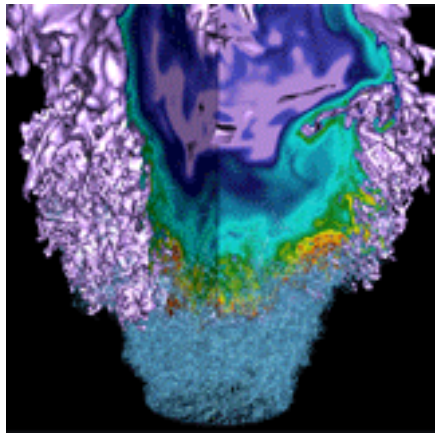
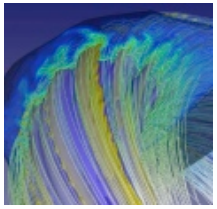
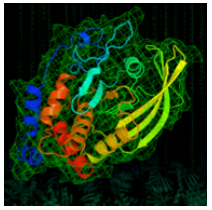
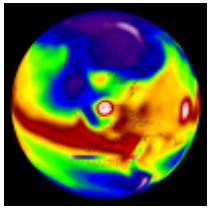
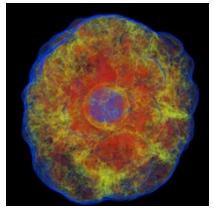
Data Intensive Science at NERSC

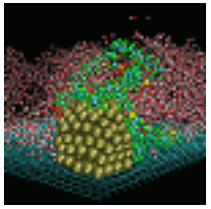










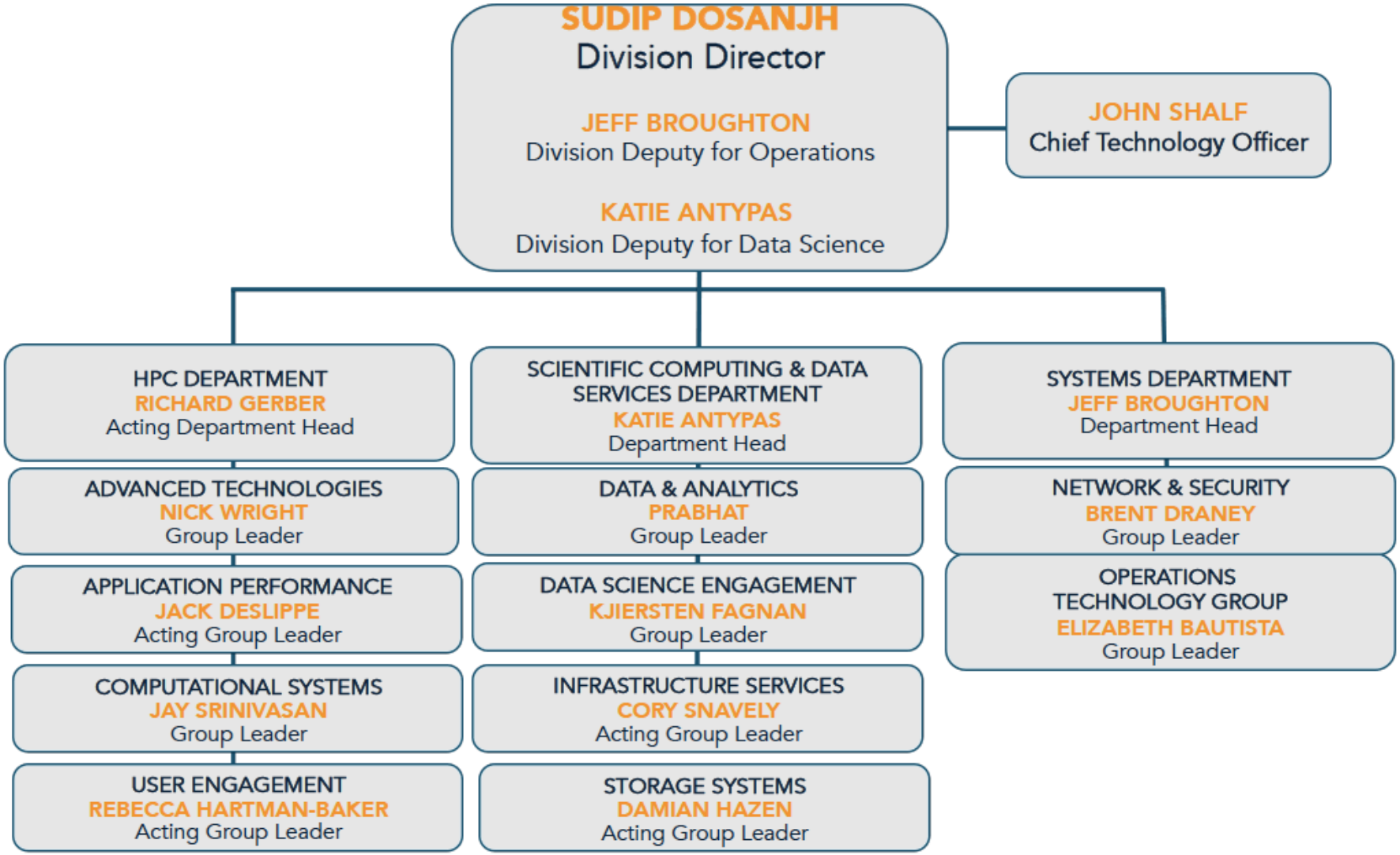


Prabhat

Data and Analytics Services Group Lead

March 24, 2016

NERSC Organization



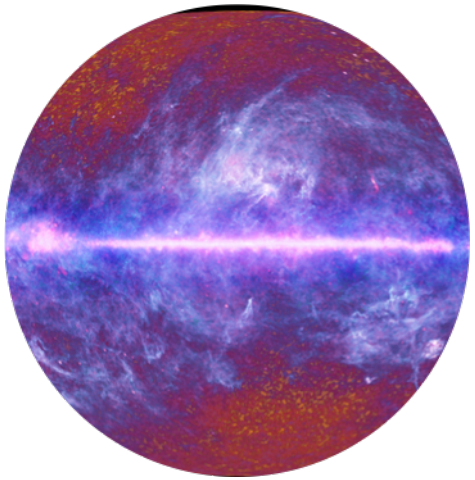
DAS Goal: "Enable Data-Intensive Science at Scale"

- **Internal Goals**
  - Provide production quality software services for all major Data capabilities:
    - Analytics, Management, Workflows, Transfer, Access, Visualization
  - Pioneer evaluation, research and deployment of Big Data technologies
    - Focusing on productivity and performance
  - Engage with stakeholders to enable scientific discovery in a data-driven world
    - Users
    - Computing Sciences Staff
    - Vendors
    - Researchers

DAS Team

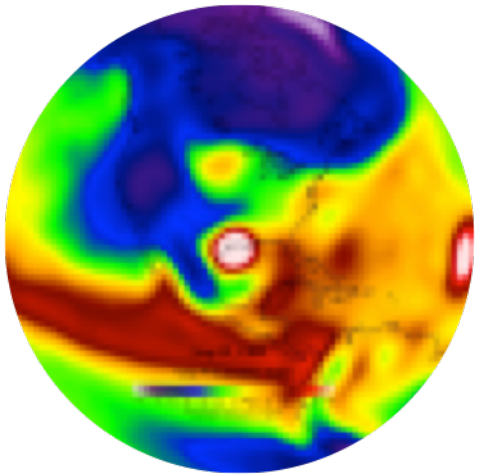
DAS member	DAS capability area	Science expertise
Debbie Bard	Analytics, Workflows	Cosmology, Astronomy
Wahid Bhimji	Management, Transfer	High Energy Physics
Shane Canon	Workflows	KBase, JGI
Shreyas Cholia	Access, Management	
Lisa Gerhardt	Management, Workflows	Particle Physics
Annette Greiner	Access	Systems Biology
Quincey Koziol	Management	
Jialin Liu	Management	
Jeff Porter	Transfer	Nuclear Physics
Prabhat	Analytics, Management	Climate
Evan Racah	Analytics	
Rollin Thomas	Analytics, Access	Astrophysics, Cosmology

DOE Facilities are Facing a Data Deluge

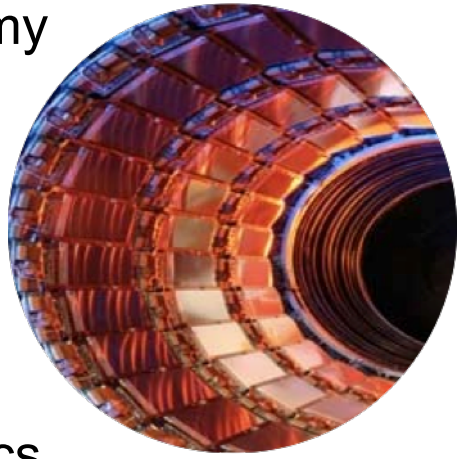


Astronomy

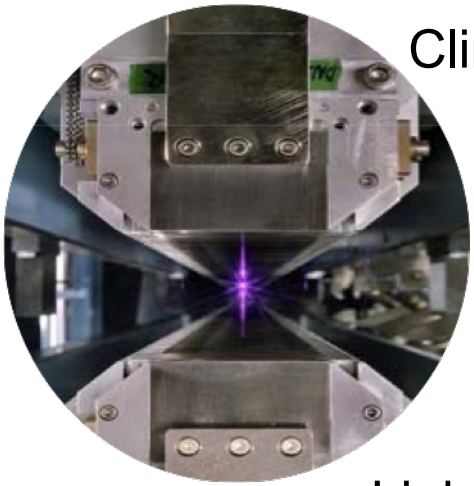
Genomics



Climate



Physics

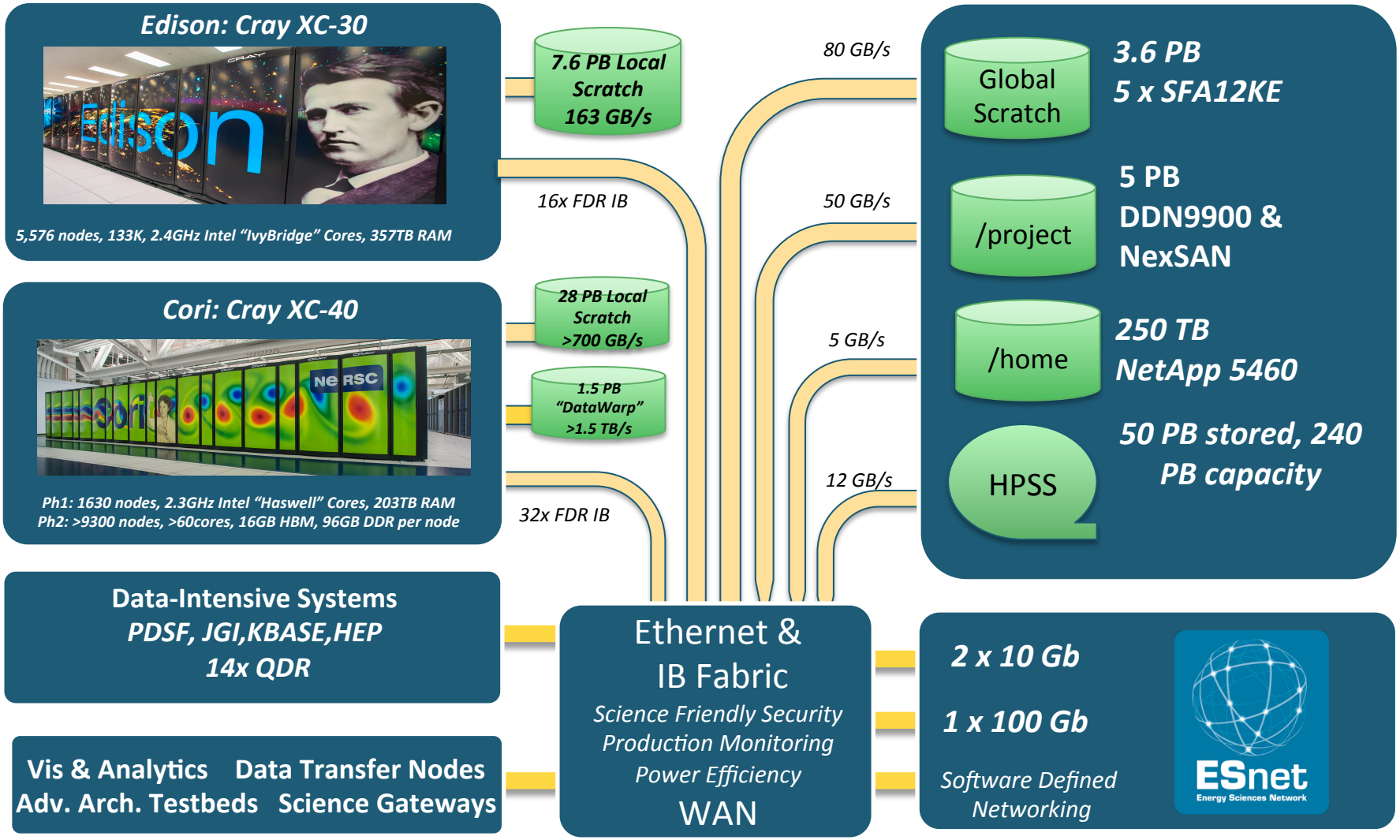


Light Sources

4 V's of Scientific Big Data

Science Domain	Variety	Volume	Velocity	Veracity
Astronomy	Multiple Telescopes, multi-band/spectra	O(100) TB	100 GB/night – 10 TB/night	Noisy, acquisition artefacts
Light Sources	Multiple imaging modalities	O(100) GB	1 Gb/s-1 Tb/s	Noisy, sample preparation/acquisition artefacts
Genomics	Sequencers, Mass-spec, proteomics	O(1-10) TB	TB/week	Missing data, errors
HEP: LHC, Daya Bay	Multiple detectors	O(100) TB – O(10) PB	1-10 PB/s reduced to GB/s	Noisy, artefacts, spatio-temporal
Climate	Simulations Multi-variate, spatio-temporal	O(10) TB	100 GB/s	'Clean', need to account for multiple sources of uncertainty

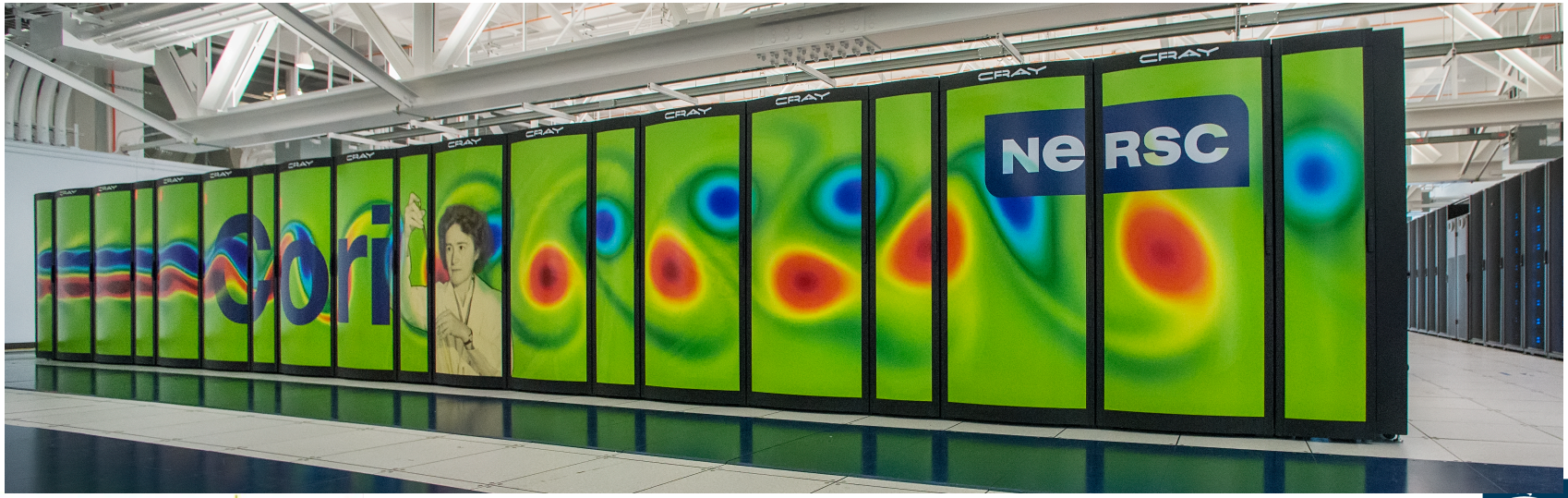
NERSC - 2016


Edison: Cray XC-30
5,576 nodes, 133K, 2.4GHz Intel "IvyBridge" Cores, 357TB RAM
7.6 PB Local Scratch 163 GB/s
16x FDR IB
80 GB/s
50 GB/s
Global Scratch
3.6 PB
5 x SFA12KE
/project
5 PB
DDN9900 & NexSAN
/home
250 TB
NetApp 5460
HPSS
50 PB stored, 240 PB capacity
Cori: Cray XC-40
Ph1: 1630 nodes, 2.3GHz Intel "Haswell" Cores, 203TB RAM
Ph2: >9300 nodes, >60cores, 16GB HBM, 96GB DDR per node
28 PB Local Scratch >700 GB/s
1.5 PB "DataWarp" >1.5 TB/s
5 GB/s
12 GB/s
32x FDR IB
Data-Intensive Systems
PDSF, JGI,KBASE,HEP
14x QDR
Vis & Analytics Data Transfer Nodes
Adv. Arch. Testbeds Science Gateways
Ethernet & IB Fabric
Science Friendly Security
Production Monitoring
Power Efficiency
WAN
2 x 10 Gb
1 x 100 Gb
Software Defined Networking
ESnet
Energy Sciences Network

The Cori System

- **Cori will transition HPC and Data-intensive workloads to energy efficient architectures**

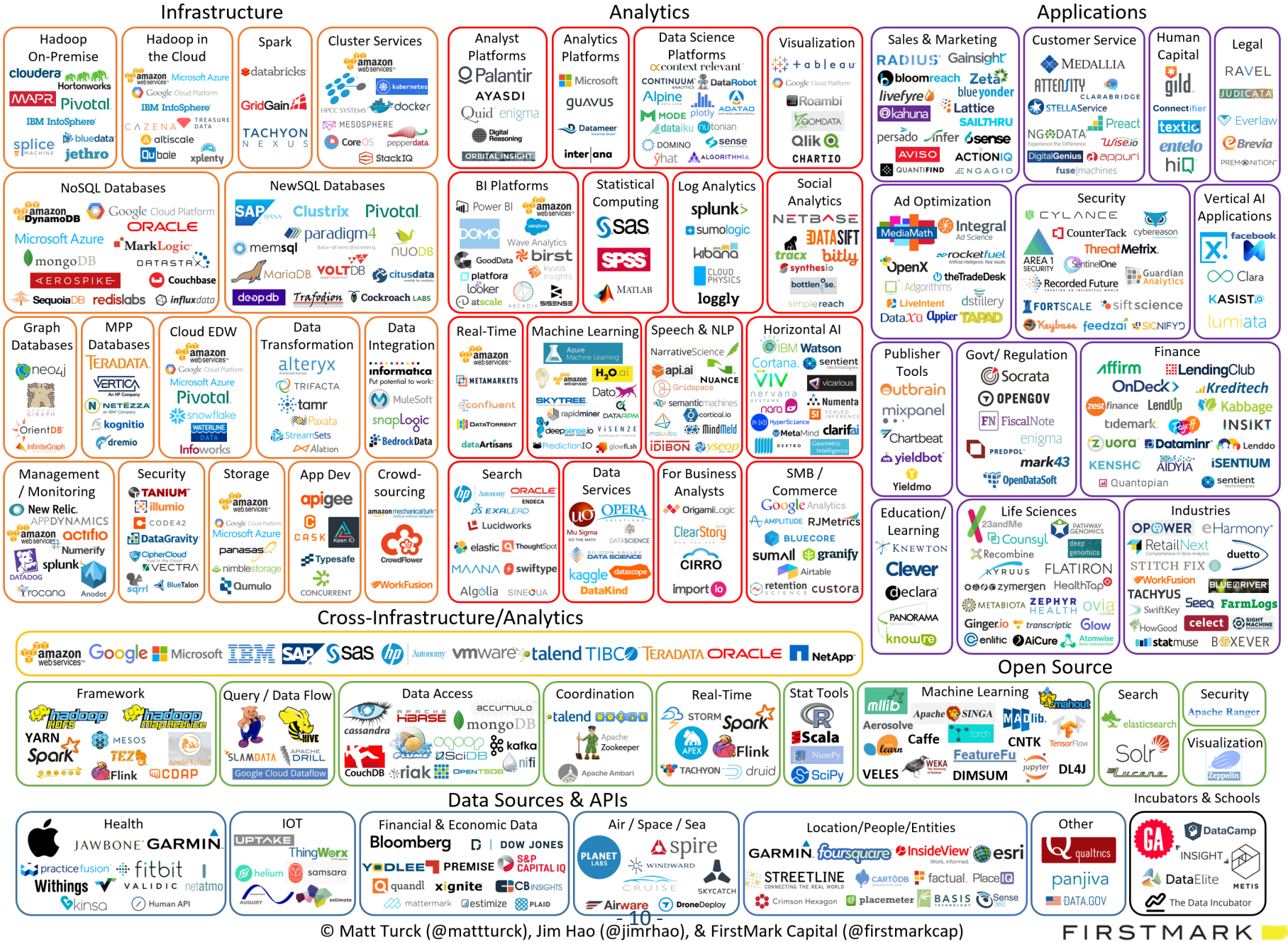
System named after Gerty Cori, Biochemist and first American woman to receive the Nobel prize in science.



Data-friendly features on Cori

Data Intensive Workload Need	Cori Solution
Fast I/O (local disk)	NVRAM 'burst buffer' Lustre filesystem
Large memory nodes	128 GB/node on Haswell High memory 775 GB/ login node
Complex workflows	14 external login nodes Spark/JupyterHub/workflow software
Communicate with databases from compute nodes	Improved RSIP handles most cases Actively exploring SDN
Stream Data from observational facilities	Software Defined Networking
Easy to customize environment	User-defined images w/ Shifter
Policy Flexibility Turnaround time	Real-time; High throughput; Serial queues Reservations

Big Data Landscape 2016



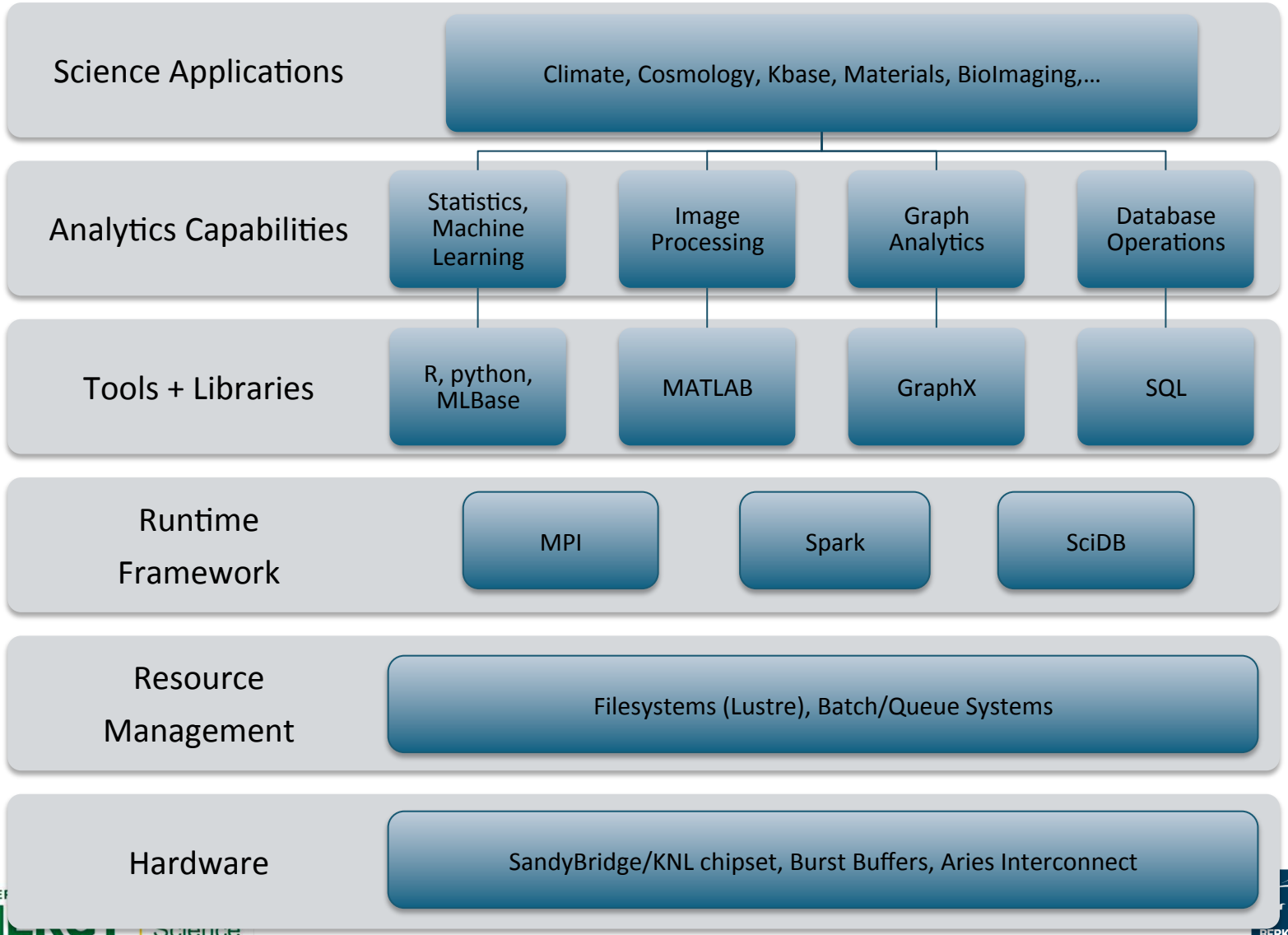
FIRSTMARK

Big Data Software Portfolio

Capabilities	Technology Areas	Tools, Libraries
Data Transfer + Access	Globus, Grid Stack, Authentication	Globus Online, Grid FTP
	Portals, Gateways, RESTful APIs	Drupal/PHP, Django/Python, NEWT
Data Processing	Workflows	Swift, Fireworks
Data Management	Formats, Models	HDF5, NetCDF, ROOT
	Databases	MongoDB, SciDB, PostgreSQL, MySQL
Data Analytics	Statistics, Machine Learning	python, R, ROOT, Spark iPython/Jupyter
	Imaging	OMERO, Fiji, MATLAB
Data Visualization	SciVis InfoVis	VisIt, Paraview

	Astronomy	Cosmology	Climate	Systems Biology	Neuroscience	BioImaging	Mass-spec Imaging	Personalized Toxicology	Materials	Particle Physics
Classification	X		X		X	X	X			X
Regression								X	X	
Clustering		X	X		X		X			X
Dimensionality Reduction			X		X		X			
Inference	X						X			X
Model Estimation	X				X			X		
Design of Experiments		X	X						X	
Semantic Analysis			X	X					X	
Feature Learning			X		X		X	X	X	X
Anomaly Detection	X		X							X

Analytics Software Strategy

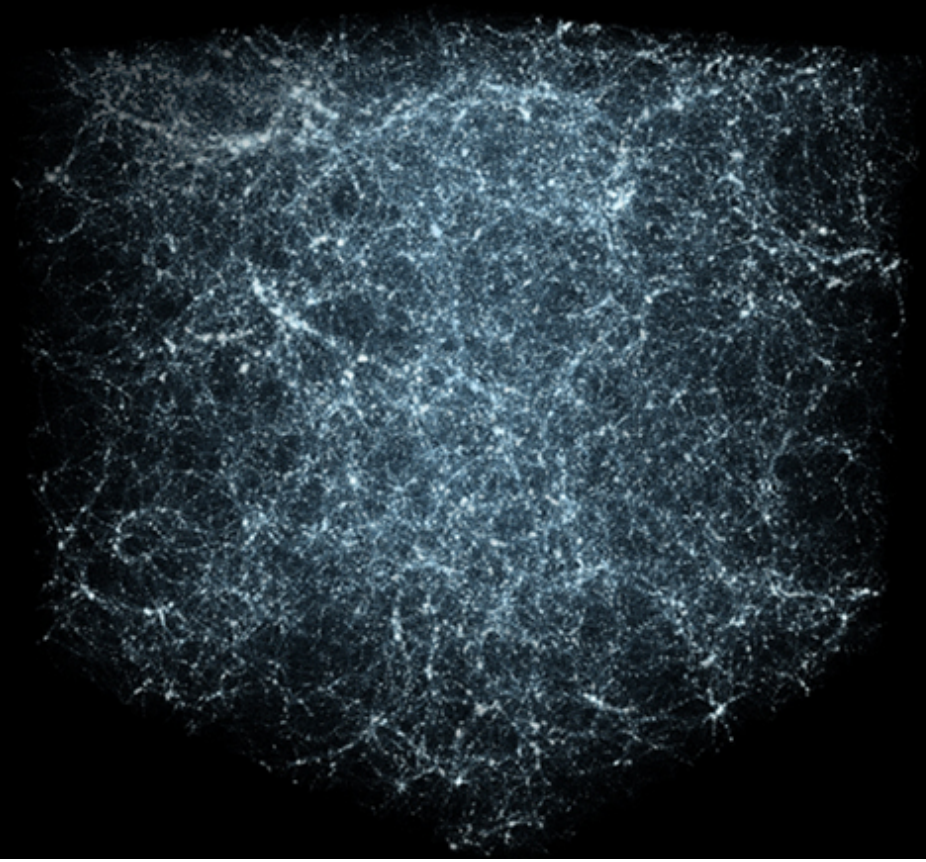


Data Analytics: Plans for next year

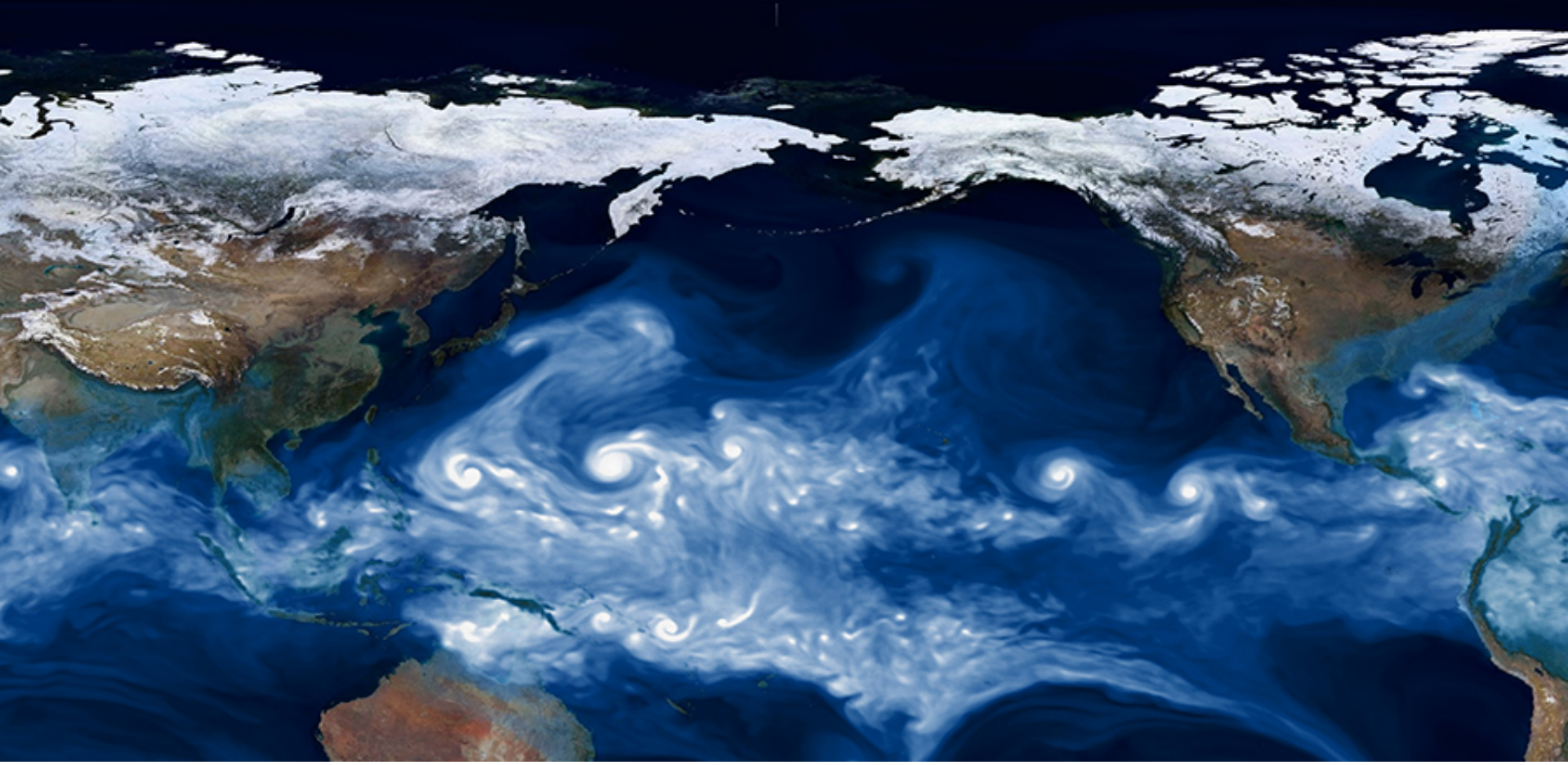
- **Productivity**
  - Jupyter/iPython notebooks for running codes and interfacing with Cori
- **Performance (HPC and HTC)**
  - Multi-core, multi-node versions of python, R, Spark
  - Processing 1TB datasets on 1,000 cores should be possible *now*
  - Aiming for processing 10TB datasets on 10,000 cores
- **Similar plans are being developed for Management, Transfer/Access, Workflows and Visualization.**

1 Creating a catalog of all objects in the Universe

2 Determining the Fundamental Constants of Cosmology



3 Characterizing Extreme Weather in a Changing Climate

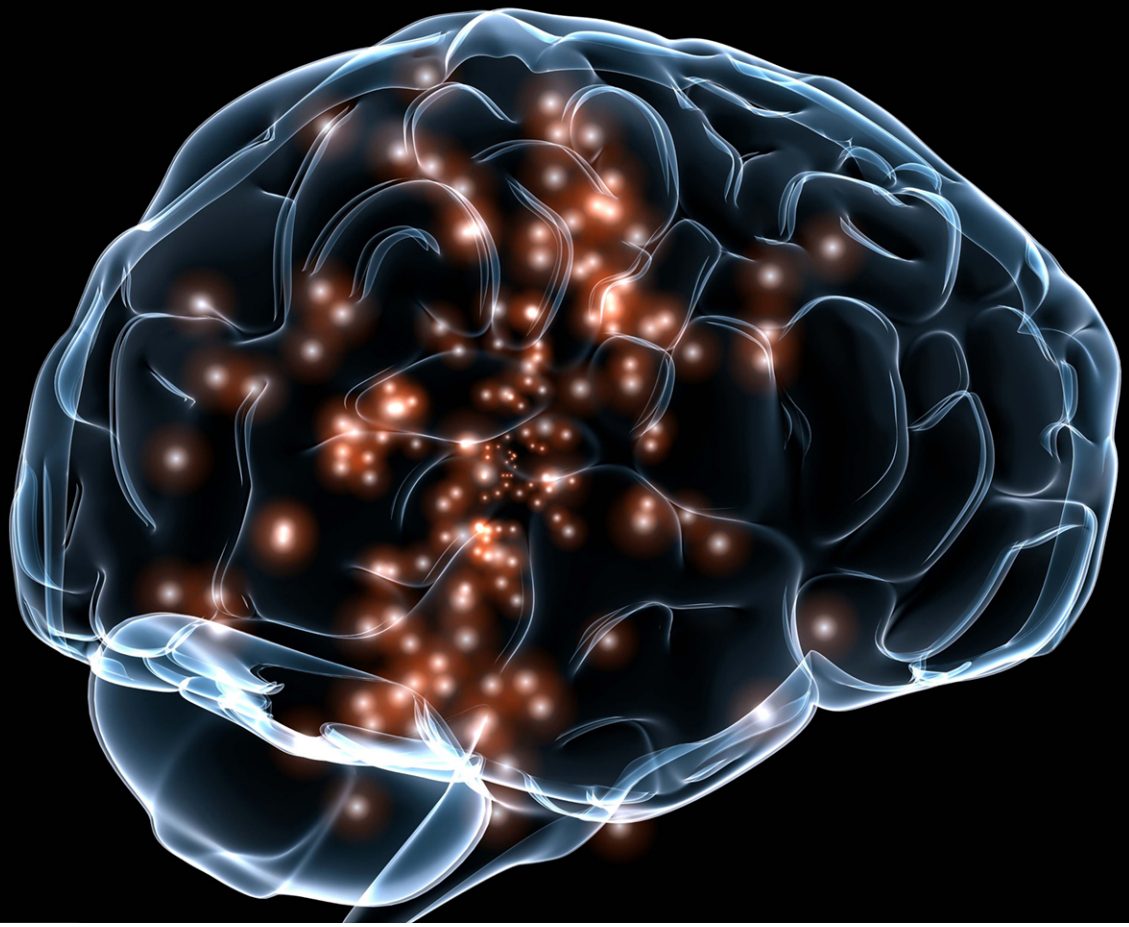


U.S. DEPARTMENT OF ENERGY | Office of Science

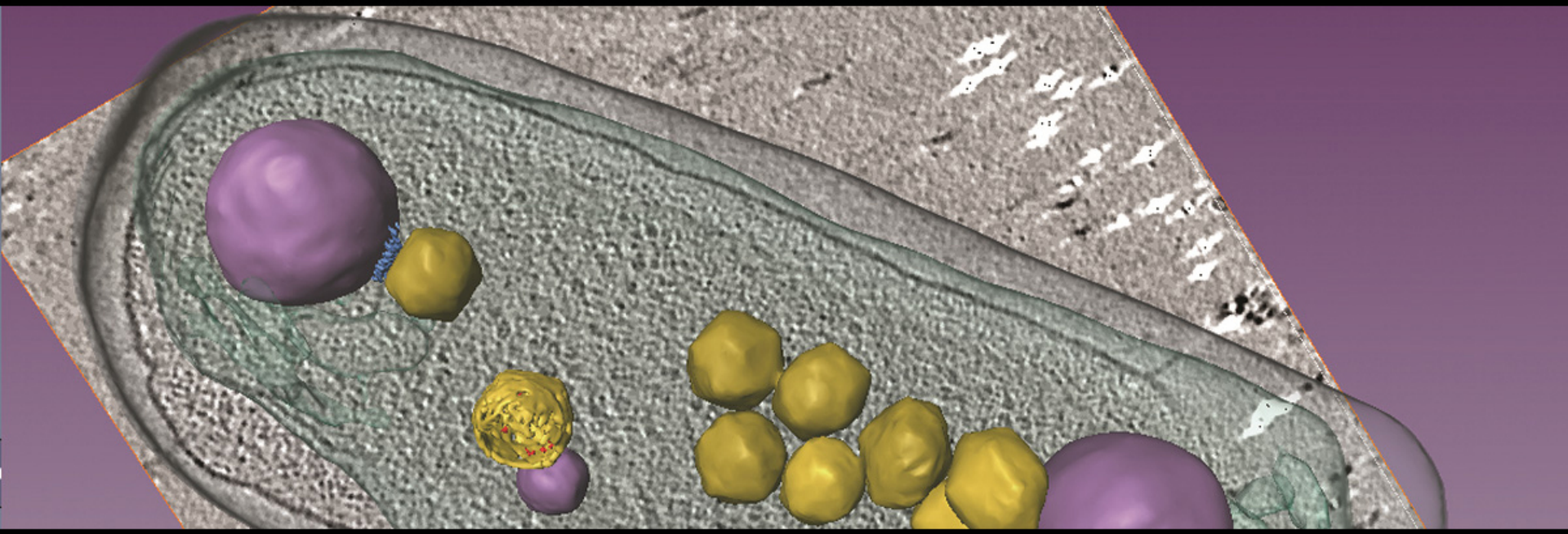
BERKELEY LAB
Lawrence Berkeley National Laboratory

4 Knowledge Extraction from Scientific Literature

5 Understanding Speech Production



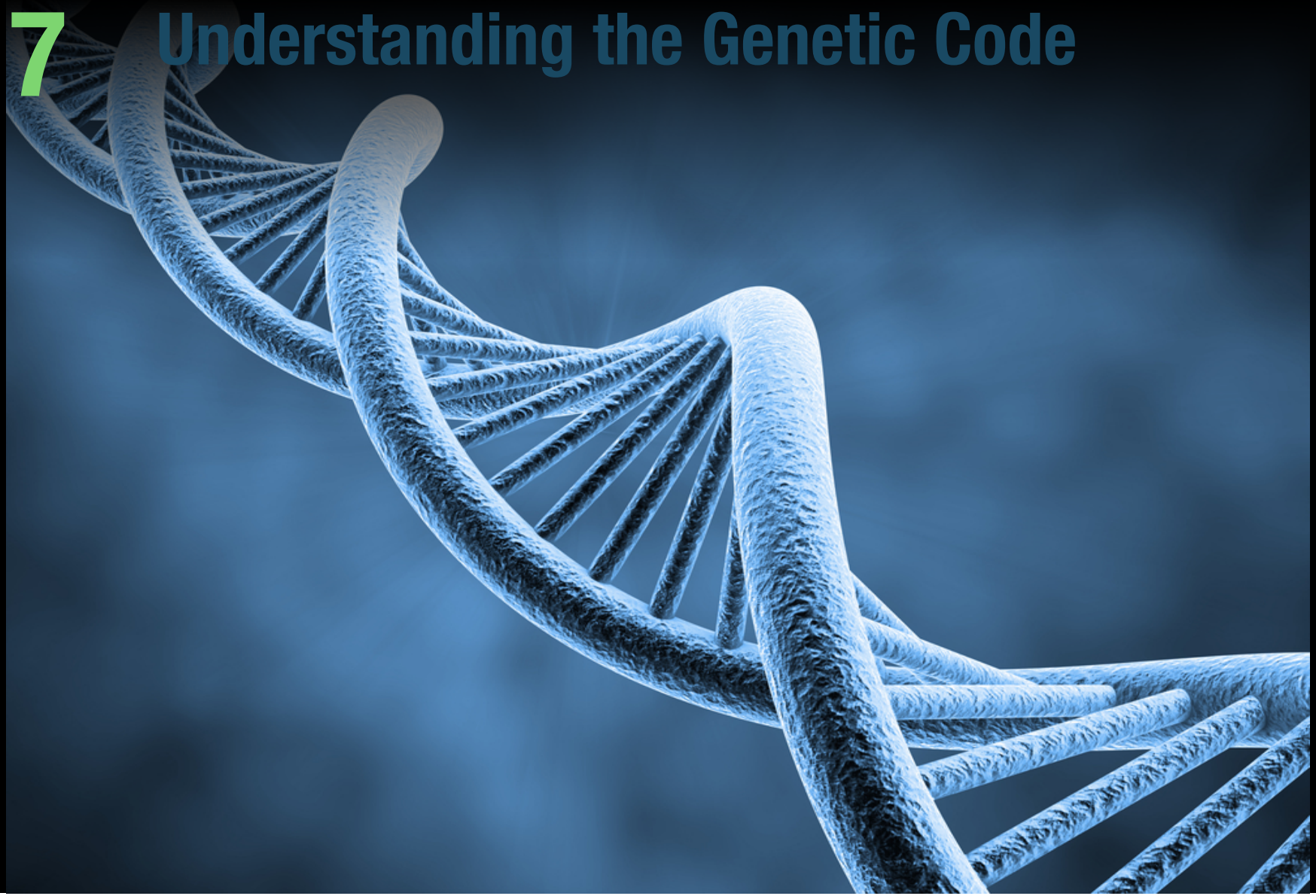
6 Quantitative and Predictive Biology



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

7 Understanding the Genetic Code

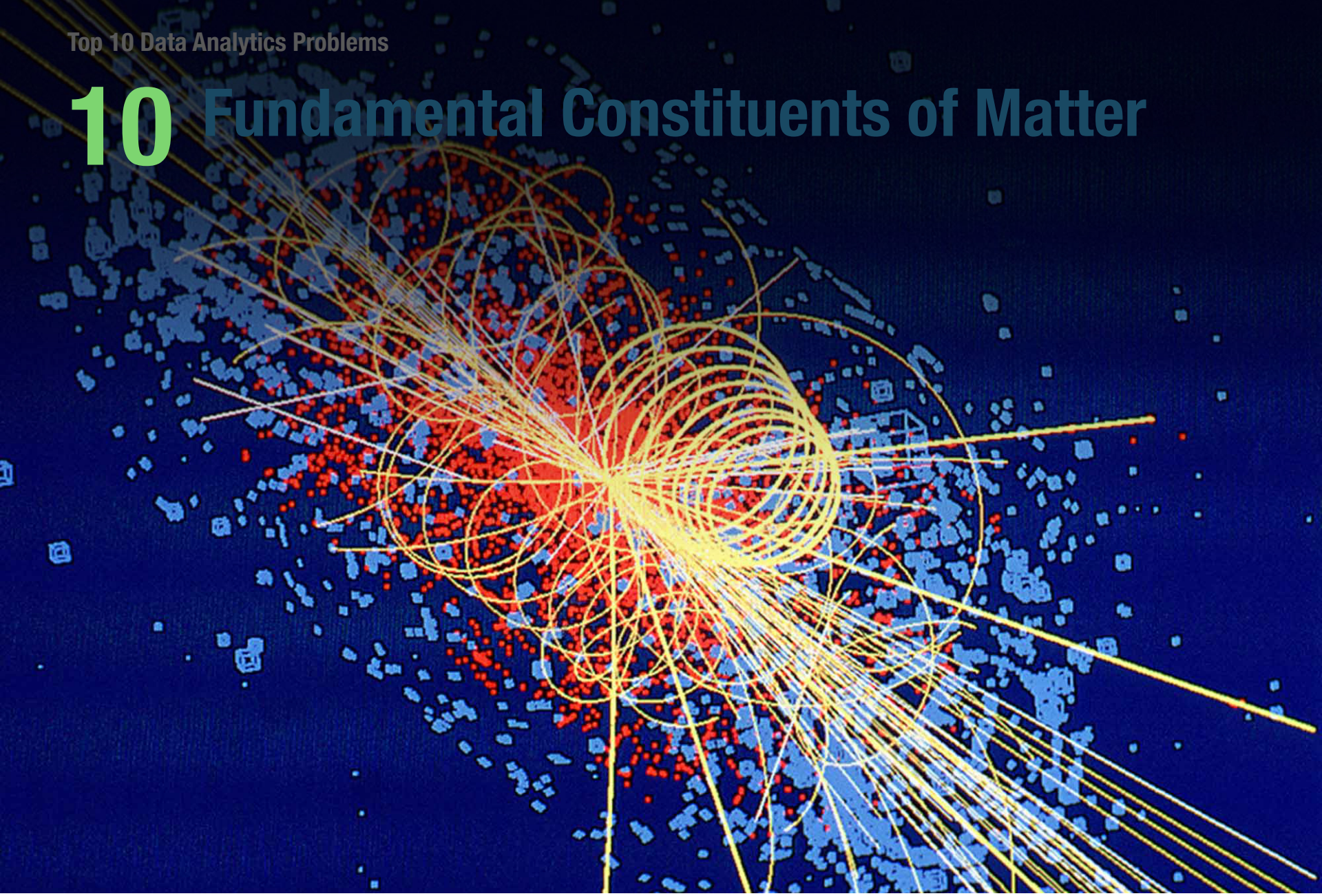


8 Personalized Toxicology

9 Designer Materials

Top 10 Data Analytics Problems

10 Fundamental Constituents of Matter



U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

Looking to the future: NERSC-9

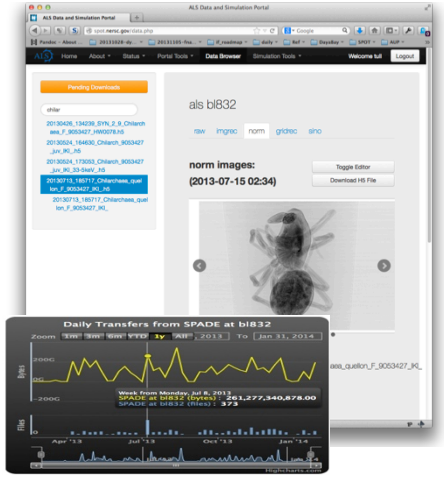
- **NERSC-9 will build upon the successes of the data components of Cori**
- **End to end workflow requirements and performance are critical for the design and optimization of the system**
- **Overall goal is to enable seamless data motion with dynamic allocation and scheduling of resources**
  - Enable first steps towards exascale-era storage system
  - Vendor community excited about engagement and collaboration opportunities

Looking to the future: Superfacility

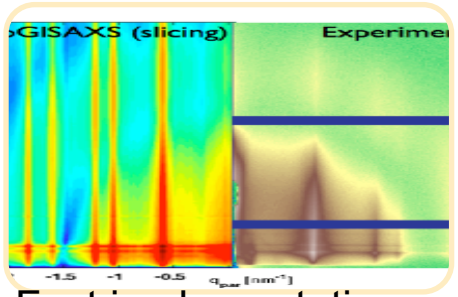
Experimental Facilities

New mathematical analyses

Integrated with ESnet: Designed for Big Science Data



Real-time analysis and data management



Fast implementations on latest computers

Computing Facilities

Conclusions

- **Exciting time for data-intensive science!**
- **NERSC is in a leadership position**
  - Current: Cori Phase I, Burst Buffer, Policies, Software
  - Future: Cori Phase II, N-9, Superfacility
- **We need your input**
  - Please talk to us about your science problems
  - World-class team is looking forward to working with you

Questions? Comments?

Contact: prabhat@lbl.gov