NERSC@50: Continually Changing

Sudip Dosanjh
Director, NERSC

ASCAC

January 17, 2025

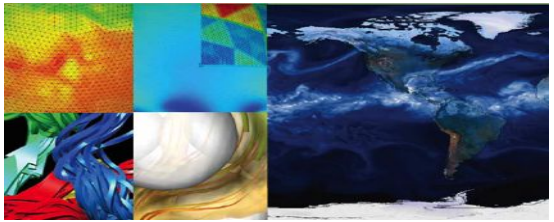
Highlights

- ASCR update by Ceren Susut
- Science Highlights – past accomplishments and future challenges
- Panel sessions on AI, IRI, Superfacility, Quantum Computing, Workforce Development

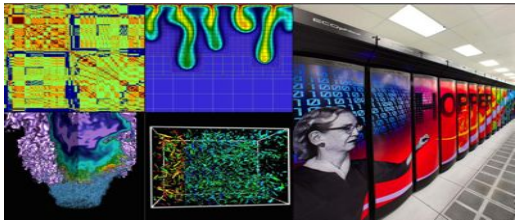
NERSC: Mission HPC for DOE Office of Science Research

U.S. DEPARTMENT OF ENERGY | Office of Science

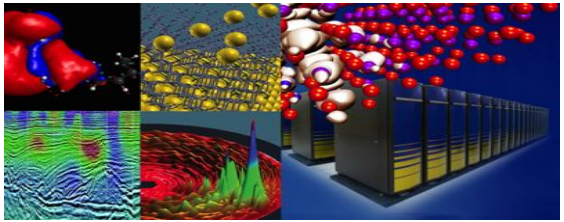
Largest funder of physical science research in the U.S.



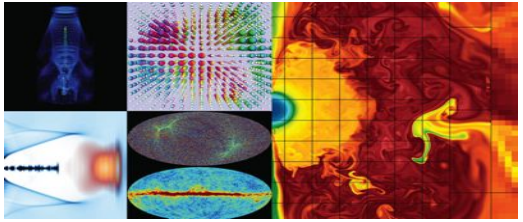
Biological and Environmental Research



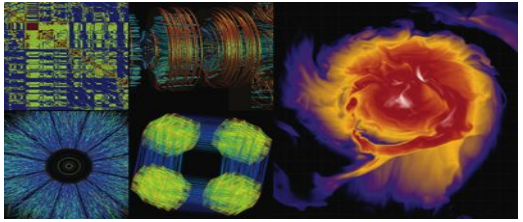
Computing



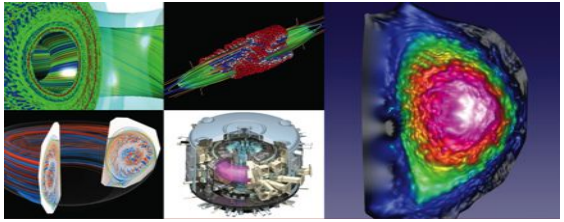
Basic Energy Sciences



High Energy Physics



Nuclear Physics



Fusion Energy, Plasma Physics

NERSC has a very broad user base

2023 NERSC USERS ACROSS US AND WORLD

50 States, Washington D.C. & Puerto Rico

46 Countries

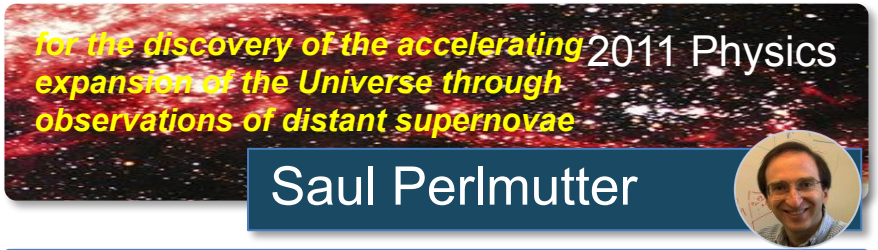
~10,000 Annual Users from **~800** Institutions + National Labs

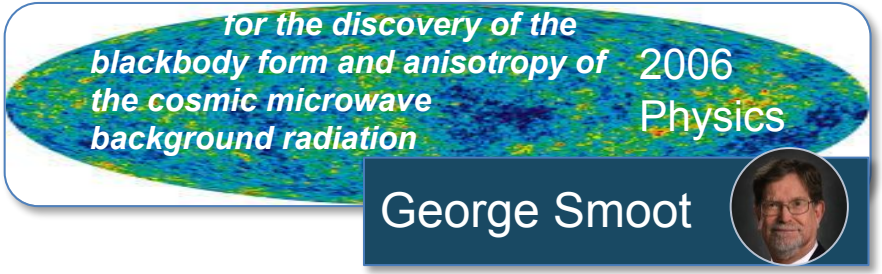
- 30% Graduate Students
- 18% Postdoctoral Fellows
- 16% Staff Scientists
- 14% University Faculty
- 9% Undergraduate Students
- 6% Professional Staff

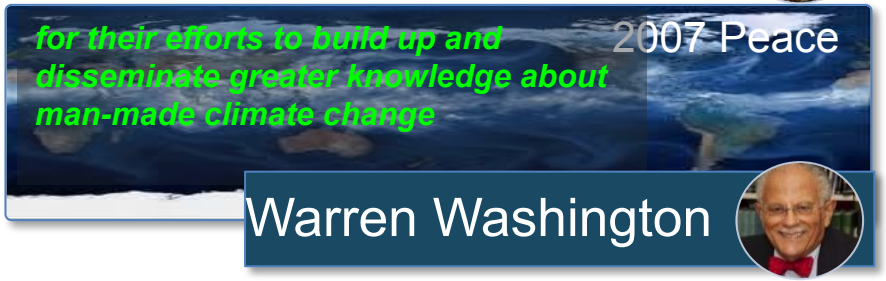
- 60% Universities
- 28% DOE Labs
- 5% Other Government Labs
- 4% Industry
- 1% Small Businesses
- <1% Private Labs

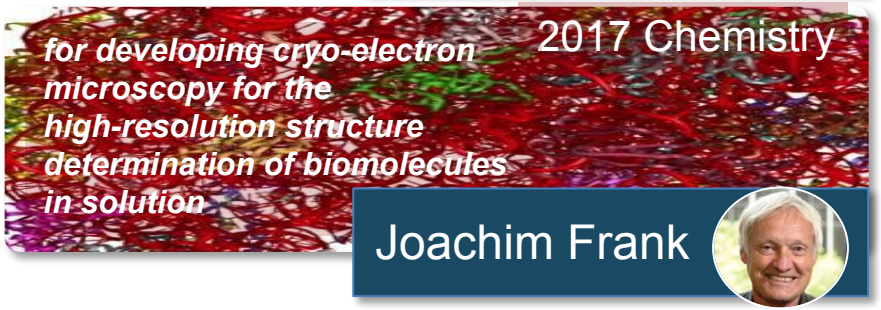
CDC 6600 at LLNL 1974

Nobel-Prize Winning Users







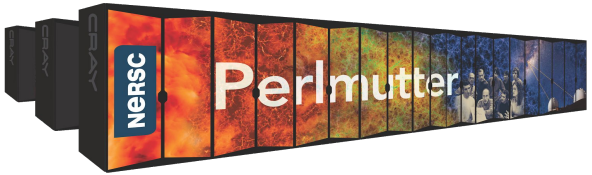


Nobel Prize in Chemistry 2024

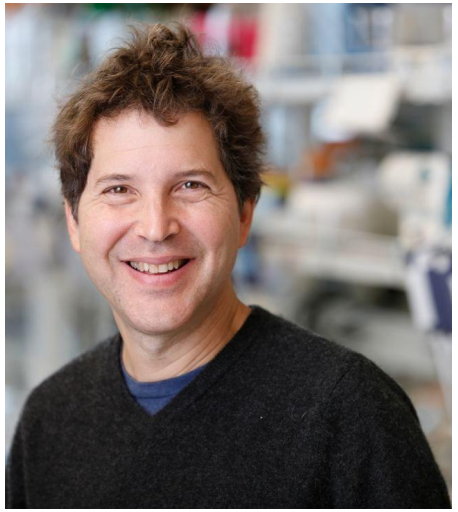
"The Nobel Prize in Chemistry 2024 is about proteins, life's ingenious chemical tools. David Baker has succeeded with the almost impossible feat of building entirely new kinds of proteins." - Nobel committee press release

David Baker and his team were among the first to have early access to Perlmutter when it arrived in 2021.

The team has used 1.5 Million GPU hours on Perlmutter and continues today.



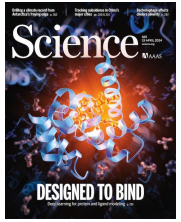
"Perlmutter is an amazing fit for our current computing needs! This is truly a game changer!!" - NERSC PI David Baker, 2021



David Baker is professor of biochemistry, Howard Hughes Medical Institute investigator, and director of the Institute for Protein Design at the University of Washington. He and his research group use AI and the power of Perlmutter to design proteins.

Recent Work at NERSC

- "*Over the past year, enabled by the extensive resources provided by NERSC, we have made significant progress with both new methods development and application of protein structure prediction, design, and interaction prediction to reveal new biological insights.*"
- "*We have been able to predict human protein-protein interactions at a proteome-wide scale. While we had previously predicted protein-protein interactions in yeast and bacteria, these previous methods were not powerful enough to predict the considerably more challenging and larger scale of the human proteome.*"



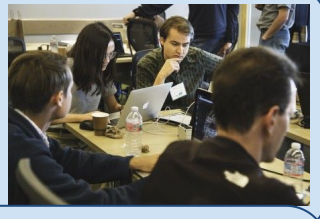
8 NERSC acknowledgements

Nature (4)
Science (3)
Nature Structural & Molecular Biology (1)

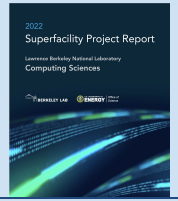
Learn more: "*Prediction and Design of Protein Interactions*", David Baker, NERSC 50th Seminar Series talk, July 3, 2024
https://go.lbl.gov/Baker-NERSC-Seminar

Success Is Depth and Breadth of Scientific Impact

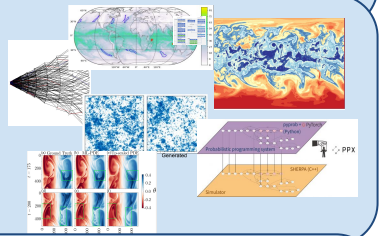
NERSC Science Acceleration Program (NESAP)



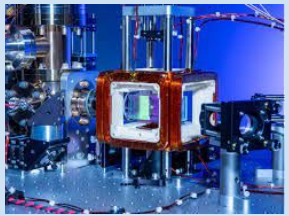
The Superfacility project



AI4Science @ NERSC



QIS @ NERSC

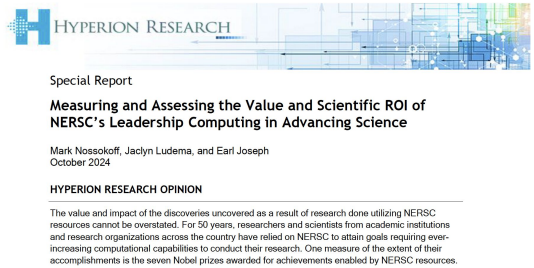


NERSC has been acknowledged in 11,173 refereed scientific publications & high profile journals since 2020

- Nature [59]
- Nature Family of Journals [587]
- Proc. of the National Academy of Sciences [100]
- Science [43]
- Monthly Notices of the Royal Astron. Society [498]
- Physical Review* [1,551]
- Astrophysical Journal [428]
- Physics of Plasmas [222]

Assessing NERSC's Value and Scientific Return

- NERSC commissioned Hyperion Research to measure and convey the impact, value, and scientific return on research using the company's evaluation framework, which was developed with DOE support.
- The process included 42 interviews with leading researchers across all Office of Science domains and reviews of literature and reports.



HYPERION RESEARCH

Special Report

Measuring and Assessing the Value and Scientific ROI of NERSC's Leadership Computing in Advancing Science

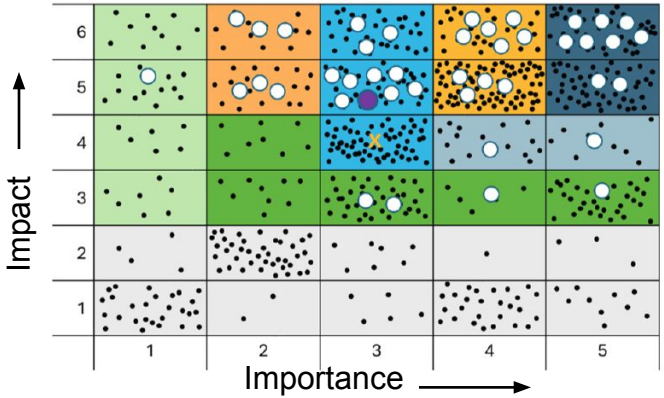
Mark Nossokoff, Jaclyn Ludema, and Earl Joseph
October 2024

HYPERION RESEARCH OPINION

The value and impact of the discoveries uncovered as a result of research done utilizing NERSC resources cannot be overstated. For 50 years, researchers and scientists from academic institutions and research organizations across the country have relied on NERSC to attain goals requiring ever-increasing computational capabilities to conduct their research. One measure of the extent of their accomplishments is the seven Nobel prizes awarded for achievements enabled by NERSC resources.

Key Findings and Summary from Hyperion

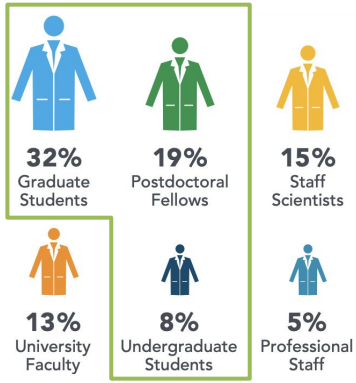
- The value and impact of research performed at NERSC is among the best of global leadership computing sites.
- NERSC facilitates groundbreaking research, which significantly impacts the global scientific community, establishing itself as a leading force in enabling high-value computational research.
- NERSC's infrastructure is indispensable for addressing societal grand challenges in scientific research.
- Superior NERSC processes and personnel are key to accelerating researchers' time to science. NERSC was resoundingly praised by scientists and researchers.



NERSC projects (large circles) compared to 650 HPC projects Hyperion has evaluated.

NERSC is a Catalyst for HPC Workforce Development

- NERSC staff & postdocs
  - NERSC staff training encouraged & provided on critical topics
  - NESAP postdoc program has prepared 35 professionals for career in HPC in past decade; alumni in important roles at HPC vendors (Intel, NVIDIA, AMD), LBNL & other National/International labs
- NERSC users
  - Approximately 60% of users are early career (students/postdocs)
  - World-leading HPC Documentation
  - New user experience is a major thrust of user engagement efforts
  - Extensive training program with broad range of topics offered to NERSC's 10,000 users; 4,000+ training-event participants (2023)
  - NERSC leads all institutions in GPUHackathon.org mentors!
- Future HPC careers
  - NERSC hosted HPC Bootcamp for underrepresented students in collaboration with ALCF/OLCF in 2023
  - Early supporter of Sustainable Research Horizons program to match underrepresented students with research projects at DOE labs; ~15 students placed at NERSC



User Demographics

NERSC is Creating an HPC Workforce

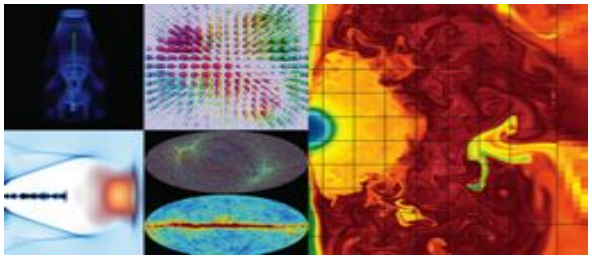
Current NESAP Postdocs

Name	Workflow
Soham Ghosh	Materials Science
Vinicius Mikuni	HEP-ML
Andrew Naylor	U.S. CMS, HEP Mini-apps
Mukul Dave	WarpX
Wenbin Xu	Catalysis, BonDNet
Jared Willard	iNAIADS, FourCastNeXT
Sam Welborn	Superfacility, NCEM
Madan Timalsina	Superfacility, U.S. CMS

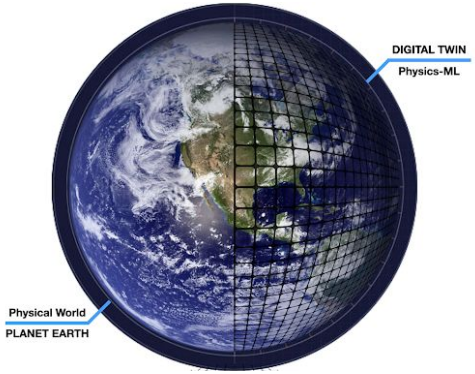
Recent NESAP PostDocs

Name	Current Position
Taylor Barnes	NSF MOLSSI
Andrey Ovsyannikov	Intel
Tuomas Koskela	University of Helsinki
Mathieu Lobet	CEA
Tareq Malas	Meta
Zahra Ronaghi	NVIDIA
Bill Arndt	LBNL / NERSC
Jonathan Madsen	AMD
Michael Rowan	AMD
Kevin Gott	LBNL / NERSC
Muaaz Awan	LBNL / NERSC
Laurie Stephey	CalEPA
Oisin Creaner	Dublin Institute of Advanced Tech.
Ozgur Cekmer	CSIRO
Yunsong Wang	NVIDIA
Yan Zhang	Velodyne Lidar
Hugo Brunie	CEA (France)
Dossay Oryspayev	Brookhaven Lab
Brandon Wood	AI at Meta
Dhruva Kulkarni	LBNL / NERSC
Amanda Dufek	LBNL / NERSC
Raphael Prat	CEA (France)
Jaideep Pathak	NVIDIA
Daniel Margala	LBNL / NERSC
Muhammad Haseeb	NVIDIA
Lipi Gupta	LBNL / NERSC
Nestor Demeure	LBNL / NERSC

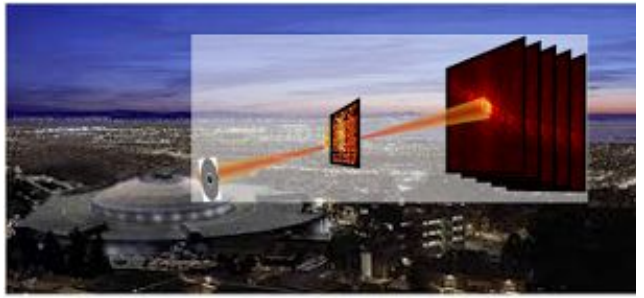
We Accelerate Scientific Discovery for Thousands of Office of Science Users with 3 Advanced Capability Thrusts



Large-scale applications for simulation, modeling and data analysis



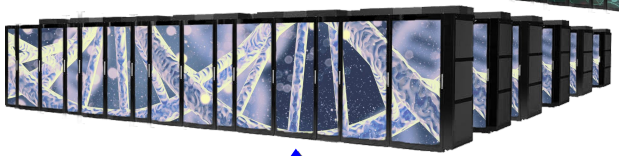
Complex experimental and AI driven workflows



Time-sensitive and interactive computing

The NERSC workload is diverse with growing emphasis on integrated research workflows

Responding to the DOE Mission



NERSC-8: Cori
Manycore CPU architectures

2016

NERSC-9: Perlmutter
CPU and GPU nodes
Expanded Simulation, Learning & Data

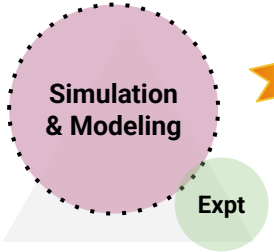
2020

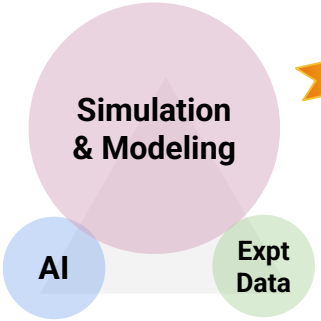
NERSC-10:
Accelerating end-to-end workflows

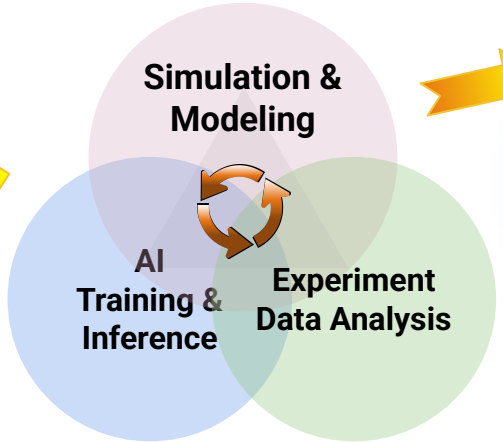
2026

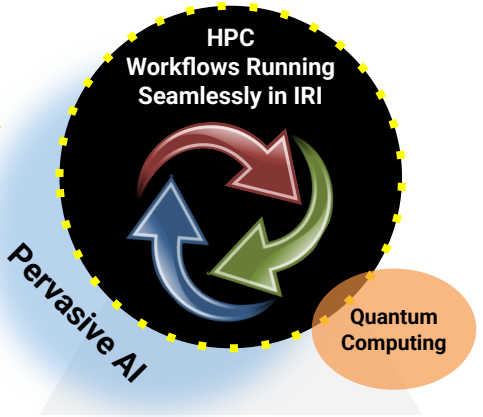
NERSC-11:
Beyond Moore

2030+

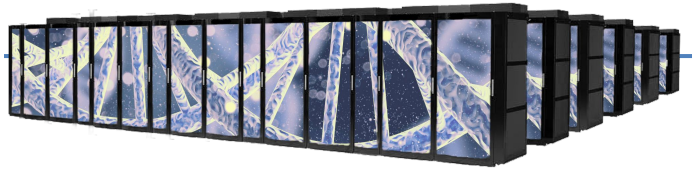








NERSC-10: A Productive System for Complex Workflows



Designed to provide > 10x performance over Perlmutter and support diverse and complex DOE SC workflows.

Workload Optimized Compute Capabilities

- Able to support GPU and CPU portions of complex workflows
- AI and Compute Optimized Partitions

2 Types of Storage Capabilities

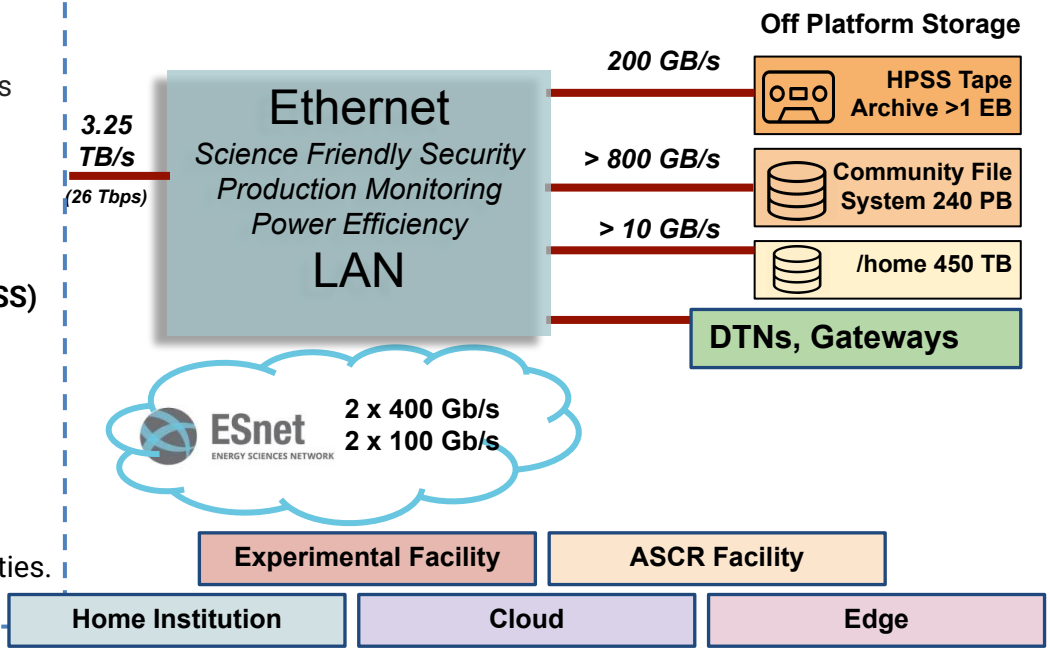
Platform Storage System (PSS)

Quality of Service Storage System (QSS)

Workflow Environment

Workflow Environment Nodes

Reconfigurable to support complex IRI workflows (e.g. batch, compile, Jupyter, cloud-native, data transfer, etc.) with integrated Spin capabilities.



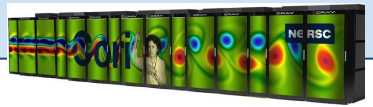
NERSC Exascale Science Applications Program (NESAP) was founded to transition the broad science user base

NESAP Strategy: (1) Form deep partnerships with ~20 Science teams at a time. Directly improve strategic workflows + gain expertise.

(2) Take lessons learned from deep-dives & apply to full NERSC community.

2011-2014
NERSC partners with DOE Researchers on GPU tech exploration via Dirac testbed

2014-2018
NESAP supports NERSC users transitioning to energy-efficient architectures on Cori (KNL)



2019-2020
~40 NESAP teams have access to GPU testbed; 50% of NERSC workload GPU-ready.

2020-2023
Scaling NESAP impact to 10,000 users: 75% of NERSC workload GPU-ready

2024-2028
NESAP program for NERSC-10 focuses on workflows.

NESAP is NERSC's Vehicle for Preparing Users for Current and Future Systems

Collaboration between **NERSC, Strategic Science Teams**, and **Technology Vendors** to prepare science workflows for systems at NERSC

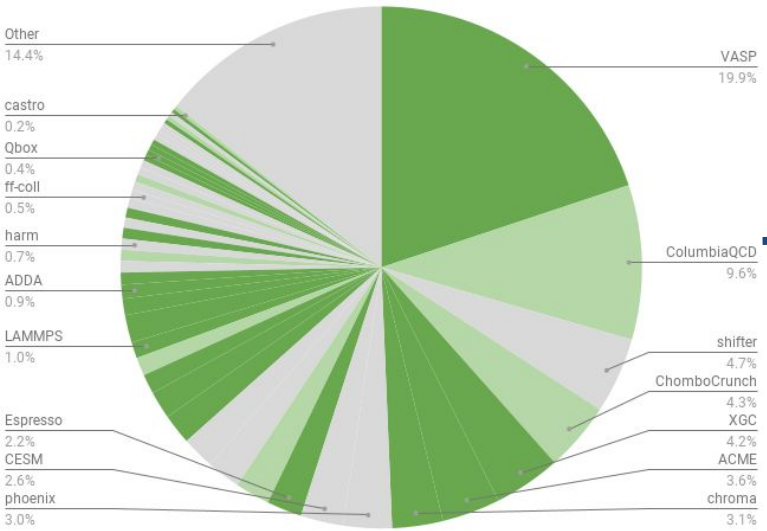
The partnership include:

- Science Application + Technology **CoDesign**
- **Early Access** to Advanced Hardware
- Collaborative **Code Development, Analysis and Optimization**
- Mutual **Staff Embedding, Training**
- Creation of Lessons Learned and Best Practices to be shared with community

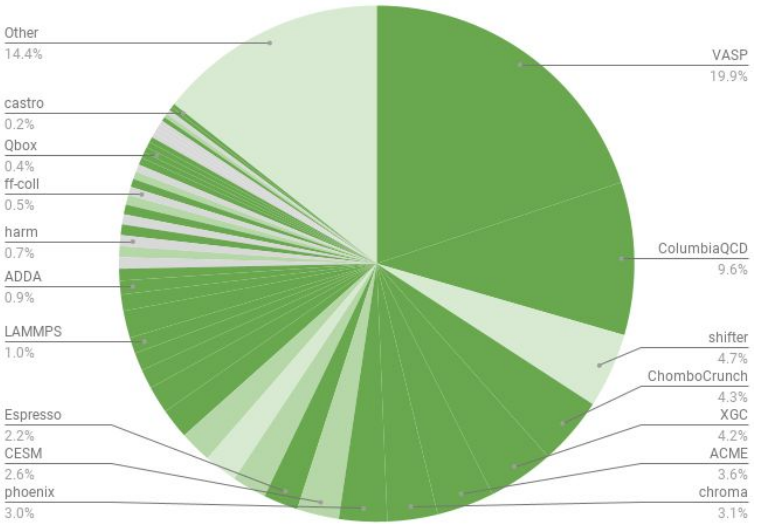
Proven Successful on Cori and Perlmutter

Application	NESAP Category	Perlmutter Speedup
FlowGAN	Learning	89.7
BerkeleyGW	Electronic Structure	78.3
ExaFEL	Data	186.9
WarpX	Particles & Grids	27.2
LAMMPs	Molecular Dynamics	79.1
Chroma	LQCD	56.0
Mean		**73.5**

NESAP is NERSC's Vehicle for Preparing Users for Current and Future Systems

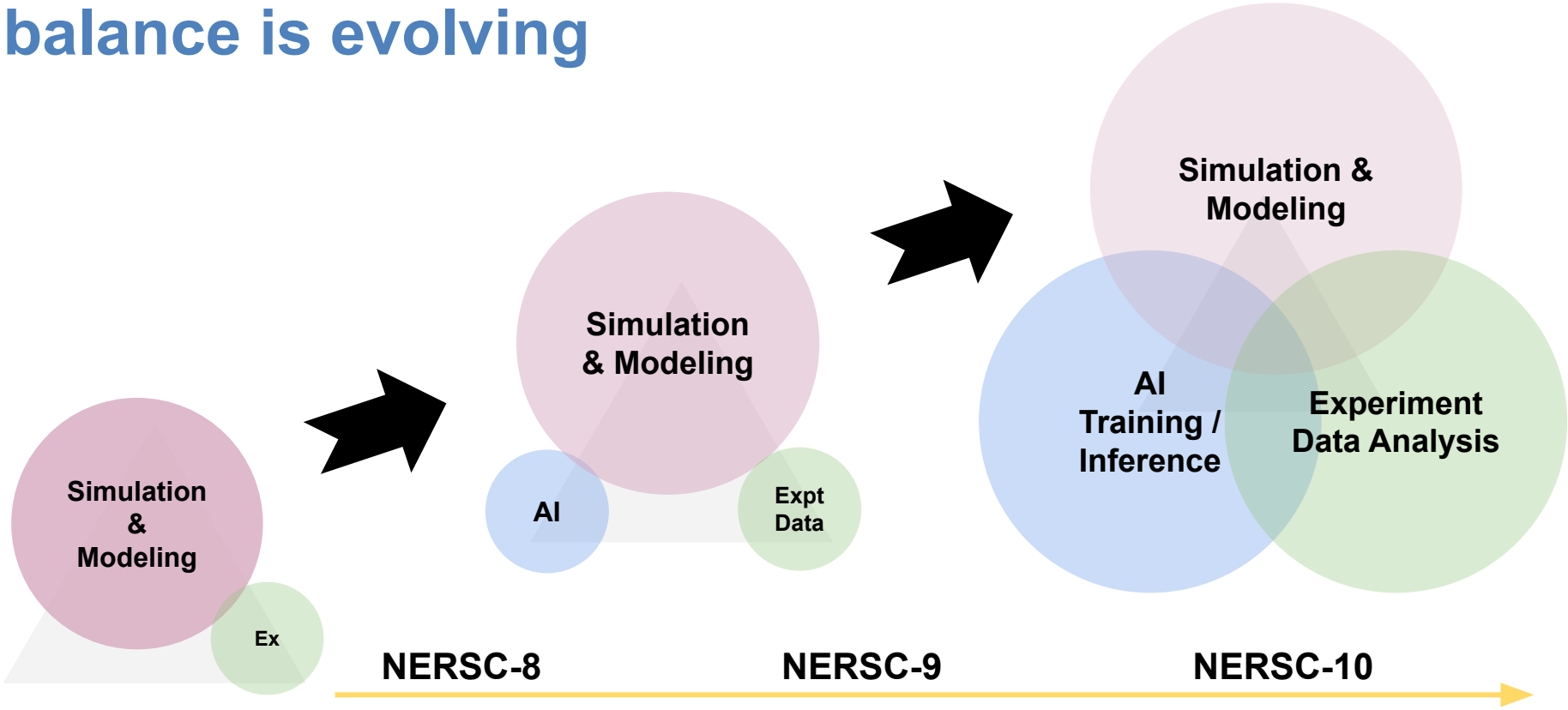


GPU Readiness of NERSC Workload Before NESAP



GPU Readiness of NERSC Workload After NESAP

The HPC facility workload balance is evolving



Broadening Mission Space to support Experimental and Observational Science

The **Superfacility** model at LBL connects experiment and compute facilities with the expertise and community they need for success

- Close collaboration between NERSC, ESnet and CS Research Divisions supports multiple DOE science teams who use automated pipelines to analyze data from remote facilities, at large scale.

<2018
NERSC supports >30 experiment teams. 2016: Cori introduces "data-friendly" features

2018-2021
Superfacility project coordinates work to enable autonomous workflows at NERSC from 8 experiment teams

2020-2023
Scaling Superfacility impact: ~2000 NERSC users leverage Superfacility- enabled tools (not just experimental science!)

2023+
IRI program enables complex workflows that cross institutional boundaries, for DOE science

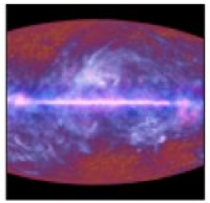
NERSC LAWRENCE BERKELEY LAB

BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY | Office of Science

NERSC supports a large number of users and projects from DOE SC's experimental and observational facilities

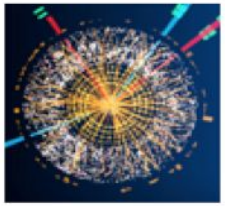
Palomar Transient Factory
Supernova



Planck Satellite
Cosmic Microwave Background Radiation

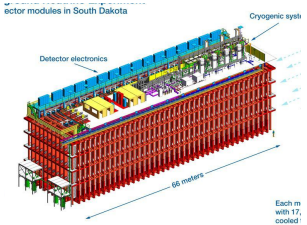
STAR

Star
Particle Physics

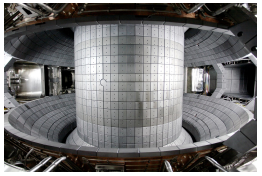


Atlas
Large Hadron Collider

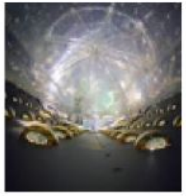
APS



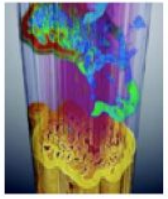
Dune



KStar



Dayabay
Neutrinos



ALS
Light Source

LCLS
Light Source

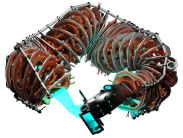
Joint Genome Institute
Bioinformatics

ARM

GLUEX

Katrin

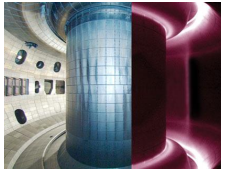
NSLS-II



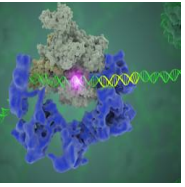
HSX

Majorana

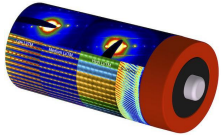
AMERIFLUX



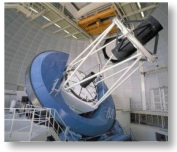
DIII-D



Cryo-EM

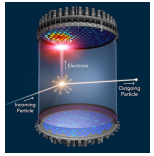


NCEM



DESI

LSST-DESC



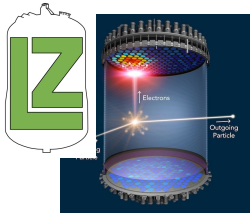
LZ

IceCube

EXO

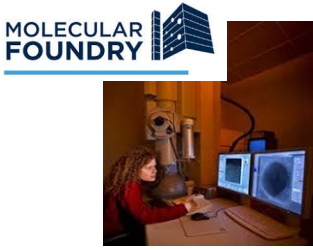
JBEI
Joint BioEnergy Institute

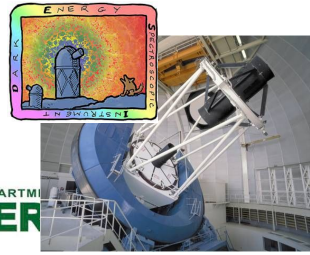
Multiple science teams are using NERSC for superfacility-enabled science, in production



The 3 year Superfacility project kick-started this work, building the base infrastructure and services. We now support **multiple science teams using automated pipelines to analyze data from remote facilities at large scale**, without routine human intervention, using:

- **Real-time** computing support
- Dynamic, high-performance **networking**
- Data management and movement tools, incl. **Globus**
- **API**-driven automation
- HPC-scale notebooks via **Jupyter**
- Authentication using **Federated Identity**
- Container-based edge services supported via **Spin**





Multiple science teams are using NERSC for superfacility-enabled science, in production

A set of 8 initial close science engagements drove this work, but the impact has scaled to benefit all NERSC users

- **Real-time** computing support
- Dynamic, high-performance **networking**
- Data management+movement tools, incl. **Globus**
- Interactive HPC via **Jupyter**
- Container-based edge services supported via **Spin**
- **API** interfaces
- **Federated Identity**/auth
- Collaboration accounts for **automated "robot" access**

>20 science teams use the **realtime** qos to process urgent data

>1500 unique **Jupyter** users per month, similar to number of users who ssh into our systems

>250 users, >85 projects use **Spin**

>40 projects use the NERSC **API**, ~19M logged requests since May 2022 = one request every 2 sec

>1400 users are now logging in with a home lab identity

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

Advanced Networking

Edge Sensors

Advanced Computing

Local Campus Computing

Researchers

Computing Testbeds

Cloud Computing

Data Management

High Performance Data Facility

Data Repositories PuRE Data Assets

Software

Software and Applications

AI Tools Digital Twins

New modes of integrated science

Rapid data analysis and steering of experiments

Novel workflows using multiple user facilities

AI-enabled insight from integrating vast data sources

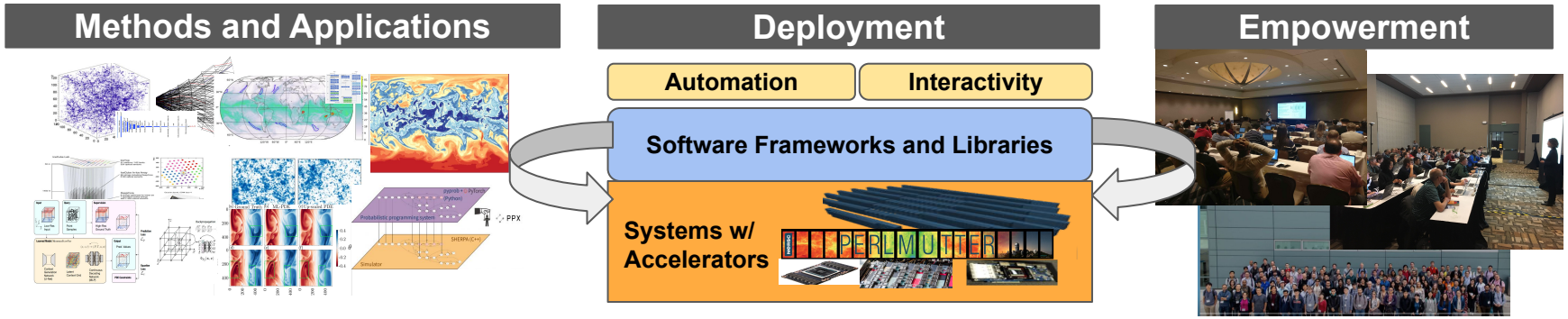
U.S. DEPARTMENT OF ENERGY | Office of Science

Slide adapted from Ben Brown, ASCR

Partnerships with ASCR Ecosystem

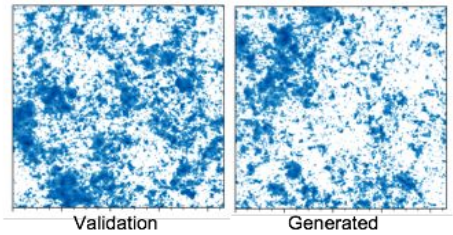
- LCFs
  - Collaborated closely to deploy current generation systems
  - Joint training, hack-a-thons
  - Outreach to underserved communities
  - Coordinated strategy on upcoming procurements
    - frequent communication, as many common requirements as possible
- ESnet
  - Collaborations to deploy new networking technologies and enable workflows
- HPDF
  - Collaborative Effort Between JLab and LBNL
  - Frequent communication with NERSC to define roles, leverage opportunities
- All ASCR Facilities working together on IRI
  - Frequent Topic on Weekly Leadership CallsEstablished 5 Facility IRI Management Council; Debbie Bard is Chair of Leadership group

NERSC Executing on Current AI Strategy

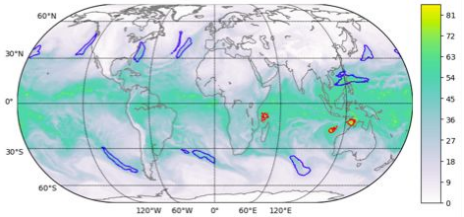


- ***Deploy*** optimized hardware and software systems
  - Work with vendors for optimized AI software
- ***Apply*** AI for science using cutting-edge techniques
  - "NESAP" and strategic projects - leverage lessons learned for scalable impact
- ***Empower*** and develop workforce through seminars, training and schools as well as staff, student intern and postdoctoral programs
  - Over 20 DL@Scale tutorials (e.g. SC18-23), 1000s of total participants

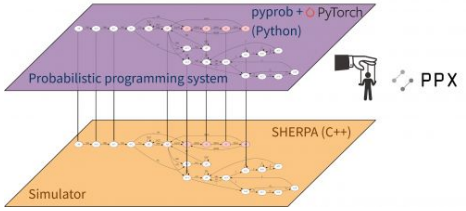
NERSC has Driven Emergence of AI for Science on HPC



First generative DL for science (*CosmoGAN* and *CaloGAN 2017*)

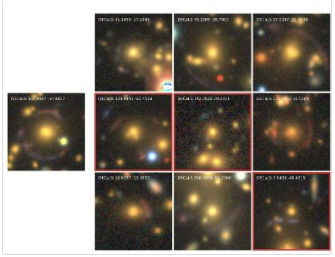


First Exascale DL application (SC18 Gordon Bell *Exascale DL for Climate Analytics*)

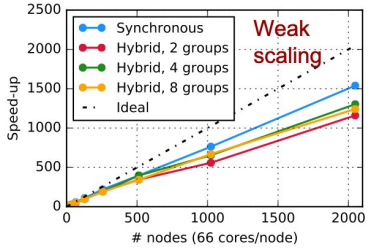


Large Scale Probabilistic Programming (SC19 *Etalumis*)

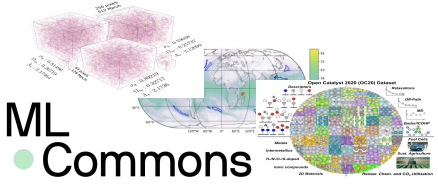
DL@Scale tutorials (e.g. SC18-SC23)



First self-supervised DL in cosmology (*2020*) - *Mining for Strong Gravitational Lenses (2021*)



First 10k node DL application (SC17 *Deep learning at 15PF*)



Founded HPC benchmarks in MLPerf

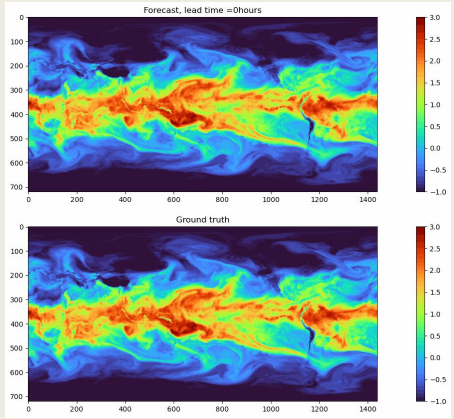
Deep learning for science school

NESAP and Perlmutter Enabled Adoption of Large-scale and Groundbreaking AI

FourCastNet

Pathak et al. 2022 **arXiv:2202.11214**
Collab with Nvidia, Caltech, …

- Forecasts global weather at high-resolution.
- Prediction skill of numerical model; 10000s times faster



Jaideep Pathak
former NERSC
Postdoc now NVIDIA

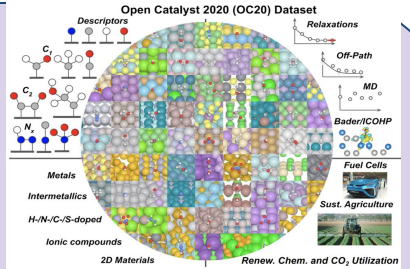
Shashank
Subramanian
Former NERSC
Postdoc now Staff

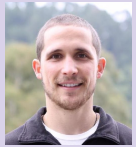
Jared Willard
NERSC Postdoc

CatalysisDL

Chanussot et al. 2021
Collab with CMU, MetaAI, …
arXiv:2010.09990

- NeurIPS 2021-23 Competitions
- Pre-trained models now used with DFT - e.g. FineTuna; AdsorbML





Brandon Wood
former NERSC
Postdoc now Meta AI

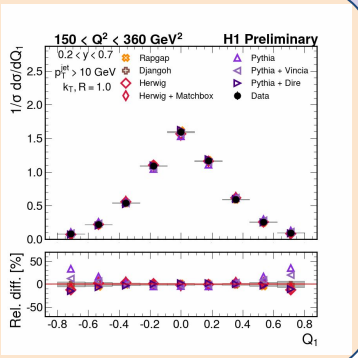
Wenbin Xu
NERSC postdoc

HEP-ML

Collab with LBL Physics division (and H1 Collaboration)

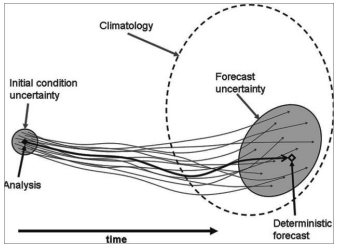
- AI "Unfolding" extracts new physics insights from data
  - Requires Perlmutter for 1000s of UQ runs

Vinicius Mikuni
NERSC Postdoc

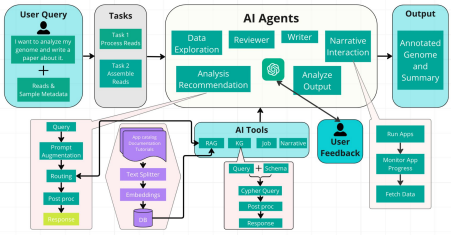


V. Andreev et al. (H1 Collaboration) Phys. Rev. Lett. 128, 132002 (2022)

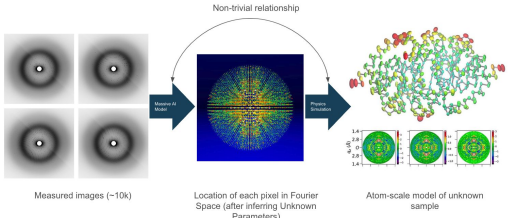
As AI Becomes Pervasive Science will be Transformed



BER: Climate hindcasting with large ensembles



BER: KBase Knowledge Assistant



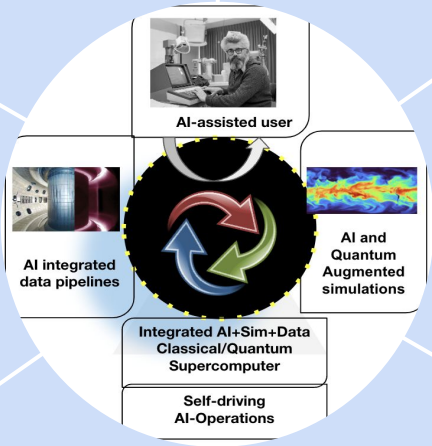
BES: Pixel-level Bragg analysis at LCLS-II

Extreme scale surrogate models

Fast novel experiment design

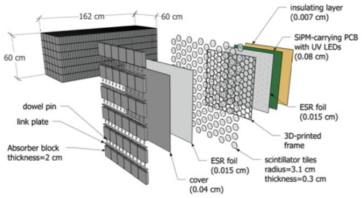
AI knowledge discovery assistants

AI-driven automation

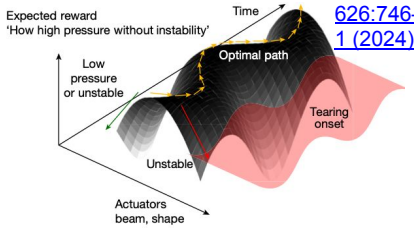


Inference with full science models

Unsupervised detection of novel science

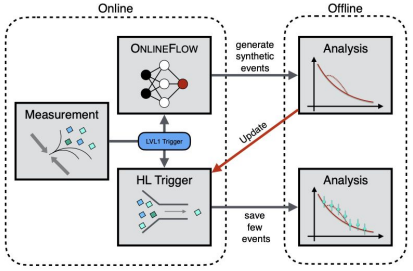


NP: EIC Calorimeter design



(ref: Nature 626:746–751 (2024))

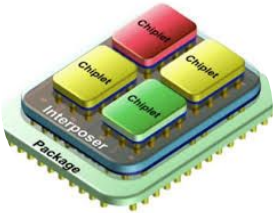
FES: Instability avoidance for ITER



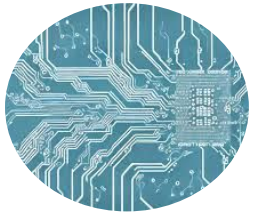
HEP: "Anomaly" detection for HL-LHC
(ref: SciPost Phys. 13, 087 (2022))

End of Moore's Law Opens Up New Approaches to Computing

FPGA



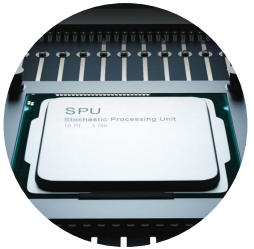
Chiplet



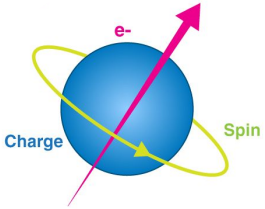
ASIC

Improved energy efficiency through hardware specialization

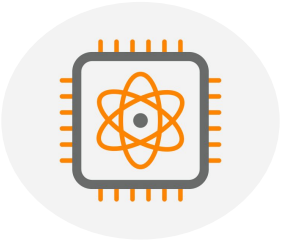
Neuromorphic



Stochastic



Spintronics

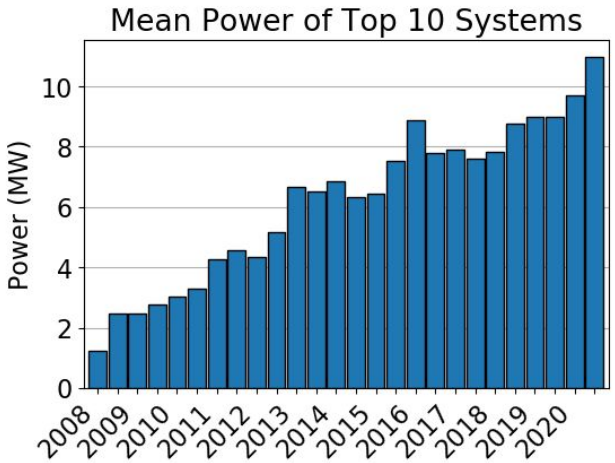


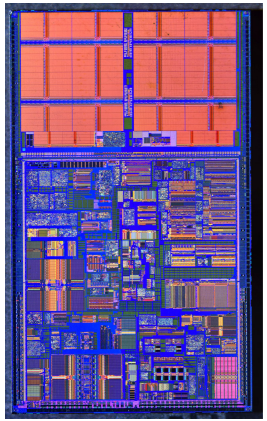
Quantum

New paradigms for computing

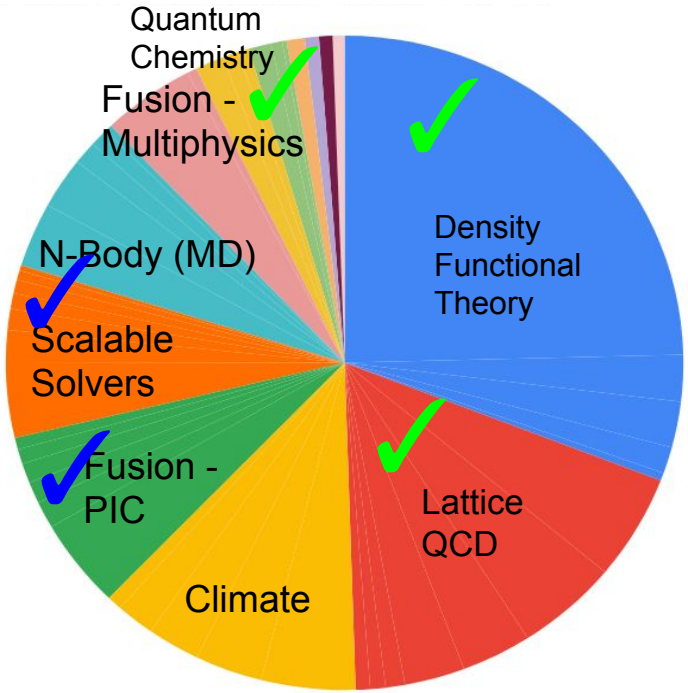
Beyond Moore's Strategy

- Maximize energy efficiency
  - Use every Joule available for science as efficiently as possible
- Deploy best in class specialized computing elements to accelerate components of users workflows
  - Deploy testbeds and collaborate with community to determine best-in-class
- Continue NRE & co-design efforts
  - develop representative benchmarks





Why Quantum Computing at NERSC?



Top Algorithms Run at NERSC

✓ Quantum mechanical problem
Quantum algorithms proposed

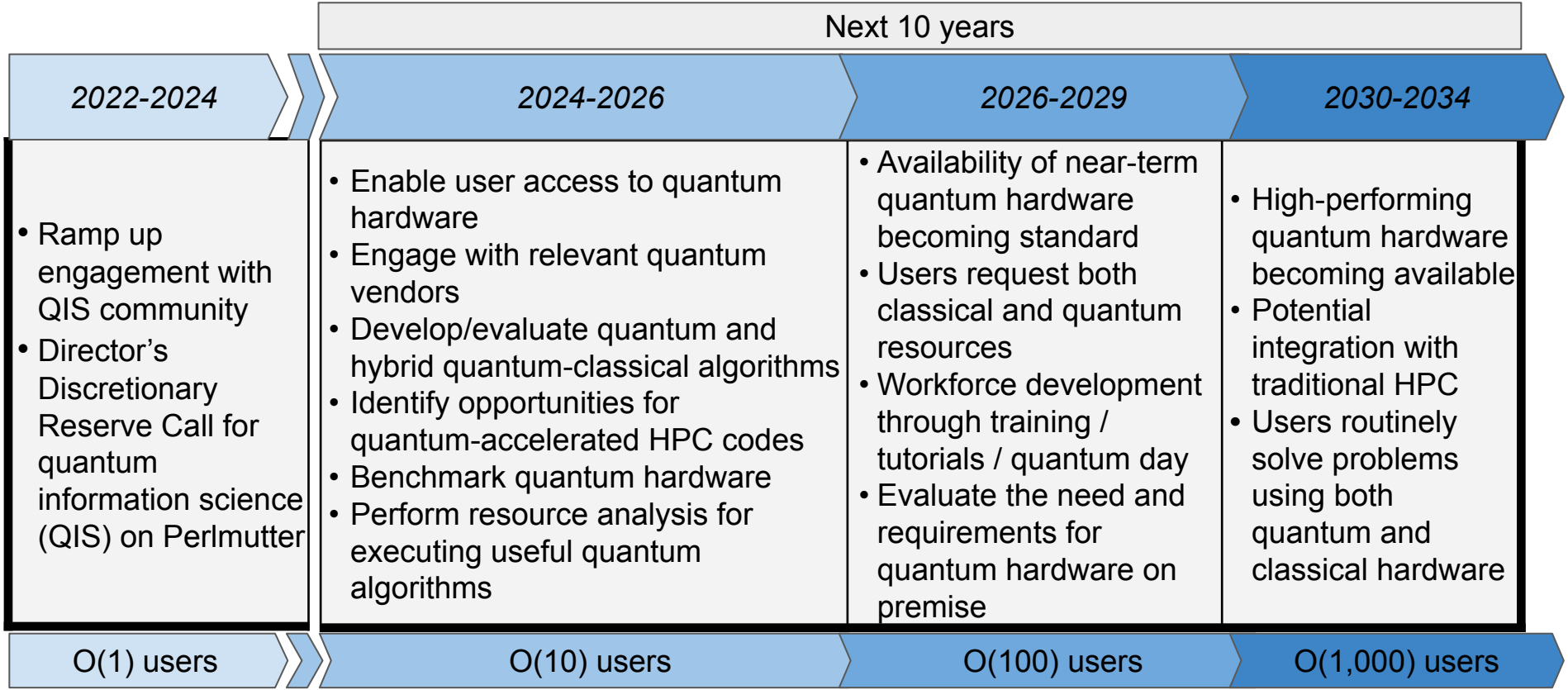
>50% of cycles

✓ non-Quantum mechanical problem
Quantum algorithms proposed

20% of cycles

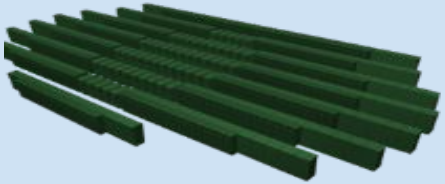
What is **not** on the pie chart? Growth areas for quantum technologies!

NERSC Quantum Computing Roadmap

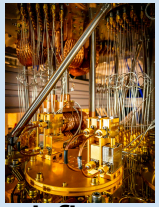


Potential Quantum Hardware in the NERSC-11 Era

NERSC-11

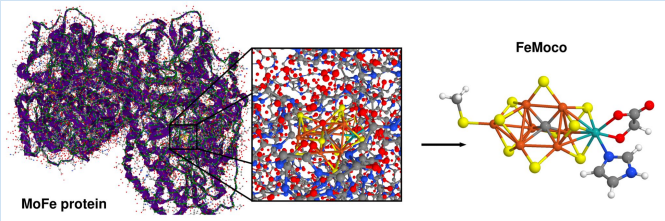


NERSC Quantum-1



Integration for hybrid workflows

Elucidating reaction mechanisms for strongly correlated molecular systems

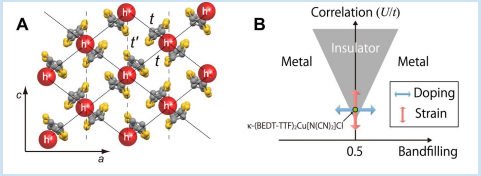


FeMoco, *PNAS*, 2017

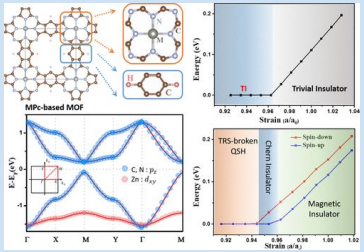
NERSC Quantum-1

Loose coupling for fully quantum-native apps

Control and exploit electronic interaction for design of bulk materials with novel functionality



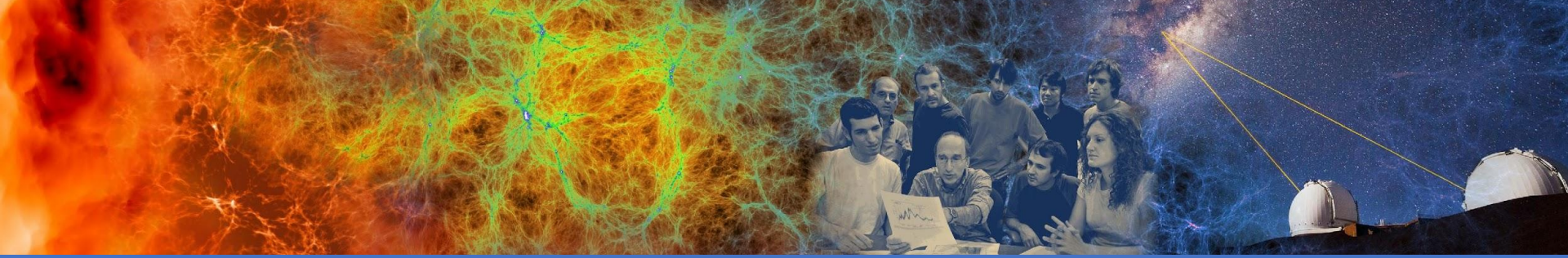
Organic superconductors, *Science*, 2021



Lieb lattice, *Nature Comm*. 2020

Final Thoughts

- NERSC has deployed numerous innovative supercomputers
- Result has been many scientific breakthroughs and publications
- The next 10-20 years will see great change
- Expansion of the mission space (complex coupled workflows, Superfacility, IRI, AI)
- End of Moore's Law
- An opportunity to use AI to revolutionize science and operations



Questions?