

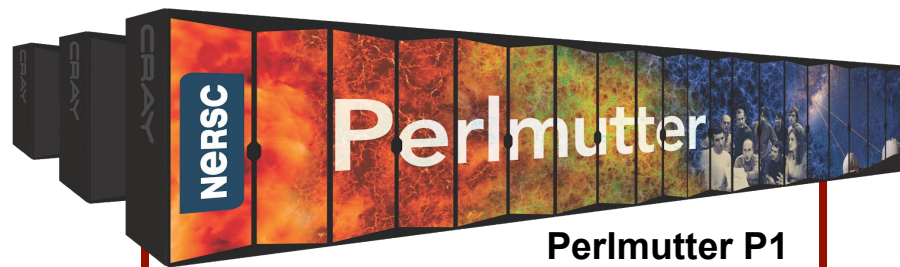
NERSC update Workshop for the USATLAS-USCMS HPC/Cloud Blueprint



*Mission HPC and data center for the
DOE Office of Science
8,000+ Users, 800+ Projects
2000+ NERSC citations per year*

Wahid Bhimji
Group Lead, Data Analytics Services
NERSC
Sep 25th 2022

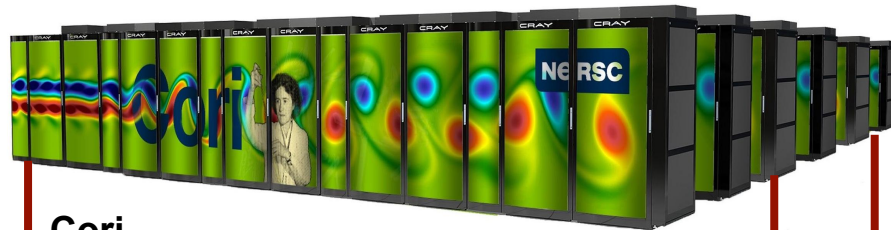
NERSC Systems Spring 2022



Perlmutter P1

1,536 NVIDIA A100 accelerated nodes
4 A100 GPUs & 1 AMD 'Milan' CPU
3,072 AMD 'Milan' CPU nodes
2x AMD Epyc 7763
HPE Cray Slingshot high speed interconnect
140 PF Peak

5 TB/s
35 PB Scratch



Cori

9,600 Intel Xeon Phi "KNL" manycore nodes
2,000 Intel Xeon "Haswell" nodes
700,000 processor cores, 1.2 PB memory
Cray XC40 / Aries Dragonfly interconnect
30 PF Peak

1.5 TB/s
2 PB Burst Buffer
700 GB/s
28 PB Scratch

50 GB/s

HPSS Tape Archive
~200 PB

DTNs, Spin, Gateways

Ethernet & IB Fabric

Science Friendly Security
Production Monitoring
Power Efficiency

LAN



2 x 100 Gb/s
SDN

100 GB/s

5 GB/s

120 PB
Common
File
System

275 TB
/home



NERSC Perlmutter

1,536 GPU nodes

1x AMD Epyc 7763
4x NVIDIA A100
4x Slingshot NICs



3,072 CPU nodes

2x AMD Epyc 7763
1x Slingshot NIC

Slingshot 200 Gb/s 2-level dragonfly

16x MDS + 274 OSS

1x AMD Epyc 7502P
2x Slingshot NICs
24x 15.36 TB NVMe

24x Gateway nodes

2x Slingshot NICs
2x 200G HCAs

2x Arista 7804 routers

400 Gb/s/port
> 10 Tb/s routing



SAN

SAN

WAN

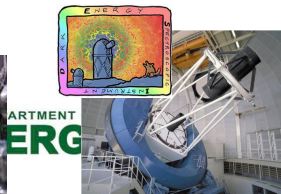
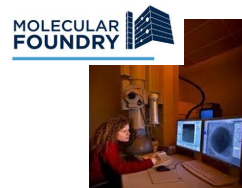
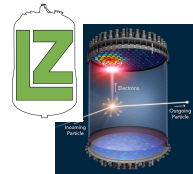


The LBNL Superfacility 'project' coordinates work across divisions to support the Superfacility Model of cross-facility science

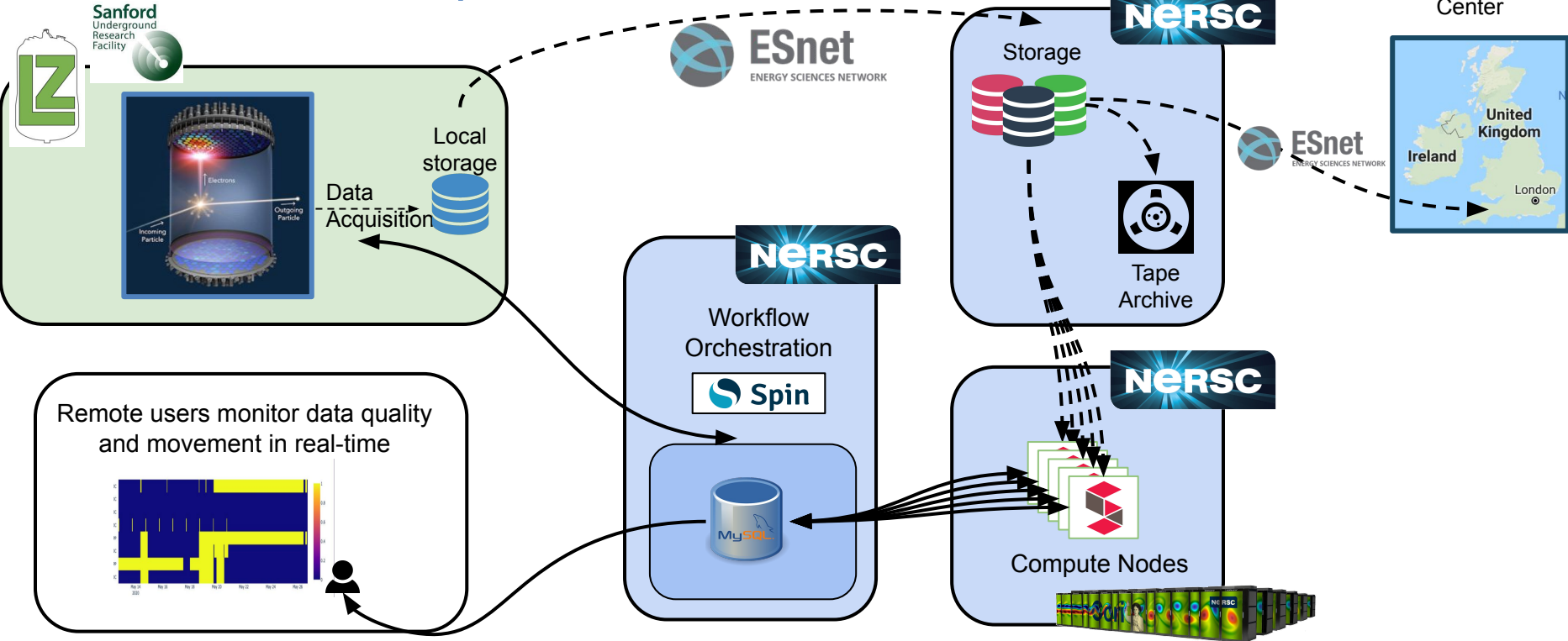
The majority of our 8 science application engagements have now demonstrated automated pipelines that analyze data from remote facilities at large scale, without routine human intervention, using these capabilities developed by the Superfacility effort:

- **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**

In several cases (LZ, DESI, LCLS, NCEM) we have gone beyond demonstrations and can now provide production-level services for their experiment teams.



LZ: real-time DAQ and data management for a dark matter direct detection experiment



Scientists increasingly needs access to a whole ecosystem of HPC services

- NESAP program is helping applications port to energy-efficient architectures
- AI strategy is accelerating adoption of cutting-edge methods in simulation and analysis
- Superfacility work is addressing many of the technical and research needs of experiment facilities
 - we have designed an integrated program of research and technical development to make workflows seamless and scalable across multiple scientific communities
- Looking to the future:
 - N10 has CD-0, will focus on workflows rather than applications
 - Increased importance of cross-facility workflows, including the Integrated Research Infrastructure effort

NERSC users require a shift in the way we design, configure and operate HPC systems

Users require an *integrated ecosystem* that supports new paradigms for *data analysis with real-time interactive feedback* between experiments and simulations.

Users need the ability to *search, analyze, reuse, and combine data from different sources* into large scale simulations and AI models.

NERSC-10 Mission Need Statement: The NERSC-10 system will accelerate end-to-end DOE SC workflows and enable new modes of scientific discovery through the integration of experiment, data analysis, and simulation.

**NERSC-10 CD-0
achieved Sept. 2021**

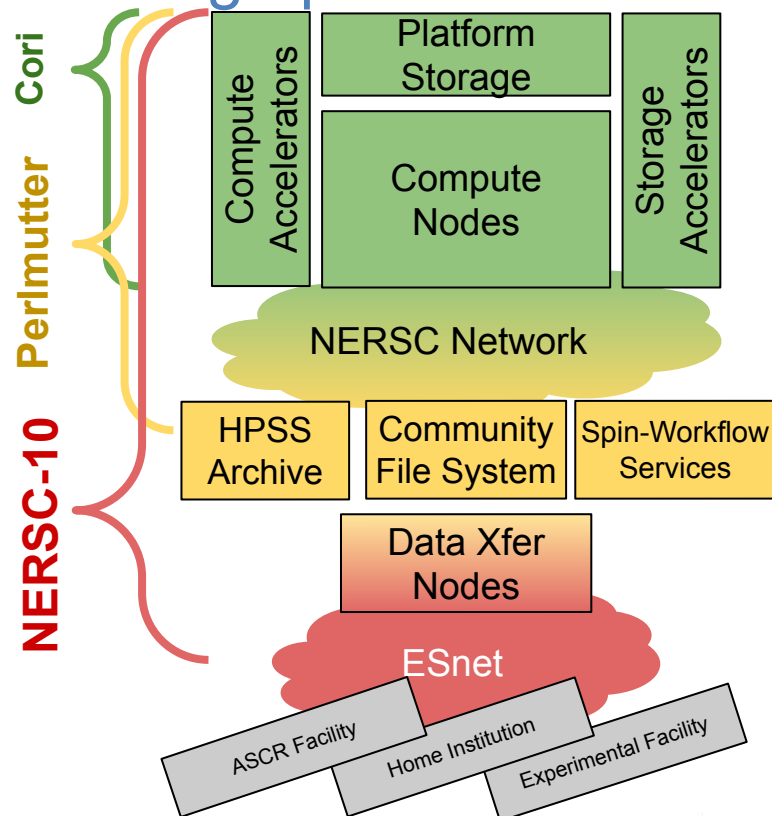


NERSC-10 Architecture: Designed to support complex simulation and data analysis workflows at high performance

NERSC-10 will provide on-demand, dynamically composable, and resilient workflows across heterogeneous elements within NERSC and extending to the edge of experimental facilities and other user endpoints

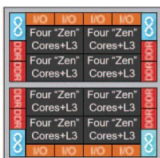
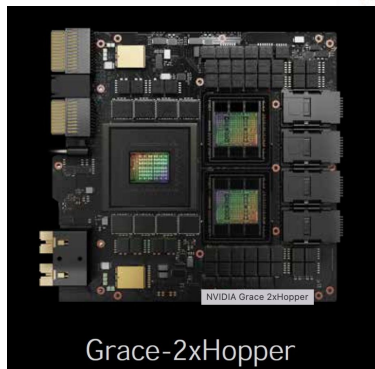
Complexity and heterogeneity managed using complementary technologies

- **Programmable infrastructure:** avoid downfalls of one-size-fits-all, monolithic architecture
- **AI and automation:** sensible selection of default behaviours to reduce complexity for users



A time of challenges and opportunities in *resources*

Increased specialization and heterogeneity

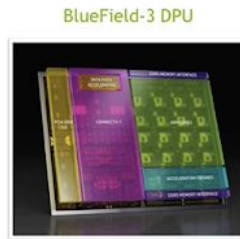
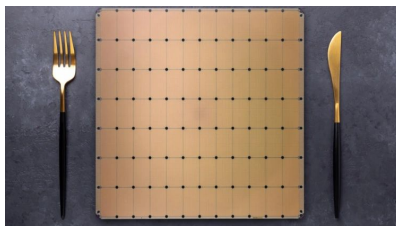


Monolithic 32-core Chip
777mm² total area
1.0x Cost



4 x 8-core Chiplet, 213mm² per chiplet
852mm² total area (+9.7%)
0.59x Cost

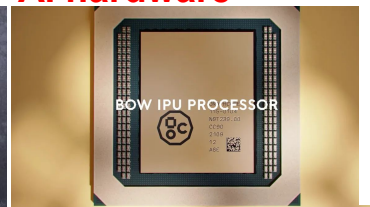
Cpu chiplets



BlueField-3 DPU

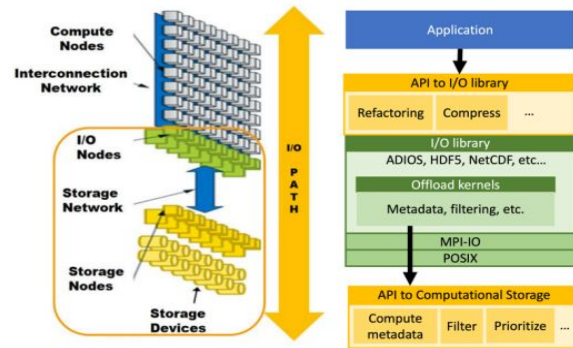
Everything Gets Better

AI hardware

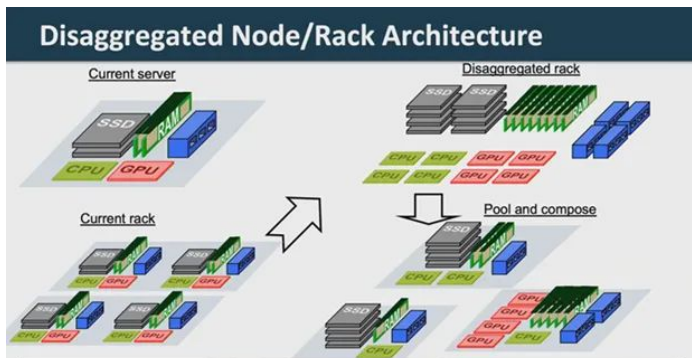


Processing in Storage

Image from Bonnet et al. ASCR [position paper](#)



Disaggregation and more efficient use of resources



A time of challenges and opportunities in *interfaces*

Huge investments from the cloud that can be exploited. But it is reshaping

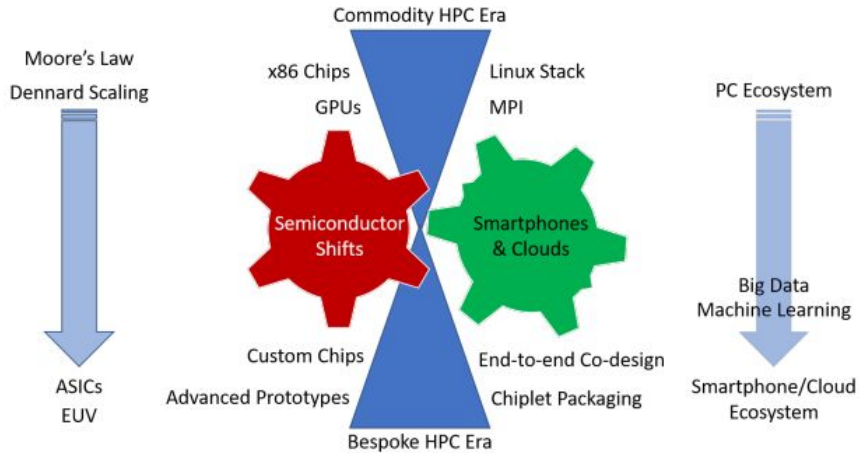


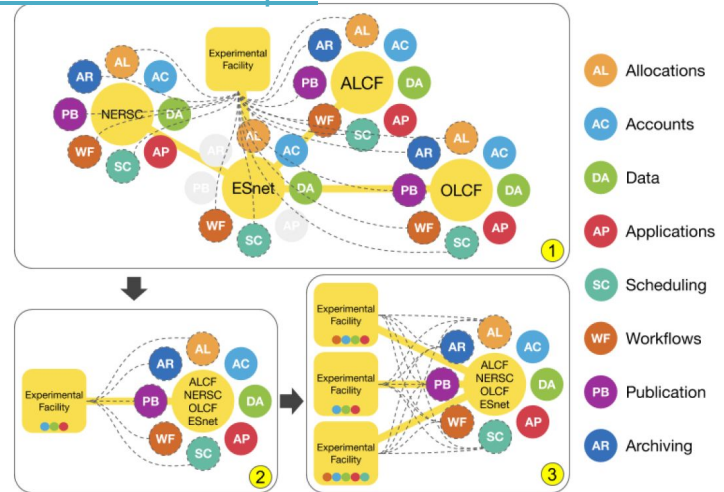
Figure 1: Technical and Economic Forces Reshaping HPC

E.g. Reed, Gannon, Dongarra Reinventing High Performance Computing: Challenges and Opportunities. [arXiv:2203.02544](https://arxiv.org/abs/2203.02544)

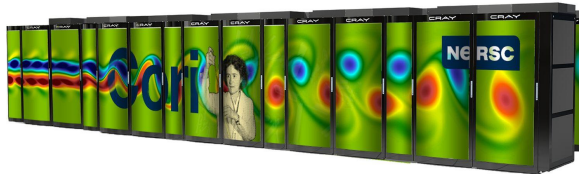
Push towards APIs and integrated access to resources



Integrated Research Infrastructure Task Force White Paper



Exploring the merging of HPC and commercial cloud technologies



2016: Traditional HPC system with a 'data partition'



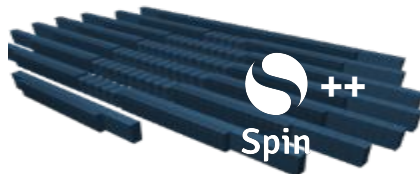
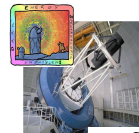
NERSC internal cloud launched: supports databases, web gateways, workflow engines and monitoring.



2020: Integrates Cloud Technologies to improve availability and flexibility; Innovative network that combines high-performance plus flexibility



Tighter system coupling, mounts HPC file systems, access to HPC scheduler. Many high visibility projects



2025:NERSC-10: Convergence to harness the power of traditional HPC with the flexibility, configurability and interactivity of cloud to support growing science requirements.

Exploring the expansion of NERSC's role in data management and stewardship

Data Repository Services: enable curation, search, tracking of metadata and dataset versions to enable FAIR Principles for DOE Science Projects

- Provide the requisite infrastructure and services
- Leverage and Integrate Key Technologies for Identity Management, Data Publishing, Data Transfer Services, Data Exploration, etc
- Provide infrastructure to rapidly deploy new services via Spin coming from CS Research Community and DOE Science Projects



Hosting data services and Portals



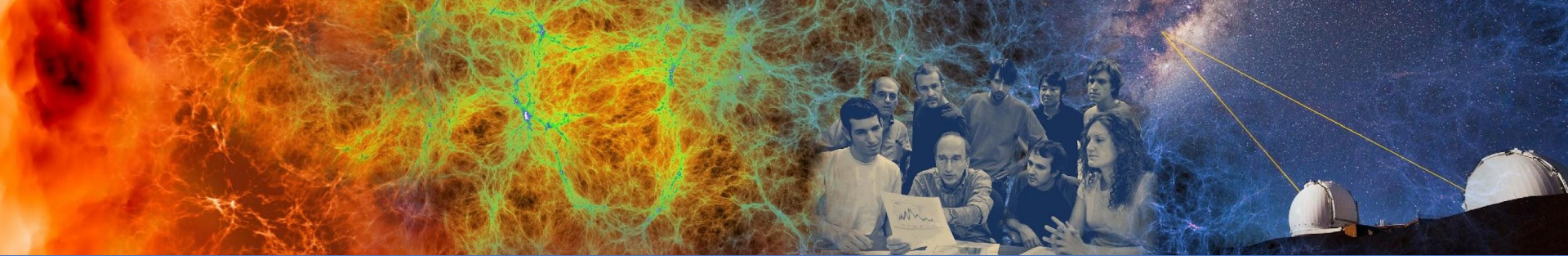
Data Transfer Services



Data storage and archive



Data Exploration and Viz



Thanks!
Questions?



BERKELEY LAB



U.S. DEPARTMENT OF
ENERGY

Office of
Science