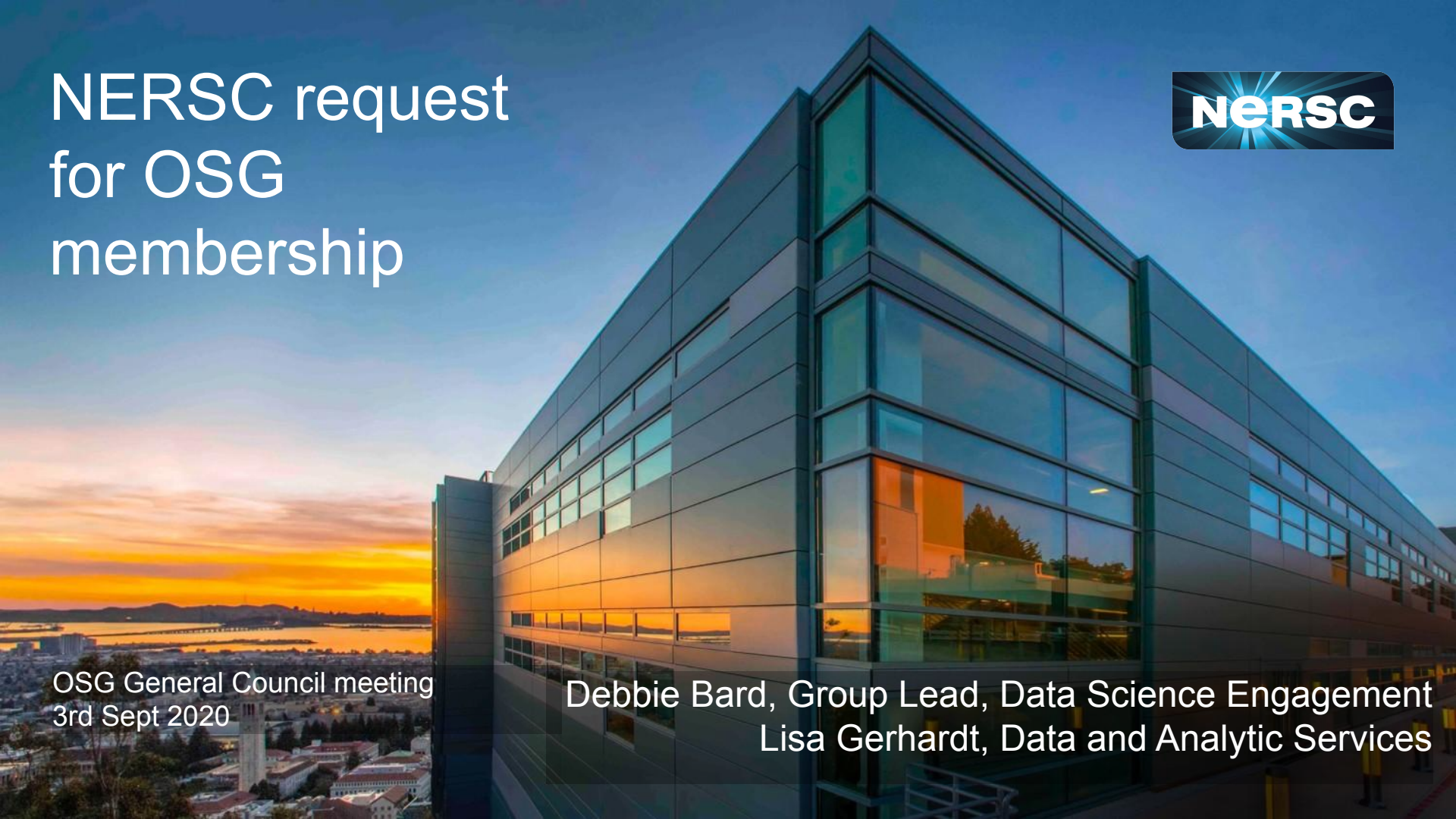
NERSC request for OSG membership



OSG General Council meeting
3rd Sept 2020

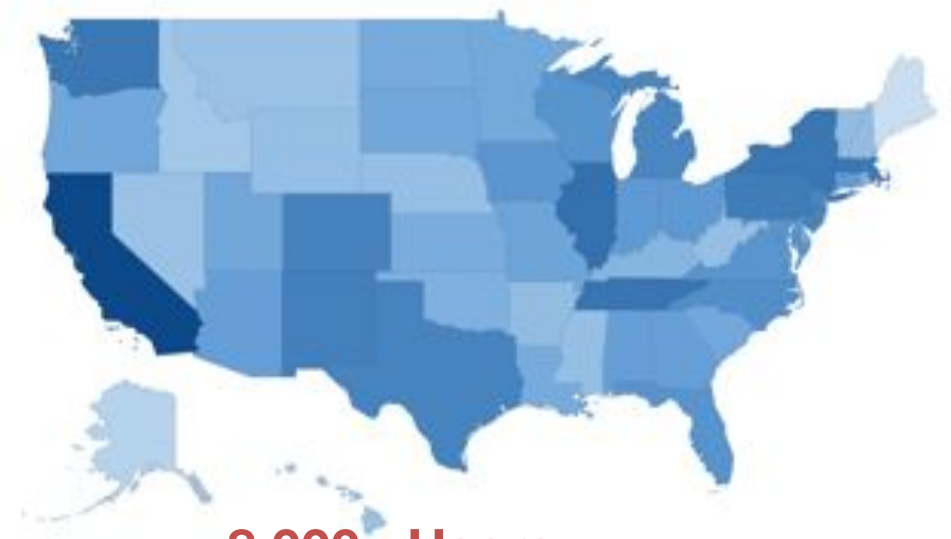
Debbie Bard, Group Lead, Data Science Engagement
Lisa Gerhardt, Data and Analytic Services

ASCR has 3 computational user facilities

- **National Energy Research Scientific Computing Center** (NERSC): ***Mission*** *computing facility for DOE Office of Science*
  - All DOE SC-funded scientists can request time on our systems
  - Time allocated by DOE program managers

- **OLCF and ALCF:**
  - Highly competitive open user allocation programs (INCITE, ALCC).
  - Tens of projects accepted, each receives huge amounts of time on highly specialised cutting-edge hardware
  - time allocated by LCF center and review committees

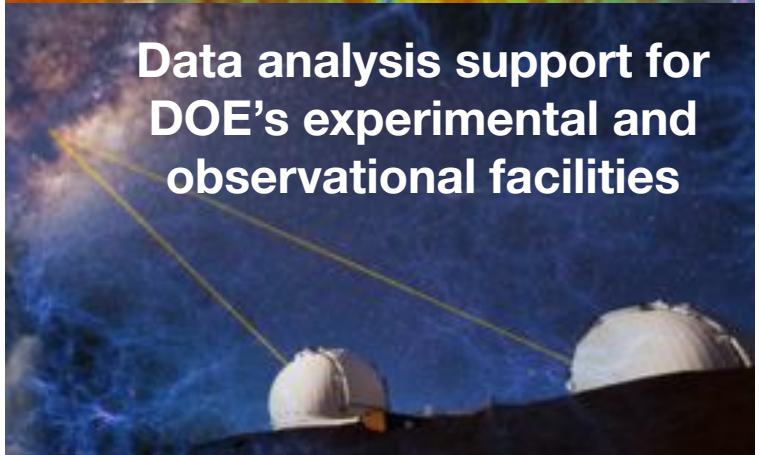
NERSC is the **mission** High Performance Computing facility for the DOE SC



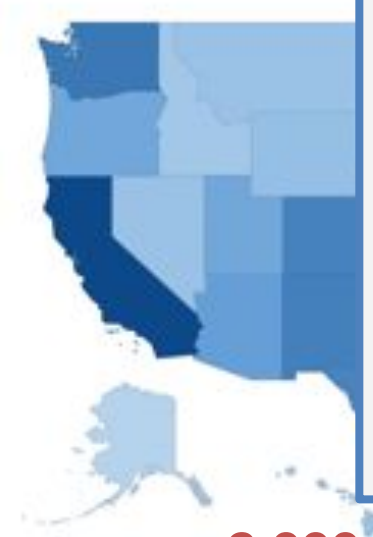
8,000+ Users

800+ Projects

2000+ NERSC citations per year



NERSC is the **mission** High Performance Computing facility for the DOE SC

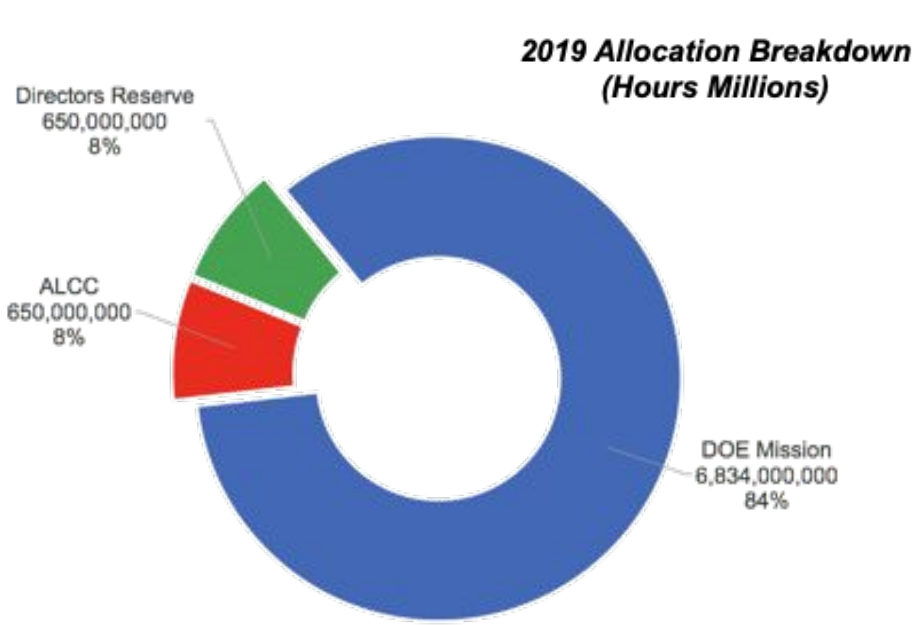


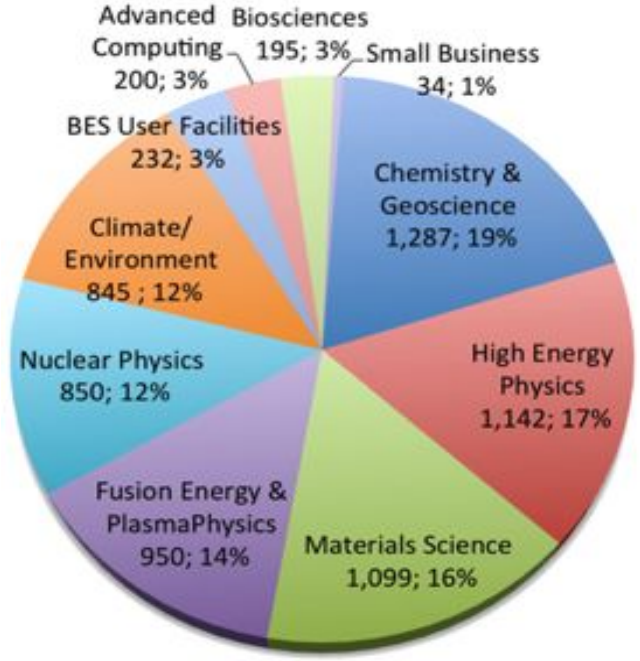
Our users include many of the same groups that constitute OSG: individual researchers, multi-institutional science teams, and 3 of the 4 “big science” projects (US-ATLAS, US-CMS and IceCube)

We want to join the OSG Council because we want to be involved in the strategic decisions that will impact a significant proportion of our user base

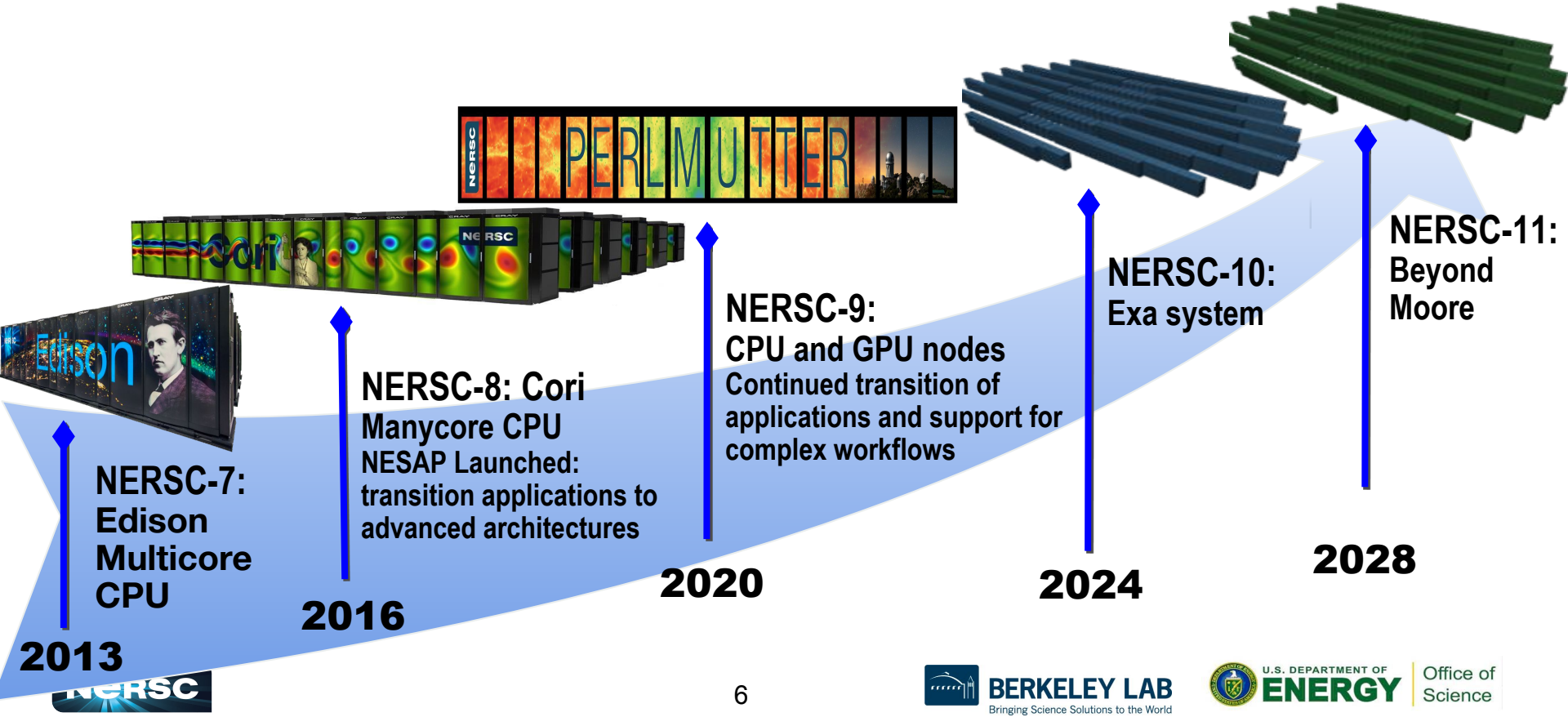
8,000+ Users
800+ Projects
2000+ NERSC citations per year

NERSC Directly Supports Office of Science Priorities



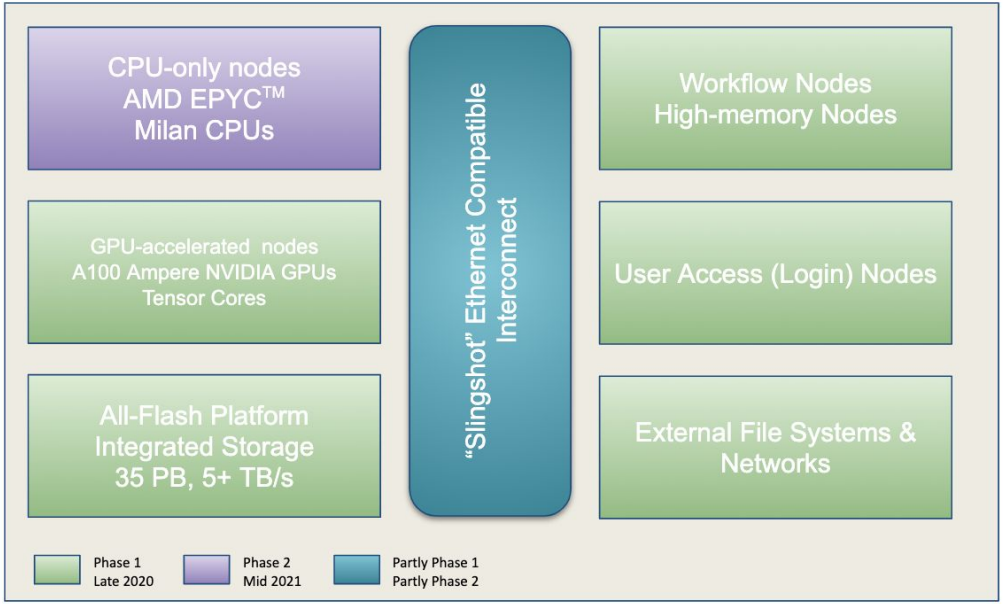


NERSC Systems Roadmap



Perlmutter: a System Optimized for Science

- AMD/NVIDIA **A100-accelerated** and **CPU-only nodes** meet the needs of large scale simulation and data analysis from experimental facilities
- Cray "**Slingshot**" - High-performance, scalable, low-latency Ethernet-compatible network
  - seamless connection between inside/outside the machine
- Single-tier **All-Flash Lustre** HPC file system, 6x Cori's bandwidth
- Dedicated login and high memory nodes to support complex workflows



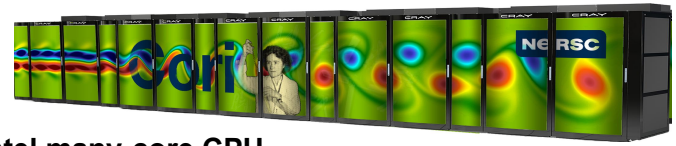
NERSC

BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science

DOE HPC Roadmap - GPUs

2016 2017 2018 2019 2020 2021 2022 2023

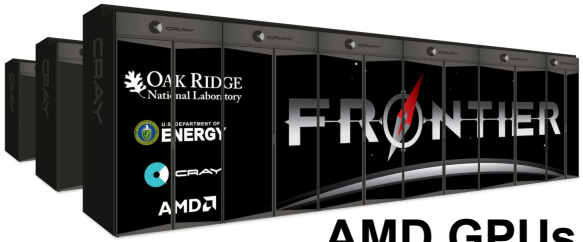


Intel many-core CPU

NVIDIA Ampere

Intel GPUs

NVIDIA Volta GPUs



AMD GPUs

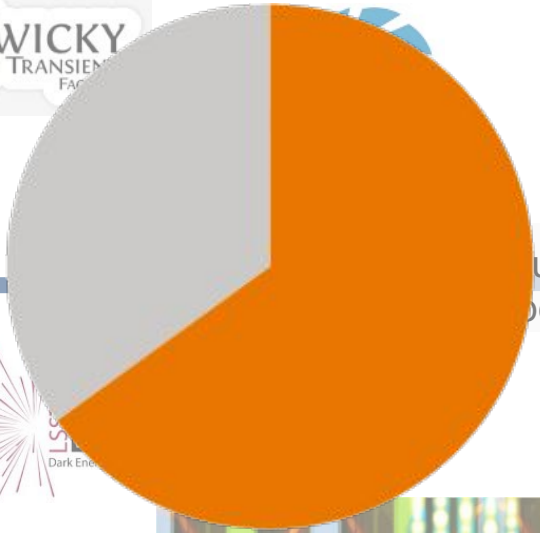
NERSC supports many users and projects from DOE SC's experimental and observational facilities - as does OSG

Experiments operating now

Future experiments

NERSC supports many users and projects from DOE SC's experimental and observational facilities - as does OSG

~35% of NERSC projects in 2018 said the primary role of the project is to work with experimental data



ALS

ZWICKY TRANSIENT FACILITY

BioEPIC

Experiments operating now

ure

periments

PALOMAR TRANSIENT FACTORY

CMB-S4
Next Generation CMB Experiment

ATLAS EXPERIMENT

STAR

MOLECULAR FOUNDRY

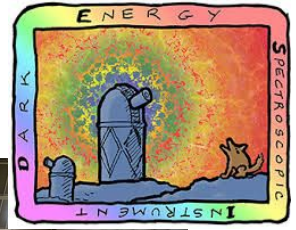
LCLS-II

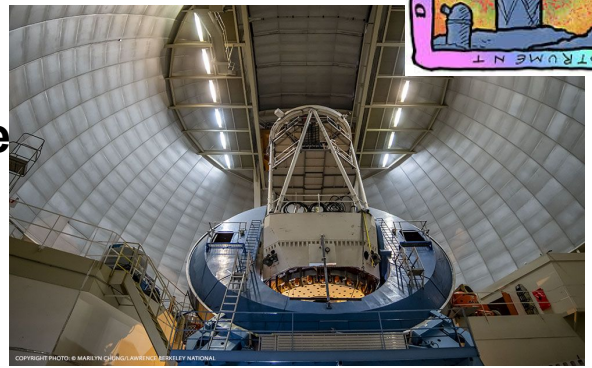
DESI: Dark Energy Spectroscopic Instrument

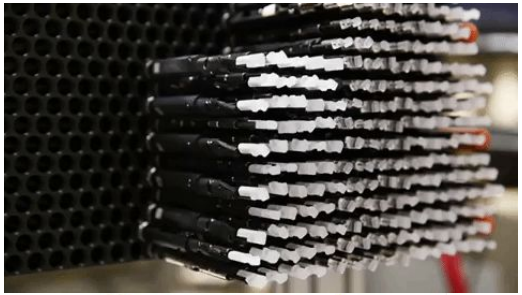
Explaining the Physics of Dark Energy with 3D map of the Universe over 10 billion years

How DESI uses NERSC:

- Analyse Kitt Peak telescope data in **quasi-realtime** to select targets each night
- Co-locate survey/sim data; HPC-scale re-processing of data
- Large collaboration monitors survey progress and **share results**.
- Part of **NESAP** program for GPU readiness







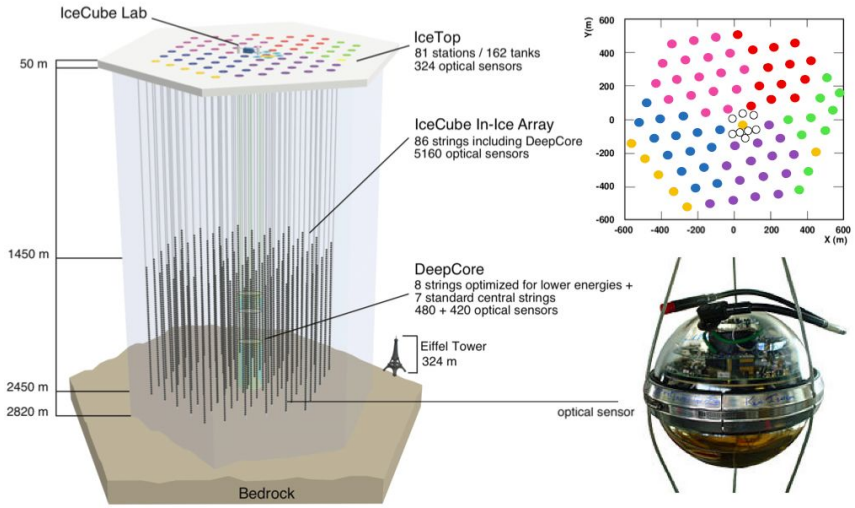
LSST-DESC

Use data from the Rubin Observatory to explain Dark Energy through multiple probes

How DESC uses NERSC:

- NERSC is primary data facility: multi-PB storage, both active and archive
- Co-locate cosmology, instrument and image simulations with data analysis
- Use **Spin** for supernova alert broker
- Use **Jupyter** for analysis
- ImSim is a **NESAP** project

IceCube



	2015	2016	2017	2018	2019
TB written during year	3300	700	700	700	700
Total TB end of year	3300	4000	4700	5400	6100

IceCube Neutrino Observatory has a long history at NERSC

- Allocations at NERSC since 2010
- Use HPSS tape archive for storage
  - IceCube and NERSC have signed an MOU to be the second archive of IceCube's experimental data
- Collaboration between NERSC and IceCube developed GraphNN for astrophysical neutrino/cosmic ray classification
  - ICMLA19 Best Paper arXiv:1809.06166

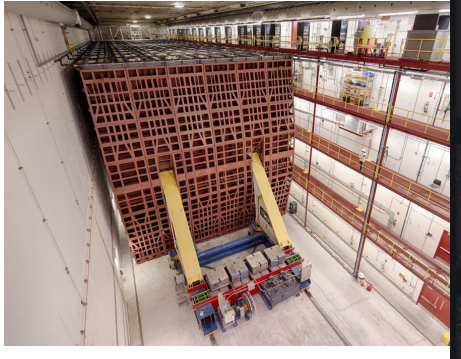
BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY | Office of Science

Particle physics

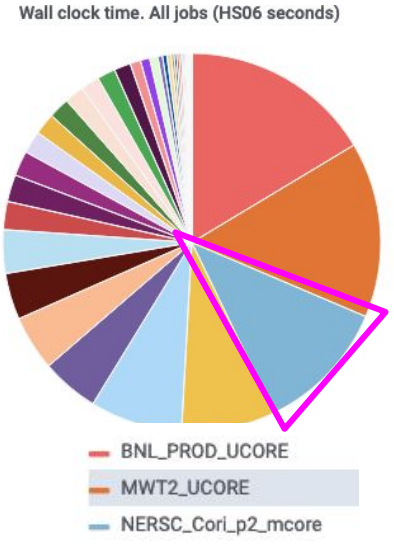
NoVA '1m cores'

- Used whole of Cori for timely re-processing of data



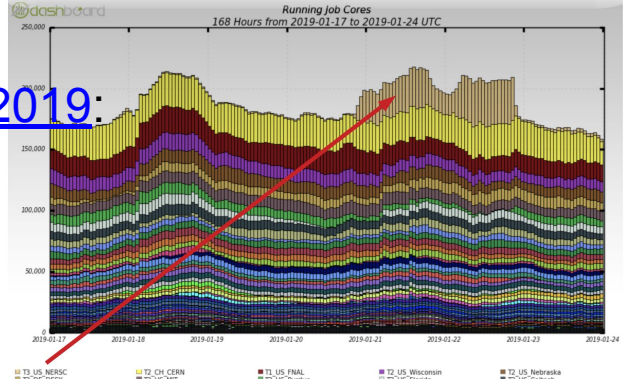
ATLAS

- Uses Cori for MC production
  - E.g. 2020 so far:

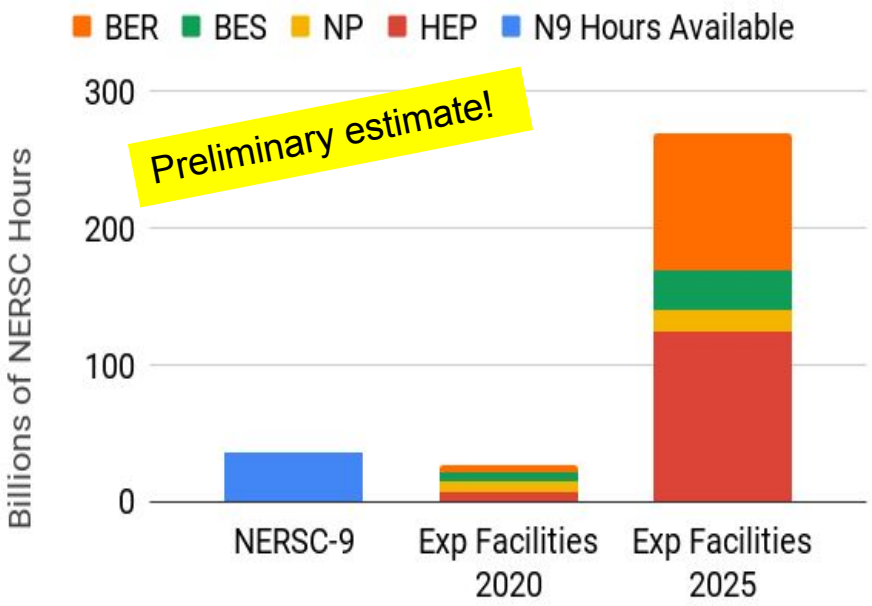


CMS

- Runs a variety of workloads:
  - E.g. Jan 2019:



Needs go beyond compute hours



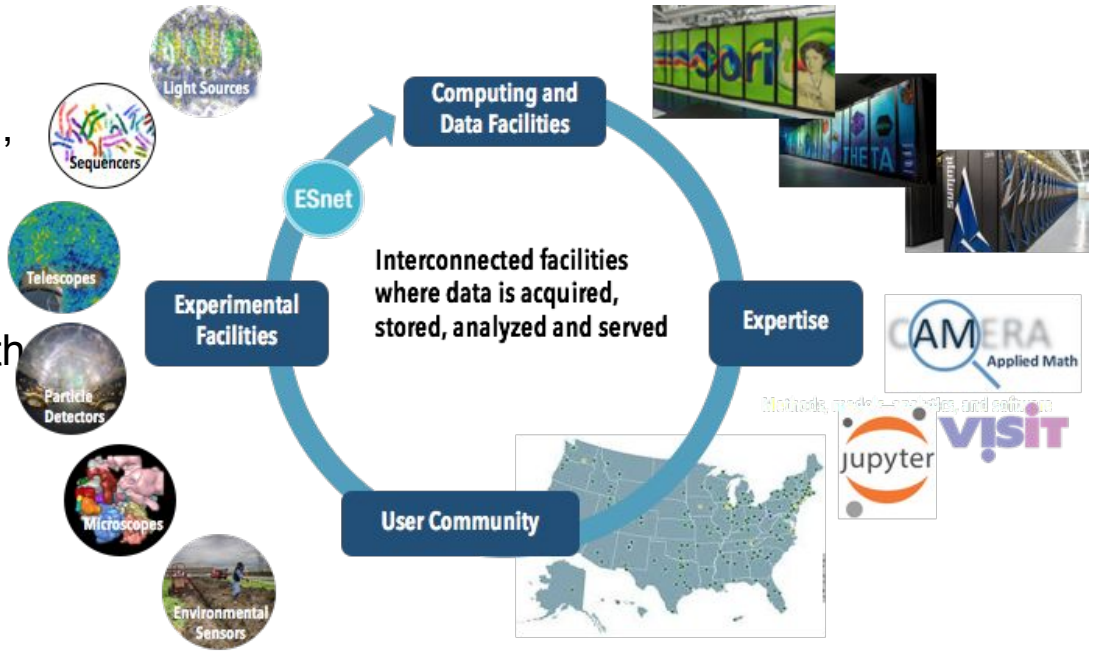
Taken from Exascale Requirements Reviews

- High data volumes (today use ~19% of computing hours, but store 78% of data) that are moved between sites
- Real-time (or near) turnaround and interactive access for running experiments
- Resilient workflows to run across multiple compute sites
- Ecosystem of persistent edge services, including workflow managers, visualization, databases, web services…

Superfacility: an ecosystem of connected facilities, software and expertise to enable new modes of discovery

Superfacility@ LBNL: NERSC, ESnet and CRD working together

- A model to integrate experimental, computational and networking facilities for reproducible science
- Enabling new discoveries by coupling experimental science with large scale data analysis and simulations
- **A lot of these goals and technologies are shared with OSG**

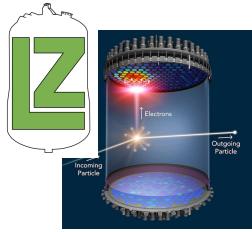


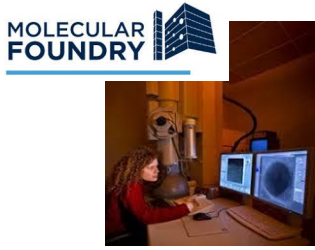
The CS Area Superfacility 'project' coordinates and tracks our work

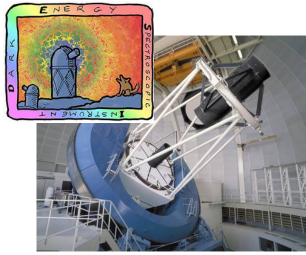
Project Goal:

By the end of CY 2021, 3 (or more) of our 7 science application engagements will demonstrate automated pipelines that analyze data from remote facilities at large scale, without routine human intervention, using these capabilities:

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







OSG software stack

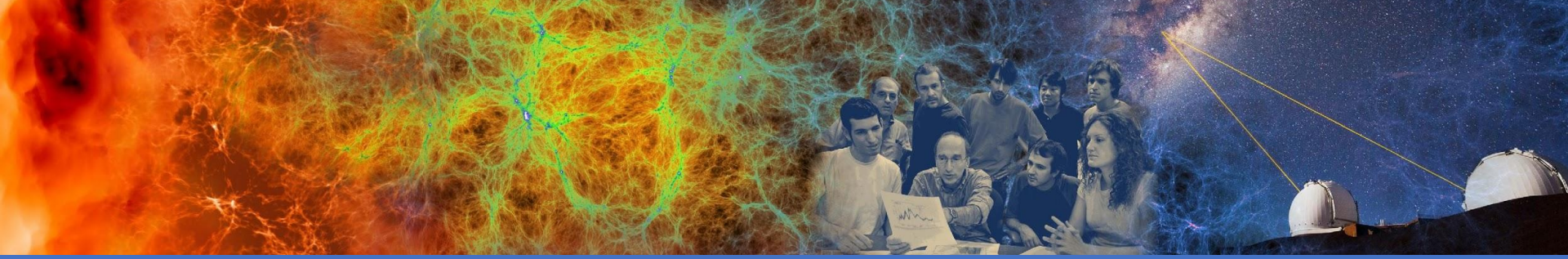
NERSC is in the process of evaluating whether we can support the OSG stack

- This is hard, due to our highly specialised hardware, system software and security requirements
- We cannot commit to running full stack, but we have had success in supporting parts…
  - eg cvmfs, globus, xrootd...

We want to get more closely engaged with OSG to understand where technology choices are heading (eg post-globus), and how we can support them for our users.

Summary

- NERSC is already used by many teams who are part of OSG
  - we have a good understanding of what OSG members need from us
  - we want to be more closely involved in how decisions are made and how strategy is set for OSG members
- We have users from every DOE-funded experimental and user facility
  - we have a wide perspective on what this community is doing computing-wise
  - we want to learn more about what OSG members need
  - the Superfacility concept/framework is a natural fit for both NERSC and OSG
- We run a supercomputing center with unique hardware and capabilities
  - we are plugged in to future computing trends and have close relationships with both large-scale vendors and startups
  - we want to build a closer connection to OSG to help guide our future technology choices to better support our users
- **We believe that joining the OSG board would be of great benefit to us, our users and OSG.**



Fin

NERSC is used complementary to OSG by many of our users

Astronomy and Astrophysics

Coming Later in 2020:
Supernova studies for DES Y3 Cosmology Analysis
(Can run embarrassingly parallel with some modifications)
OSG will provide a very important relief valve for pressure on the DES NERSC allocation

8 2020-09-03

ProtoDUNE Production Processing

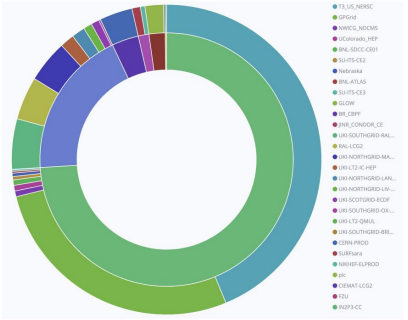
- DUNE doing excellent job inc
  Compute Elements (CE) and
  (SE) using OSG and WLCG in
  – Continue to add resources f
  the world - 36 sites
  – Addition of Storage Elemen
- Soon undertake ProtoDUNE
  Production version 3 (PD-SP
- Data processing on distribute
  – (FNAL ~50% similar to prev
- Utilizing NERSC SuperCompu
  through HEPCloud for simula
  10000 simultaneous jobs runi
  DUNE CPU hours)
- Anticipate using 80 - 100 M C
  during ProtoDUNE II operatio

15 2020-09-03 Fermilab

ProtoDUNE Production Processing

- DUNE doing excellent job incorporating new Compute Elements (CE) and Storage Elements (SE) using OSG and WLCG infrastructure
  – Continue to add resources from sites around the world - 36 sites
  – Addition of Storage Elements continues - 13
- Soon undertake ProtoDUNE Single Phase Production version 3 (PD-SPProd3)
- Data processing on distributed computing
  – (FNAL ~50% similar to previous usage)
- Utilizing NERSC SuperComputer Cori allocation through HEPCloud for simulation generation (10000 simultaneous jobs running ~40% of total DUNE CPU hours)
- Anticipate using 80 - 100 M CPU hours/year in during ProtoDUNE II operations

CPU Hour fraction Feb 1 - May 20, 2020



Adding in HPC @ NERSC Cori

15 2020-09-03 Fermilab

Taken from “Fermilab and OSG: Perspectives from neutrino, muon, and astronomy experiments”
https://indico.fnal.gov/event/22127/contributions/194483/attachments/133930/165401/OSG_AHM_2020_Non-CMS_FNAL_experiments.pdf

Support for CVMFS

- Restrictions on OS (FUSE etc.) meant cvmfs at NERSC was historically challenging.
- Copying whole or part of software stack into a container was/is a solution (automated workflows to build images used in production by e.g ATLAS/CMS)

Support up-to-date cvmfs for entire sw stack w/out large images

- Currently deploy solution via Cray DVS to mount CVMFS over NFS
  - 24 repositories now mounted inc LZ; ATLAS; CMS; AMS, DUNE, NOVA
- **CVMFS will be supported on Perlmutter**: evaluating 'normal' FUSE-based or DVS solutions