laboratory

LBNL Organization



MULTIDISCIPLINARY SCIENCE 2020

Societal needs for technical solutions to energy and environment problems will intensify



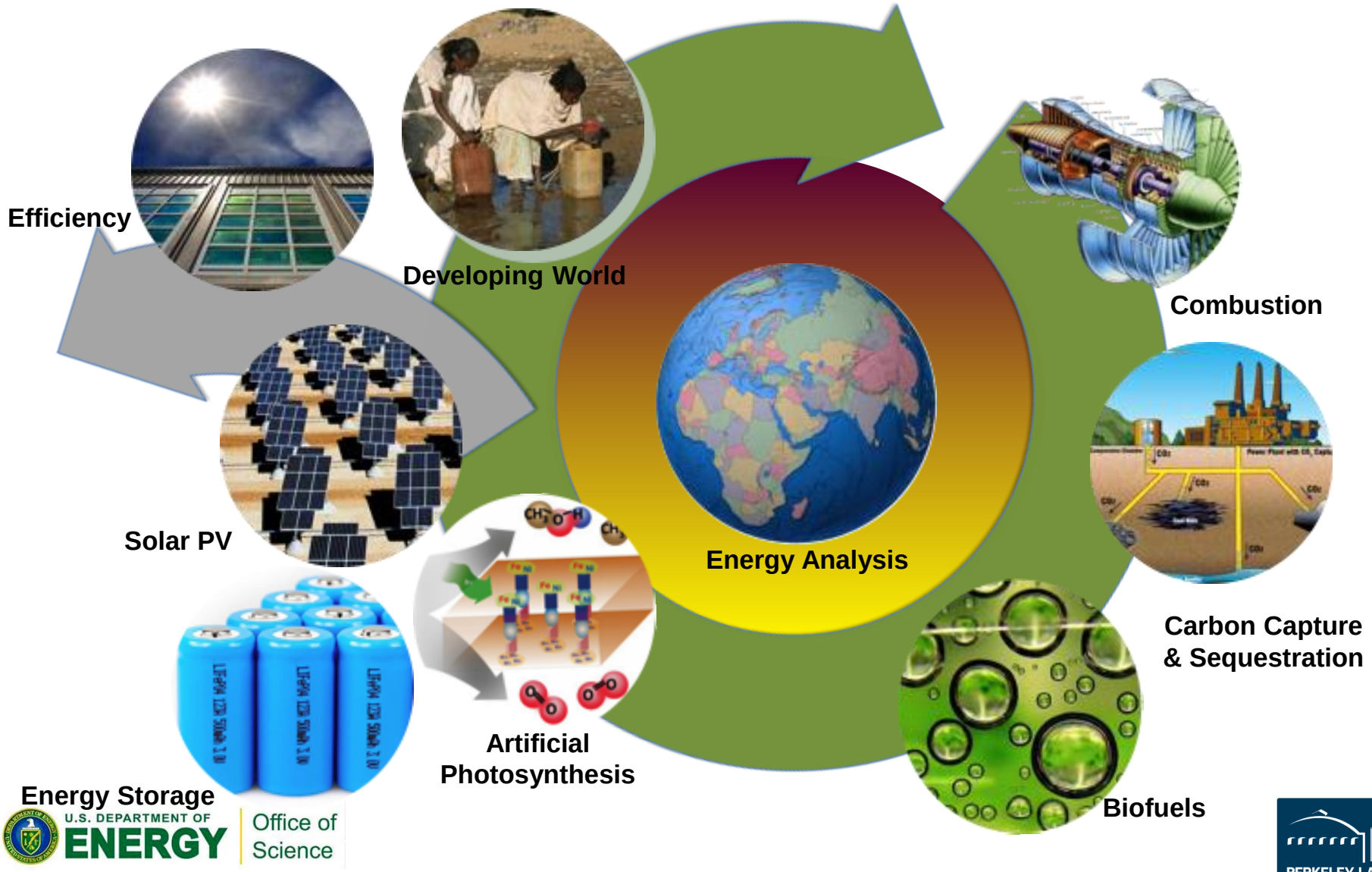
Measurement tools will open realms of inquiry

Biology revolution will impact other disciplines

Reliance on computation will grow with data

Fundamental discoveries in basic science will strengthen our foundation.

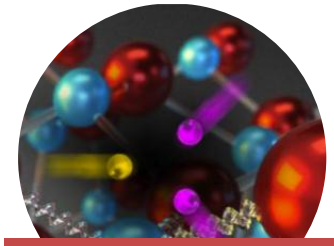
Carbon Cycle 2.0 Initiative at Berkeley Lab

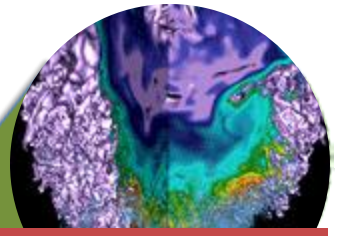


U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Computing for Carbon Cycle 2.0





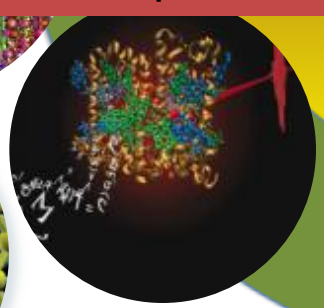
Materials for Efficient Lighting

All of these fields need more computing and more data!

1) To quantify and reduce uncertainty in simulation and to analyze the data sets from experimental devices (increase number of simulations)
2) Analyze data from experiments and simulations

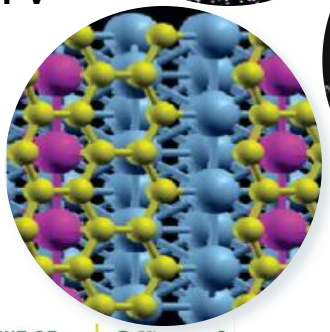
ptimized
mbustion
Burners

Materials for Solar PV



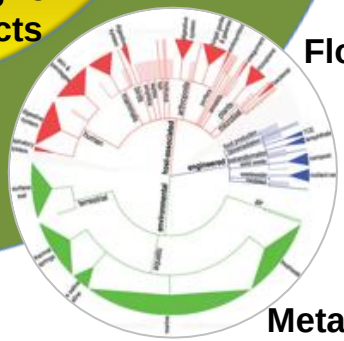
Climate Modeling for Analyzing Impacts

Flows through Porous Media for CCS



Chemistry for Catalysis in Batteries

Quantum Dynamics of Photosynthesis

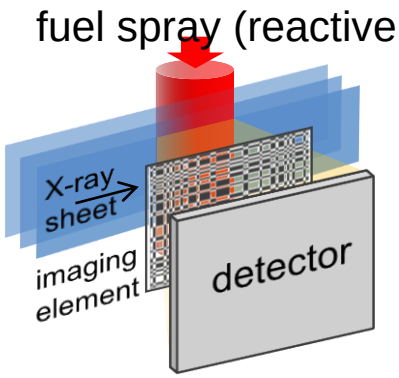


Metagenomics for BioFuels

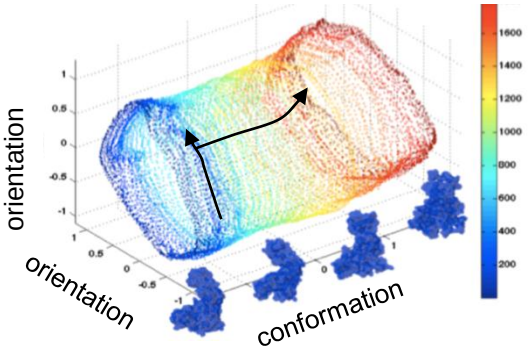
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

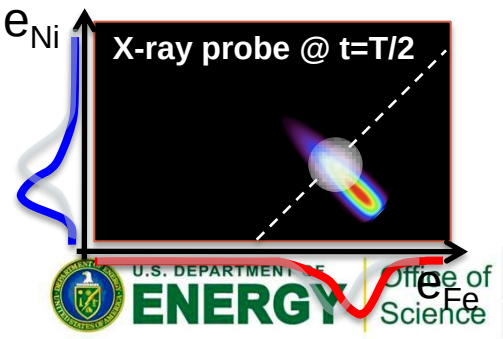
Photon science
NGLS – a transformative tool for energy science



- **Imaging:** from stills to movies



- **Structure:** systems that *change* conformation or are heterogeneous

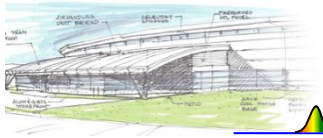


- **Spectroscopy:** multiple pulse techniques

More than 150 contributors representing more than 40 research institutions contributed to the NGLS CD-0 science case

Photon science
NGLS – high repetition rate, ultrafast pulses

Today's storage ring x-ray sources



~ nanoseconds

nanojoule

~ picoseconds

Weak pulses at high rep rate

Today's x-ray laser sources

~ milliseconds

Europe X-FEL

millijoule

~ femtoseconds

Intense pulses at low rep rate

comparable average power to SR

Tomorrow's NGLS

~ microseconds

millijoule

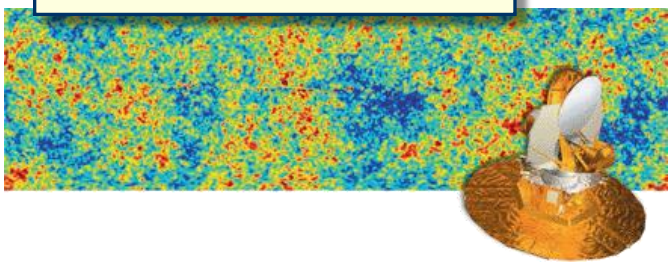
~ attoseconds to femtoseconds

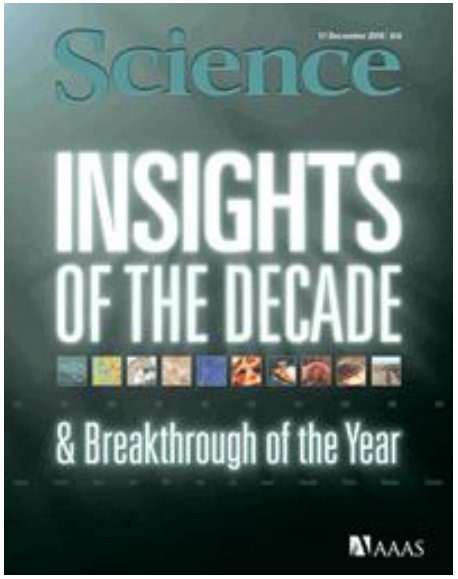
Intense pulses at high rep rate

>100x average power

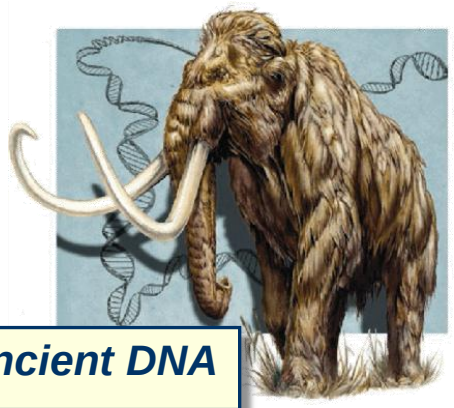
4 of 10 Science "Insights of the Decade" Involved LBNL CS Research or Facilities

Precision Cosmology

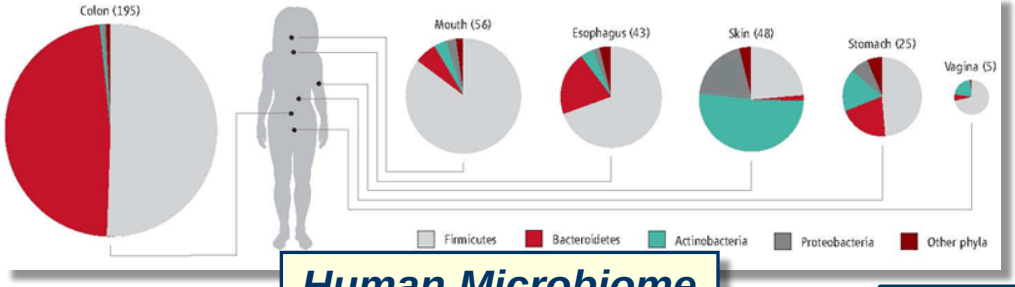




Genetic Dark Matter



Ancient DNA



Human Microbiome

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

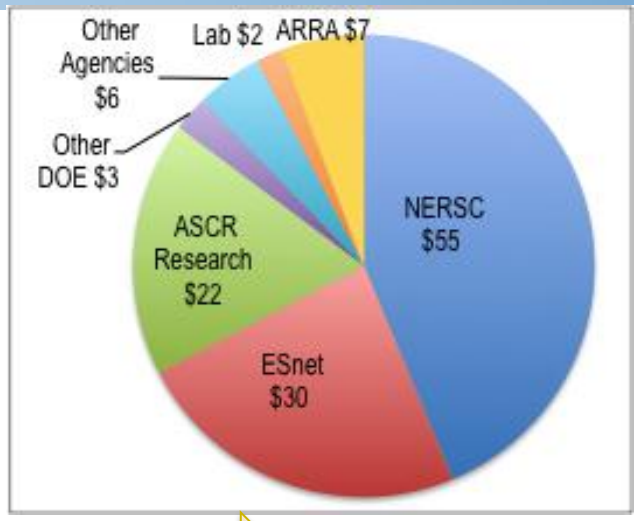
Computing Sciences at Berkeley Laboratory

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Computing Sciences at Berkeley Lab

- **Mission: Accelerate scientific discovery across a *broad community* through advanced computing & mathematics**
- **Goal: To be the leader delivering scientific output from computing through advanced research & facilities**



Computational Science

All of DOE Science and Applied Offices including science areas within the LBNL and CRD

Applied Mathematics

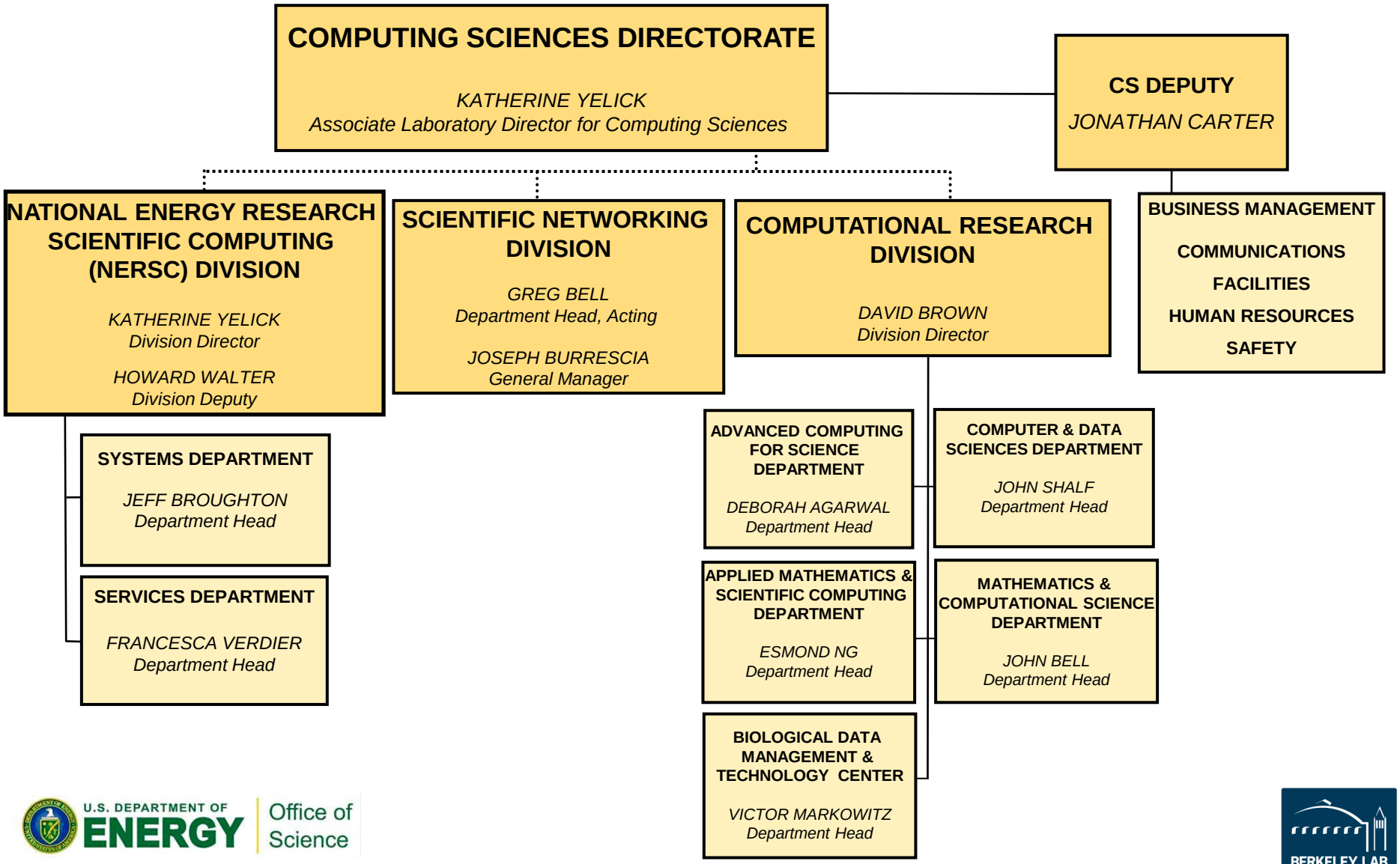
Computer Science

NERSC and ESnet Facilities

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Organization of Computing Sciences

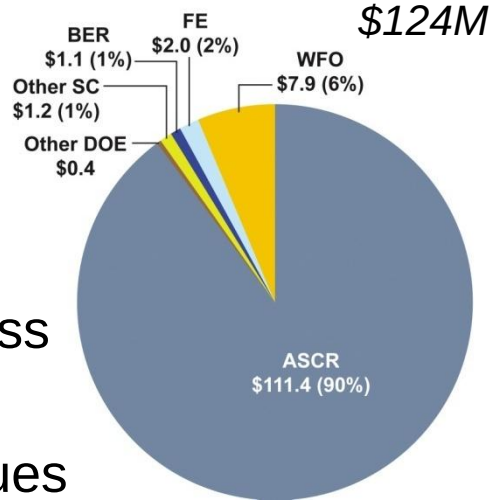


U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

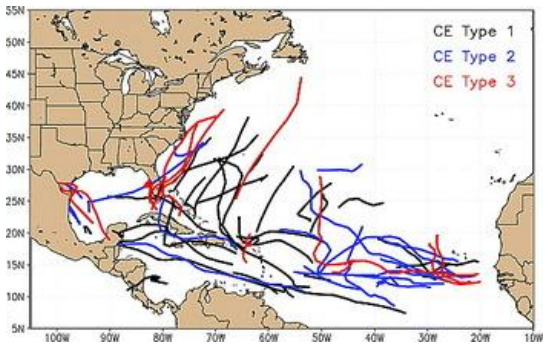
Computing sciences: petaflops to the people

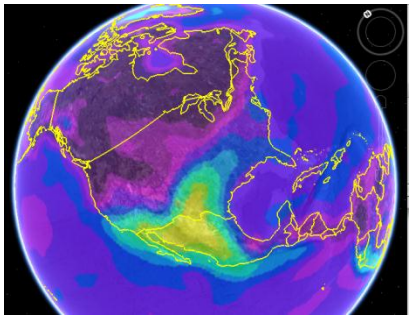
Vision: Accelerate scientific discovery across a broad community through advanced computing & math



- **Energy efficient computing:** Improve application performance per Watt by 100x necessary for exascale
- **Data driven computing:** Improve insight through access to and analysis of data
- **Mathematical Foundry:** Devise mathematical techniques for new science domains, at higher fidelities, and with higher confidence

- **Facilities**
  - **NERSC: 4000 users**
  - **ESnet: 27,000 users**





Reconstructed 3D weather maps at 6-hour intervals from 1871 to 2010. Used 20M NERSC hours resulting in public data set and 16 publications to date

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

NERSC Facility Leads DOE in Scientific Computing Productivity

NERSC computing for science

- 4000 users, 500 projects
- From 48 states; 65% from universities
- Hundreds of users each day
- 1500 publications per year

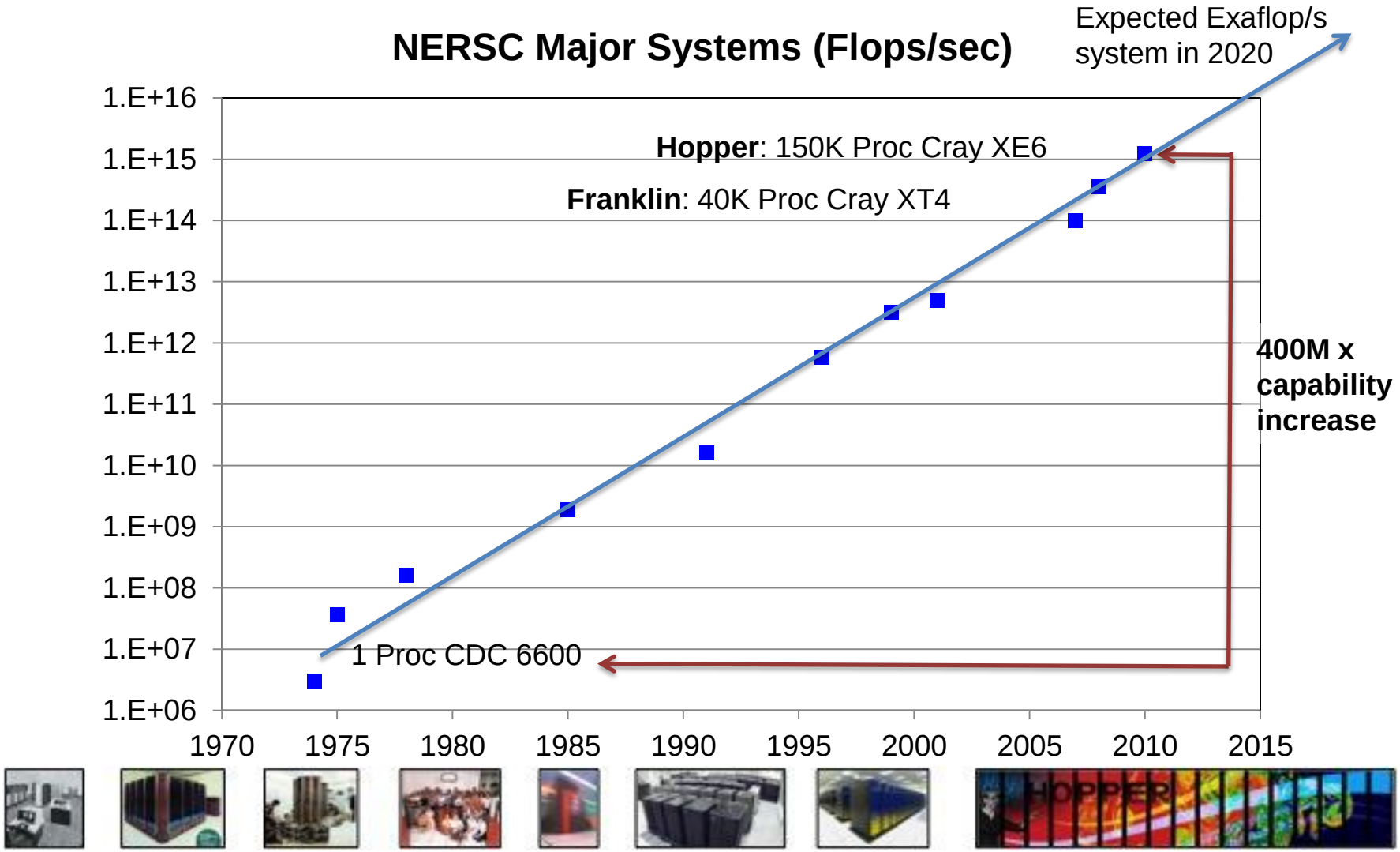
Systems designed for science

- 1.3PF Petaflop Cray system, Hopper
  - 2nd Fastest computer in US and one of 2 Petaflop systems in Office of Science
  - Additional .5 PF in Franklin system and smaller clusters

U.S. DEPARTMENT OF ENERGY | Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Computing Capability at NERSC



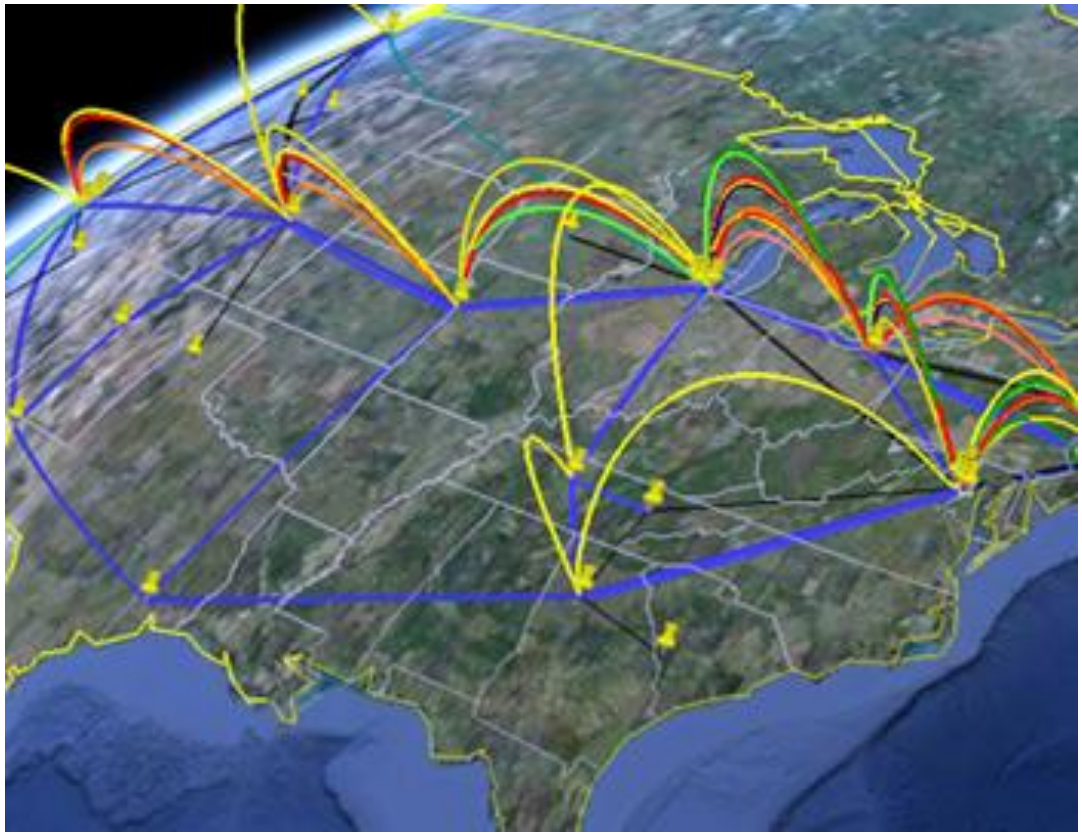
ESnet: DOE's Leadership and Production Network

DOE Science Network:

- 72% annual traffic growth
- International collaborations
- Bandwidth reservations and monitoring

Advanced Networking Initiative (ANI):

- 100 Gbps network
- Demo at SC11
- Will transition to production
- Separate network research testbed serving 17 projects



ESnet+ANI, DOE will be the world leader in networking for science.

Challenges to ~~Exascale~~ Performance Growth

1) **System power** is the primary constraint
2) **Concurrency** (1000x today)
3) **Memory** bandwidth and capacity are not keeping pace
4) **Processor** architecture is an open question
5) **Programming model** heroic compilers will not hide this
6) **Algorithms** need to minimize data movement, not flops
7) **I/O bandwidth** unlikely to keep pace with machine speed
8) **Reliability and resiliency** will be critical at this scale
9) **Bisection bandwidth** limited by cost and energy

Unlike the last 20 years most of these (1-7) are equally important across scales, for midrange, high end, leadership

DOE Facilities will have Massive Computing

Astronomy

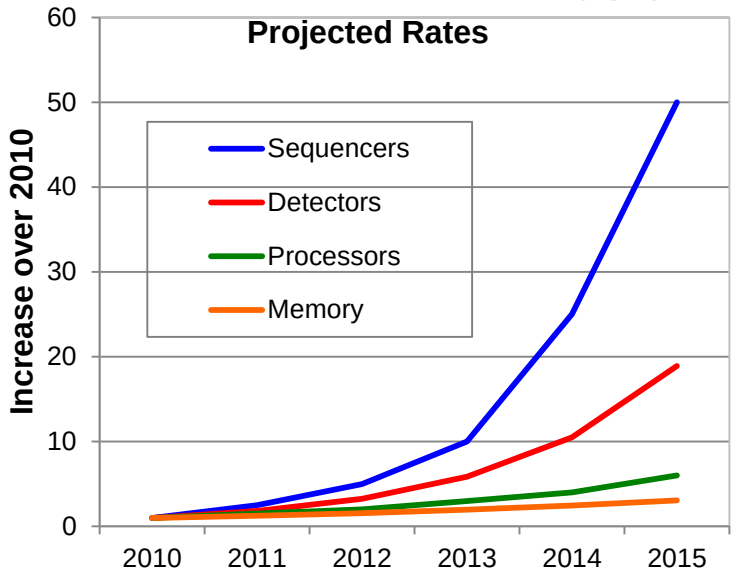
Particle Physics

Chemistry and Materials

Genomics

Fusion

- Petabyte data sets today, many growing exponentially
- Outpacing storage, networking, memory, and compute
- **Data is growing at 58% while storage is growing at 40%.**

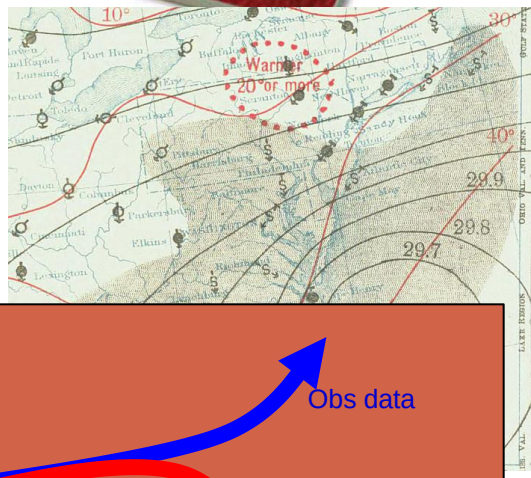


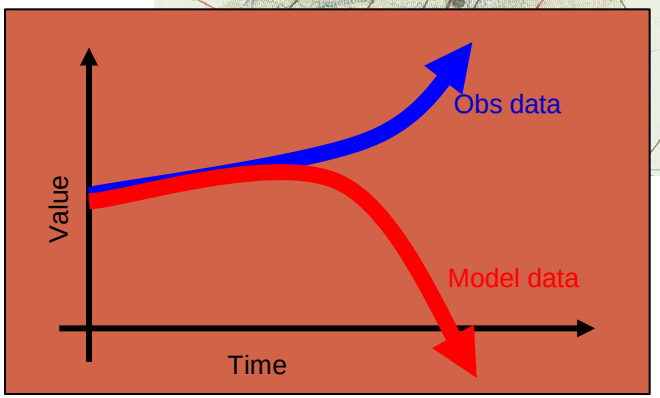
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

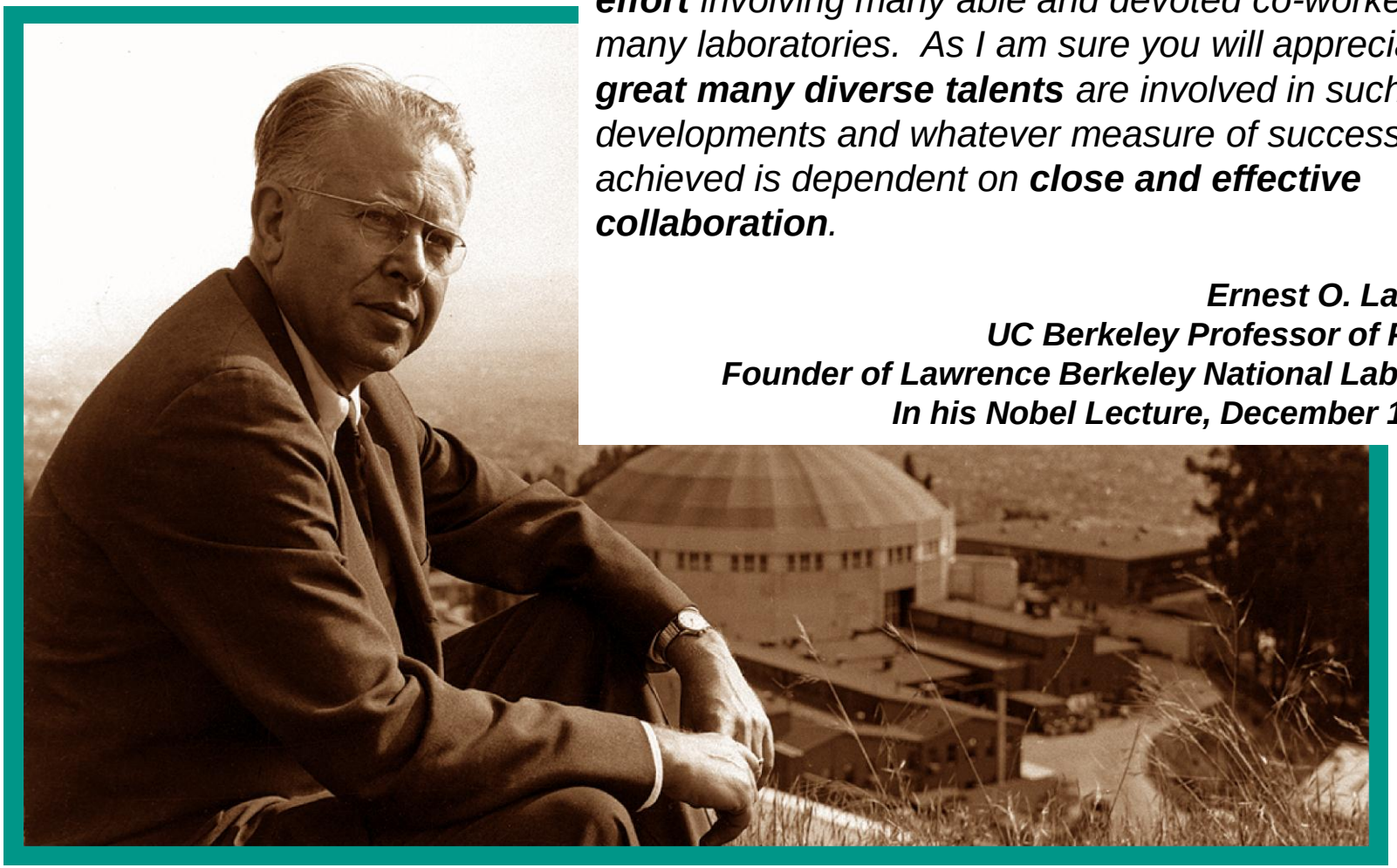
Data Challenges

- Hardware balance (data vs flops)
- Where to cut data between experimental and compute facilities?
- Analysis algorithms for incomplete or noisy data
- Archival, provenance issues
- Algorithms and data structures for indexing and searching
- Workflows
- Networking research
- Programming models





A History of Collaborative Science



*… as from the beginning the work has been a **team effort** involving many able and devoted co-workers in many laboratories. As I am sure you will appreciate, a **great many diverse talents** are involved in such developments and whatever measure of success is achieved is dependent on **close and effective collaboration**.*

Ernest O. Lawrence
UC Berkeley Professor of Physics
Founder of Lawrence Berkeley National Laboratory
In his Nobel Lecture, December 11, 1951