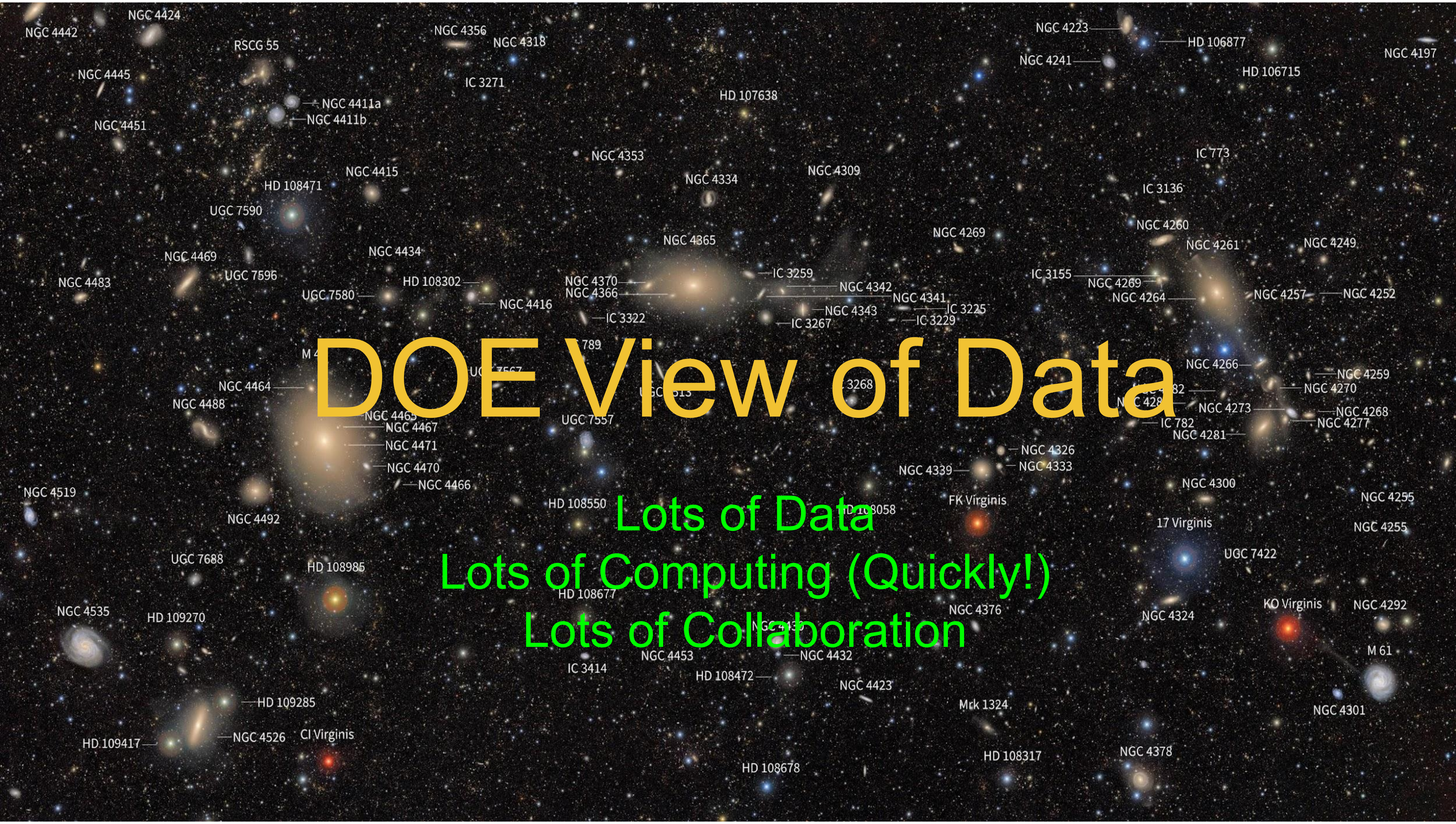
DOE View of Data



Lots of Data

Lots of Computing (Quickly!)

Lots of Collaboration

About Myself

Senior Science Advisor, National Energy Scientific Computing Center (NERSC) (Retired)
Lawrence Berkeley National Laboratory, Berkeley CA

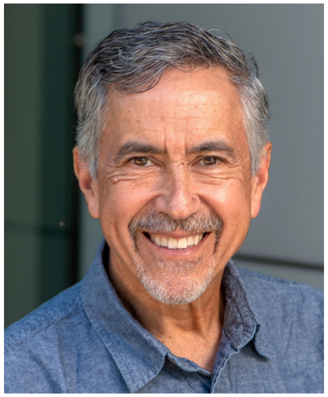
Past NERSC Roles

High Performance Computing Department Head
User Services Group Lead
HPC Consultant, NERSC User Services Group

Education & Experience

Ph.D. in Physics, University of Illinois at Urbana-Champaign
B.S. in Physics, University of Florida
National Research Council Postdoc, Space Sciences, NASA-Ames Research Center
Sportswriter, Editor, and Columnist for Gainesville Sun daily newspaper

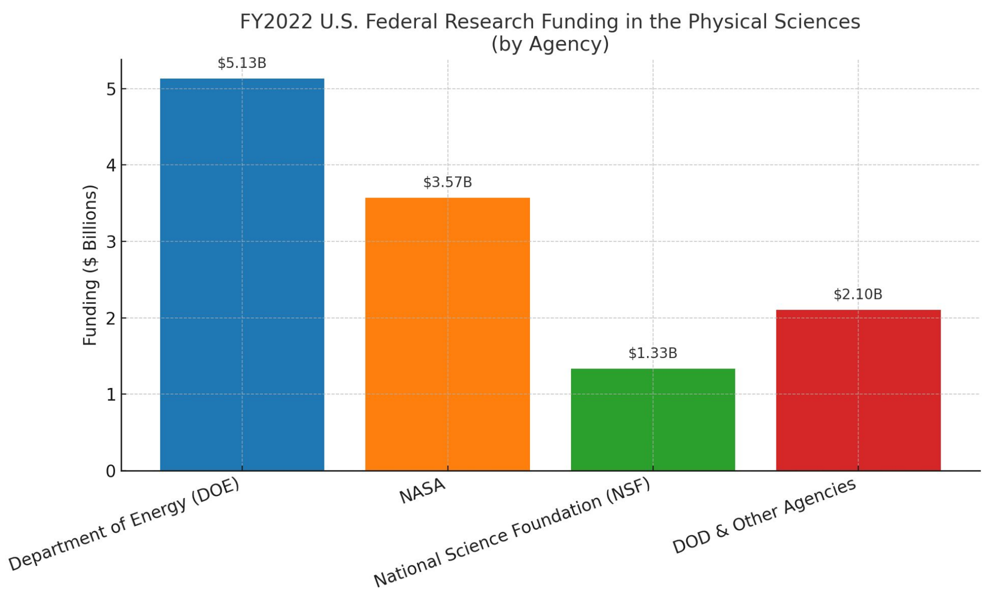
My son attended and my daughter is currently at the local Community College and it's been an amazing and irreplaceable experience for both of them.



Thesis: Stellar and Gas Dynamics in Colliding Ring Galaxies

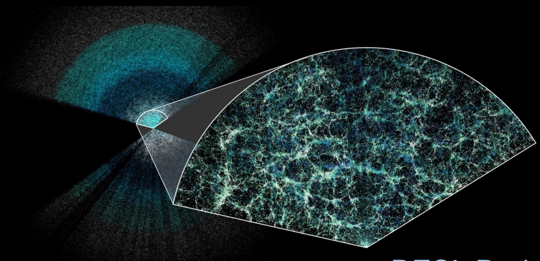
- NCSA User 1986-1993
- Led NERSC & DOE Exascale Requirements Reviews
- PEARC Steering Committee
- Past President of IBM SciCOMP user group and Intel eXtreme Performance User Group (IXPUG)
- DOE Exascale Computing Project, Hardware and Integration Director

The US Department of Energy (DOE) is the Largest Funder of U.S. Physical Science & Engineering

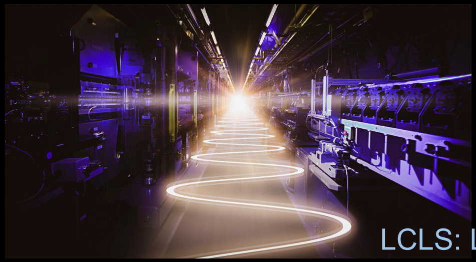


- The DOE National Laboratories are the largest scientific research system in the world
- The Laboratories are the outgrowth of scientific research funded by U.S. government during WWII
- DOE Labs work on **big science:** Multidisciplinary teams solving large-scale, complex R&D challenges
- From basic to applied research in the national interest

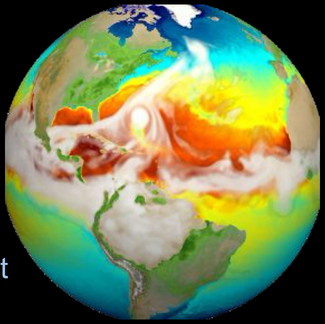
Multi-Petabyte Datasets Stored at NERSC



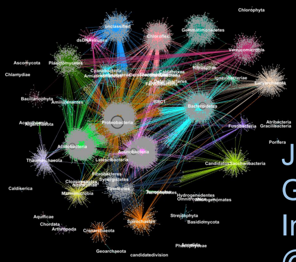
DESI: Dark Energy Spectroscopic Instrument



LCLS: Linac Coherent Light Source @SLAC

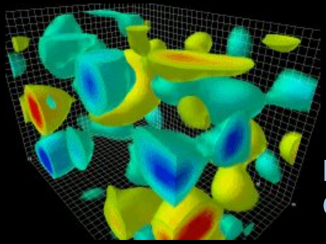


Earth Systems Models



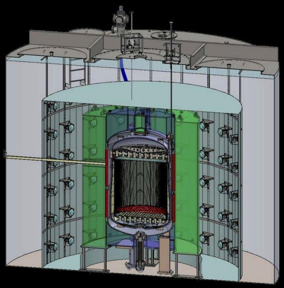
JGI: Joint Genome Institute @Berkeley Lab

DUNE: Deep Underground Neutrino Experiment

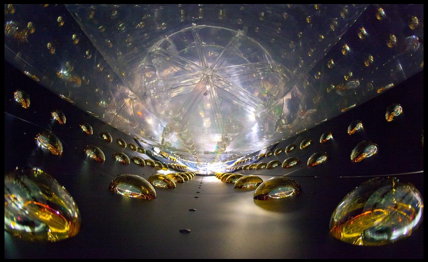


Lattice QCD

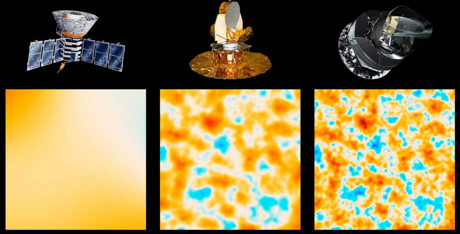
LSST / Rubin Telescope



LZ Dark Matter Experiment



Daya Bay Reactor Neutrino Experiment



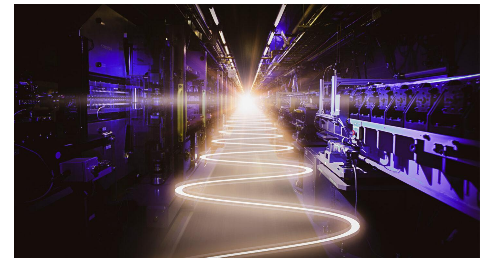
Cosmic Microwave Background Data and Mock Catalogs

From DOE Exascale Requirement Reviews (2017)

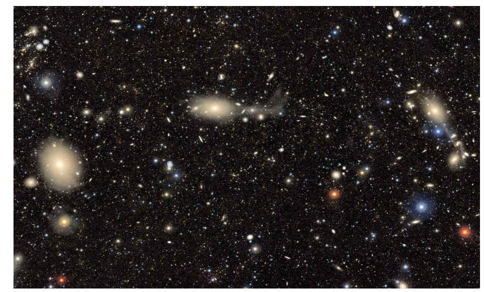
- Scientists need long-term storage space and support for analysis tools. The actions of performing analyses of big data sets and drawing inferences based on these data are revolutionizing many fields. New approaches are needed for analyzing large datasets including advanced statistics and machine learning.
- Workflows in both simulation and analysis are becoming more complex, and they need to be accommodated on HPC systems. This complexity is often related to needs for data movement, including over the wide area network. Scientists at experimental facilities want to use HPC to help guide experiments in real time, which requires co-scheduling between ASCR facilities and facilities from other DOE offices.
- The complexity, volume, and rapidity of data from experiments and simulation requires data management, archiving, and curation well beyond what is in common practice today. Scientists are looking to ASCR for help in this area.
- As science increasingly becomes a community effort, the need to share, transfer, and access data at remote sites becomes more important. Large scientific projects no longer work in isolation.
- The input/output capabilities of large HPC systems need to scale with their computational capability, and sometimes grow faster. Simulations cannot spend excessive time blocking on I/O, and data read/write rates can be the primary factor that limits performance of data analysis pipelines.

Example Use Cases

- LCLS - II: 100 GB/s in 2025. Ephemeral users, real-time feedback, scalable data analytics



- LSST - DESC: Data rates ~10 TB / day. Complicated analytics depend on fast access to large external DBs, next day turnaround, publish catalogs for general public

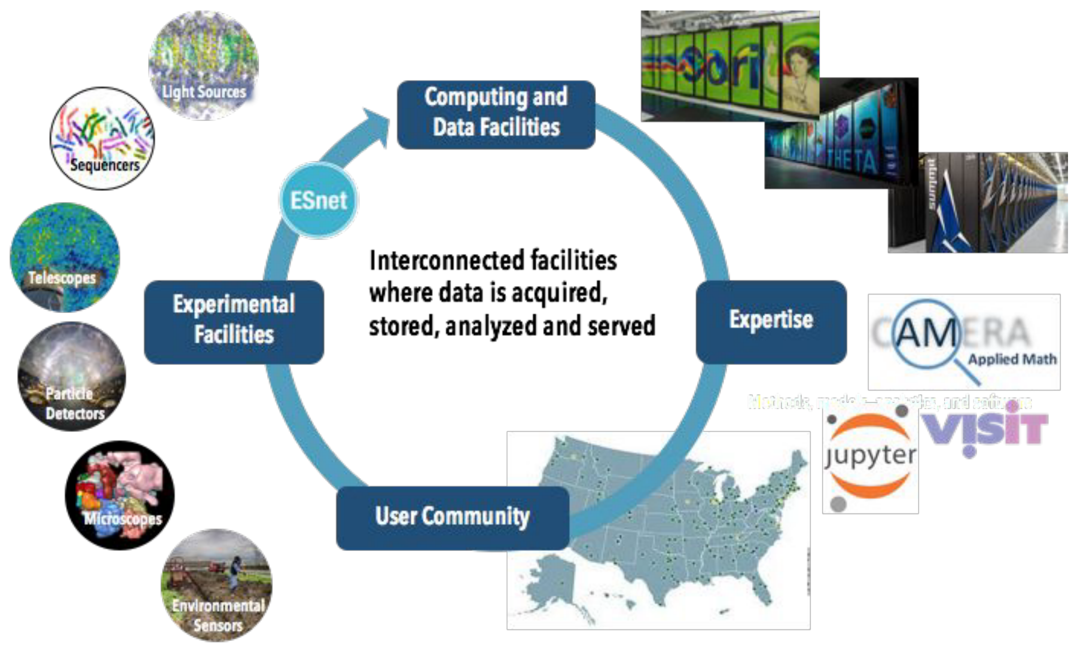


- JGI: Data rates: ~1 TB / day. Complex pipelines (gateways, DBs), unique IO pattern, fine-grained sharing control, large scale memory- and CPU-intensive analytics

JGI
JOINT GENOME INSTITUTE
DEPARTMENT OF ENERGY

The Superfacility Model at LBL: Integrating Experimental, Computational and Networking Facilities for Reproducible Science

- The Superfacility project (2018 - 2021) kick-started work at LBL to support experiment science
- Large scale computing and storage resources
- Reusable building blocks for experimental scientists to build pipelines
- Scalable infrastructure to launch services
- Expertise on how to optimize pipelines



SuperFacility API

- **Meets a critical need; automation is no longer optional**
  - Unattended operation; minimizing HITL
  - Track/submit large number of jobs
  - Interface with collaborations, workflows and machines
- **NERSC becomes "machine readable"**
  - Enables easier creation of UIs, portals, etc.
  - Allows integration with control/analysis software
- **Less DIY: simpler, standardized tooling (Python, etc)**
  - Stable refactor target for established projects or easier on-ramp for new ones
  - Contribute to HPC interface standards for portability
  - Authentication and security models

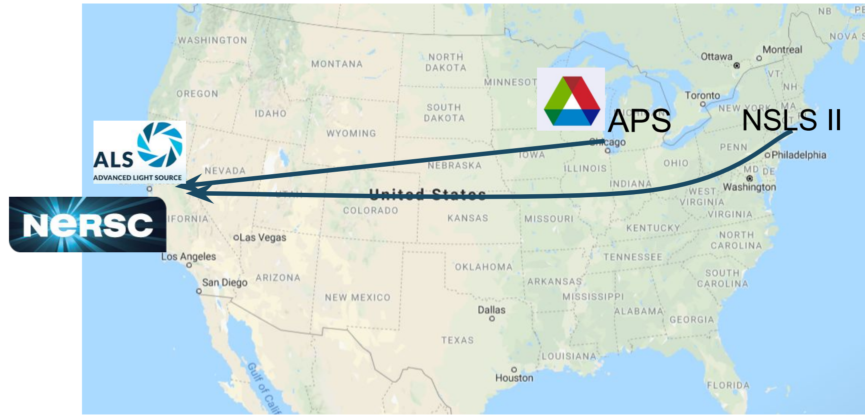
Drivers:

- **Complex workflows**
- **Data-driven projects**
- **Real-time compute and streaming data from instruments**
- **Automation**

Model Case

Experiments at ext. facilities use high frame rate 2D detectors for their science.

Hosting data & compute on site has become increasingly demanding.



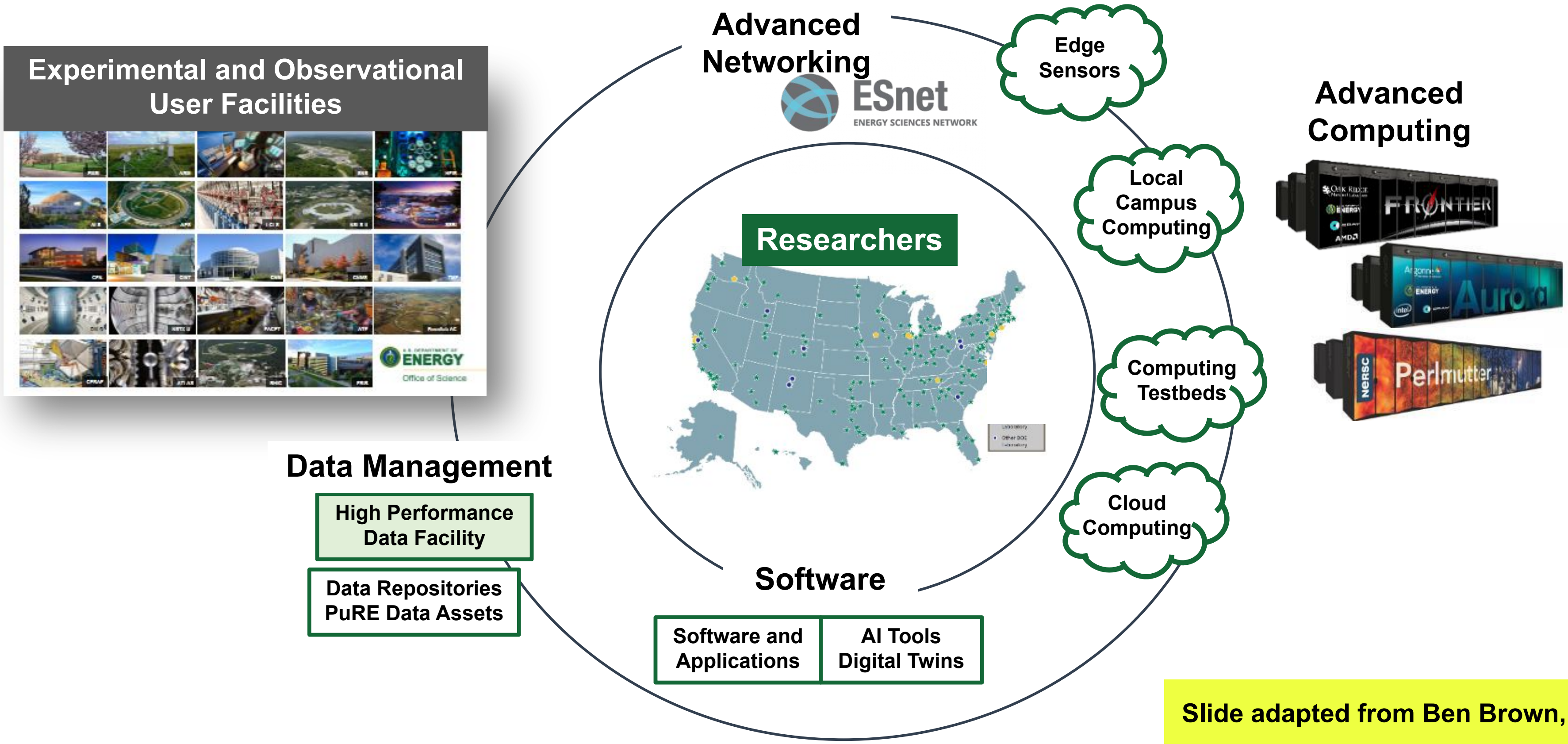
Requirements

- Planning (HPC as reliable partner)
  - machine-readable status
- Resiliency (needs failover)
  - compatible interfaces
- Realtime (can't wait in queue)
  - workflow endpoint
- Services (portals, data, db)

1. Plan / Check availability of NERSC resource for experiment.
   - check status / accounts
2. Get raw data to NERSC, when experiment is live.
   - move data
3. Start analysis job quasi synchronous with data
   - submit job / monitor job
4. Gather feedback, ideally immediate.
   - download / execute command
5. Move data and results to archive after analysis.
   - move data

DOE's Integrated Research Infrastructure (IRI) Vision:

To empower researchers to meld DOE's world-class research tools, infrastructure, and user facilities seamlessly and securely in novel ways to radically accelerate discovery and innovation



Slide adapted from Ben Brown, ASCR

The IRI Architectural Blueprint Activity collated and categorized the many challenges scientists face in building workflows integrated across DOE resources.

The IRI Framework comprises:

> **3 IRI Science Patterns** represent integrated science use cases across DOE science domains.

Time-Sensitive Patterns	Data-Integration Patterns	Long Campaign Patterns

> **6 IRI Practice Areas** represent critical topics that require close coordination to realize and sustain a thriving IRI ecosystem across DOE institutions.

Workflows, Interfaces & Automation	Scientific Data Lifecycle	User Experience
Cybersecurity & Federated Access	Portable/Scalable Solutions	Resource Co-Operations

DOE Announces High Performance Data Facility

A new scientific user facility specializing in advanced infrastructure for data-intensive science.

~$300 M Project

Hub and Spoke model lead by Jefferson Lab in partnership with Berkeley Lab

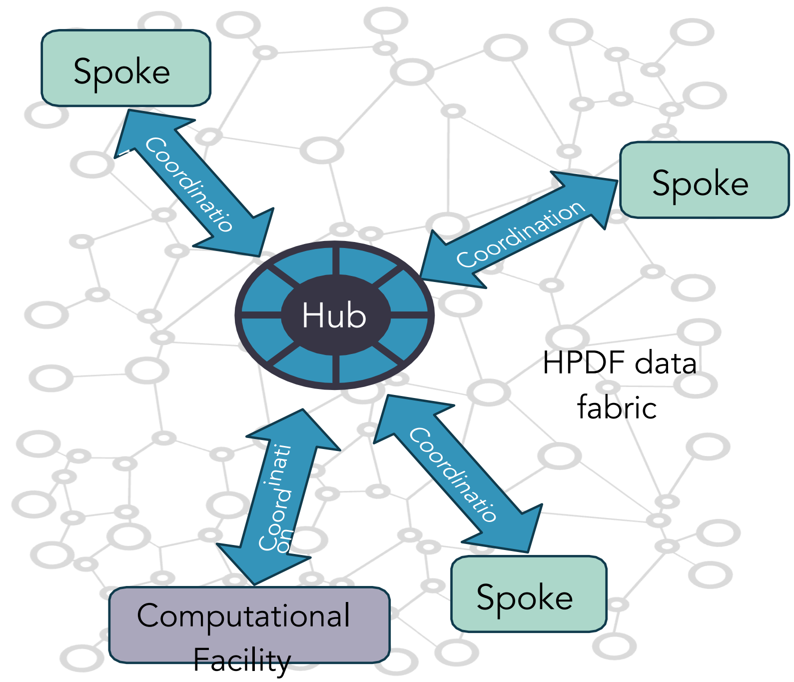
Foundation for advancing IRI vision

Provide leadership in the stewardship of the scientific data lifecycle

High Performance Data Facility — A Distributed Facility

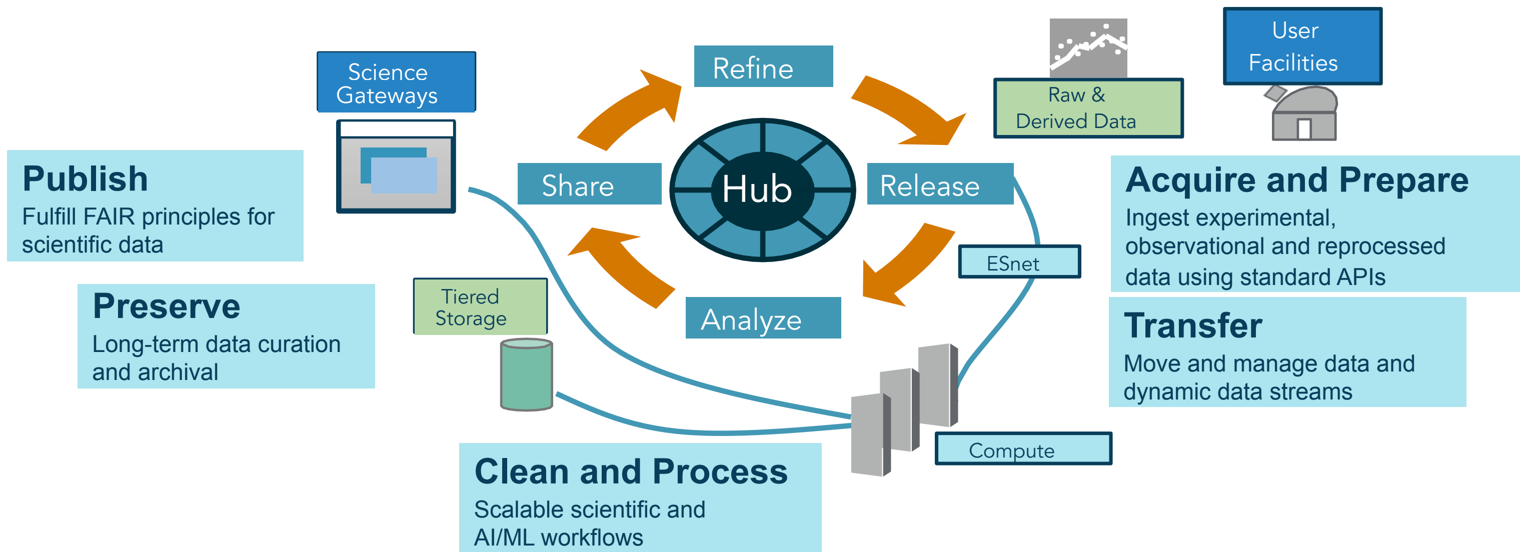
Concept: HPDF is a distributed facility with a hub and spoke architecture.

- **Hub.** Data-centric infrastructure with high availability and performance, as well as geographically and operationally resilient active-active failover.
- **Spokes.** Distributed data-centric infrastructure to enhance HPDF access and support for science users and integrate distributed computing or storage resources.
- **Integration and Services.** Orchestration hardware, software, and services for data movement, storage and retrieval, and science workflow automation. These will use a mesh data fabric building on ESnet6 capabilities.

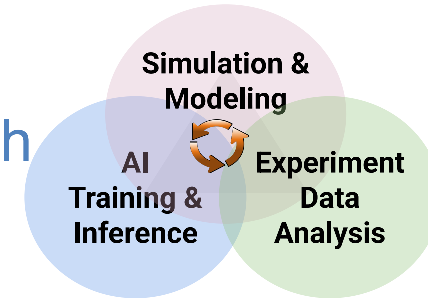


HPDF will Support Data Lifecycle Management

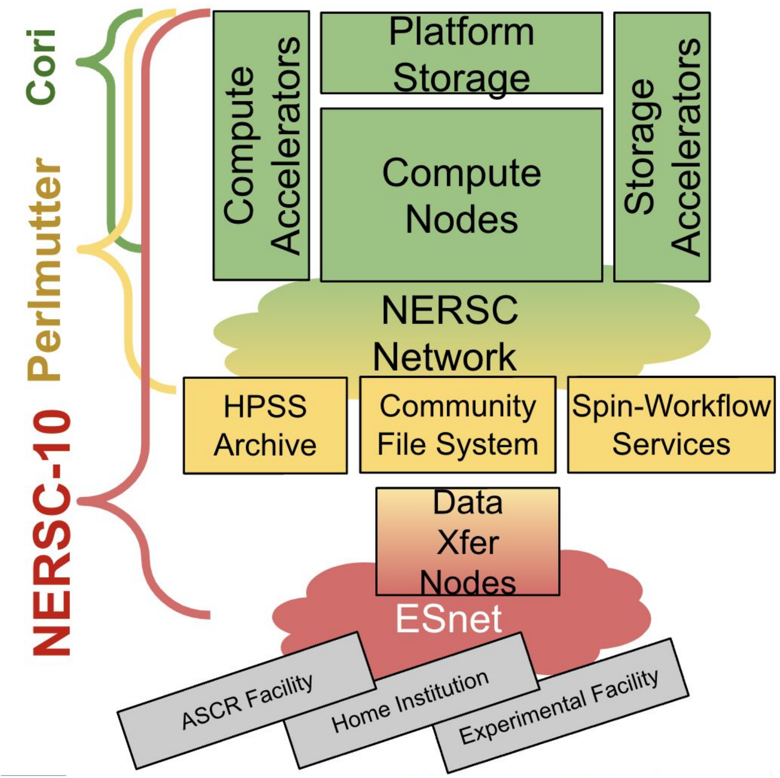
Data science requires curated and annotated data that adheres to FAIR principles, and data reuse will be a metric for HPDF. Office of Scientific and Technical Information (OSTI) services will complement HPDF to provide full life cycle coverage.



NERSC's Next System - Doudna - Will Emphasize Support for Complex Simulation and Data Analysis Workflows at High Performance



- **Quality of Service**: computation, storage and networking enables response-time plus utilization.
- **Seamlessness**: tight integration of system. components enables high performance workflows.
- **Programmability**: APIs manage data, execute code, and interact with system resources.
- **Orchestration**: coordinates resource management across domains.
- **Portability:** Modular workflow execution across IRI sites.
- **Security:** authentication, authorization and auditing.



The IRI Architectural Blueprint Activity collated and categorized the many challenges scientists face in building workflows integrated across DOE resources.

The IRI Framework comprises:

> **3 IRI Science Patterns** represent integrated science use cases across DOE science domains.

Time-Sensitive Patterns	Data-Integration Patterns	Long Campaign Patterns

> **6 IRI Practice Areas** represent critical topics that require close coordination to realize and sustain a thriving IRI ecosystem across DOE institutions.

Workflows, Interfaces & Automation	Scientific Data Lifecycle	User Experience
Cybersecurity & Federated Access	Portable/Scalable Solutions	Resource Co-Operations

IRI is influencing major Infrastructure investments

- The **NERSC-10 RFP** explicitly asks for IRI-relevant features (eg end-to-end QoS, QoS Storage system, APIs, workflow portability)
- The **OLCF-6 RFP** asks for increased bandwidth and connectivity in and around the system to extend Leadership capabilities to enable new workflows between facilities
- The **ALCF-4 RFP** will also feature IRI prominently

KPPs , benchmarks and readiness programs are explicitly designed for IRI

- HPDF was conceived as a data focused component of the IRI ecosystem; we intend to provide a full range of services to support data driven IRI use cases.
- ESnet 6/7 is designed with IRI in mind
  - Fully integrated network automation
  - High fidelity traffic monitoring at scale
  - Network service composability
  - In-network compute and storage