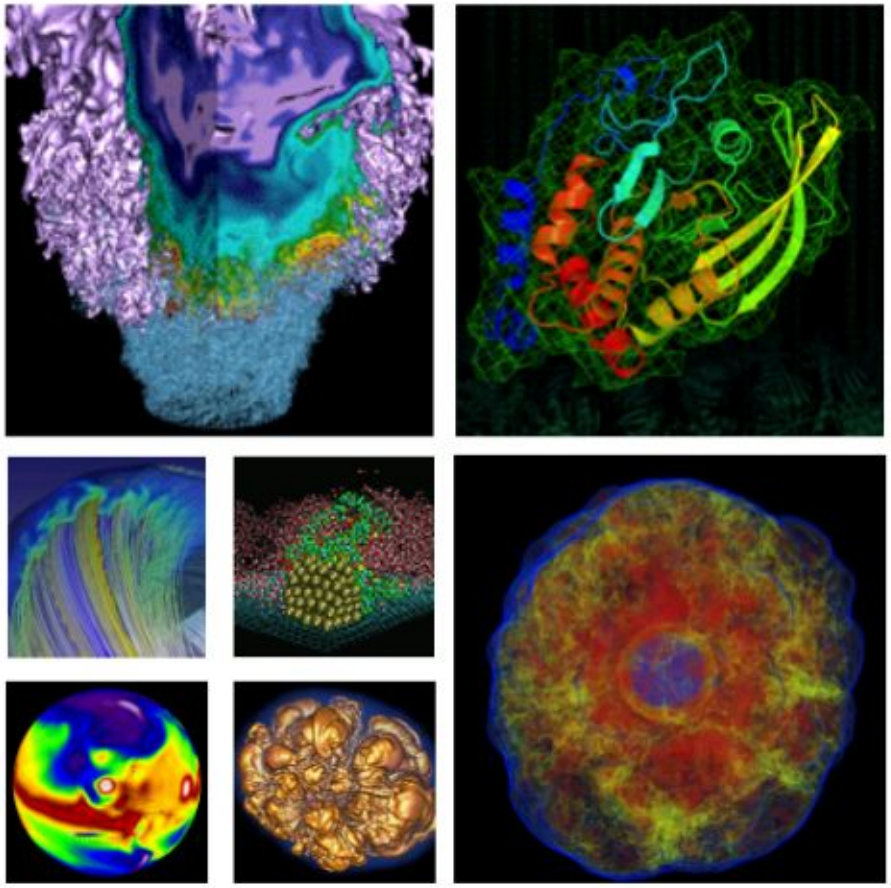
# NERSC Users Group Monthly Webinar



NERSC

June 15, 2017

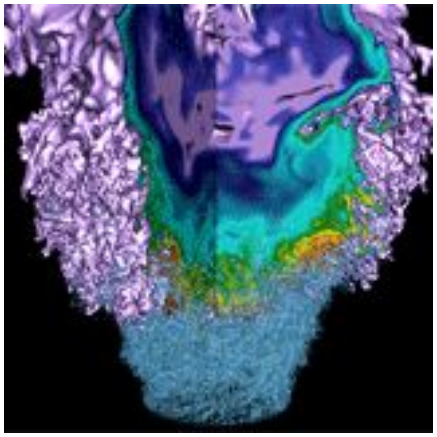
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB
Lawrence Berkeley National Laboratory

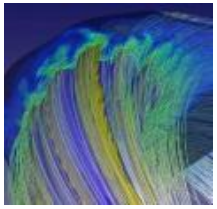
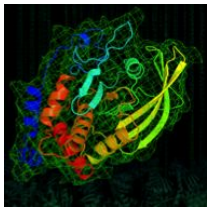
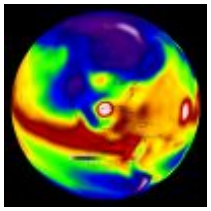
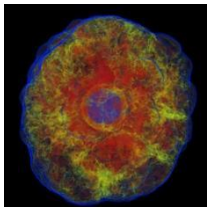
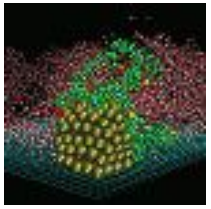
# Agenda

- Cori status and transition to production
- KNL allocations and allocations outlook
- NERSC 9 plans and status
- Storage 2020
- NESAP next steps
- Science at Scale CFP
- Data Day and NUG meeting plans
- ECP
- Q&A

# Cori status and transition to production

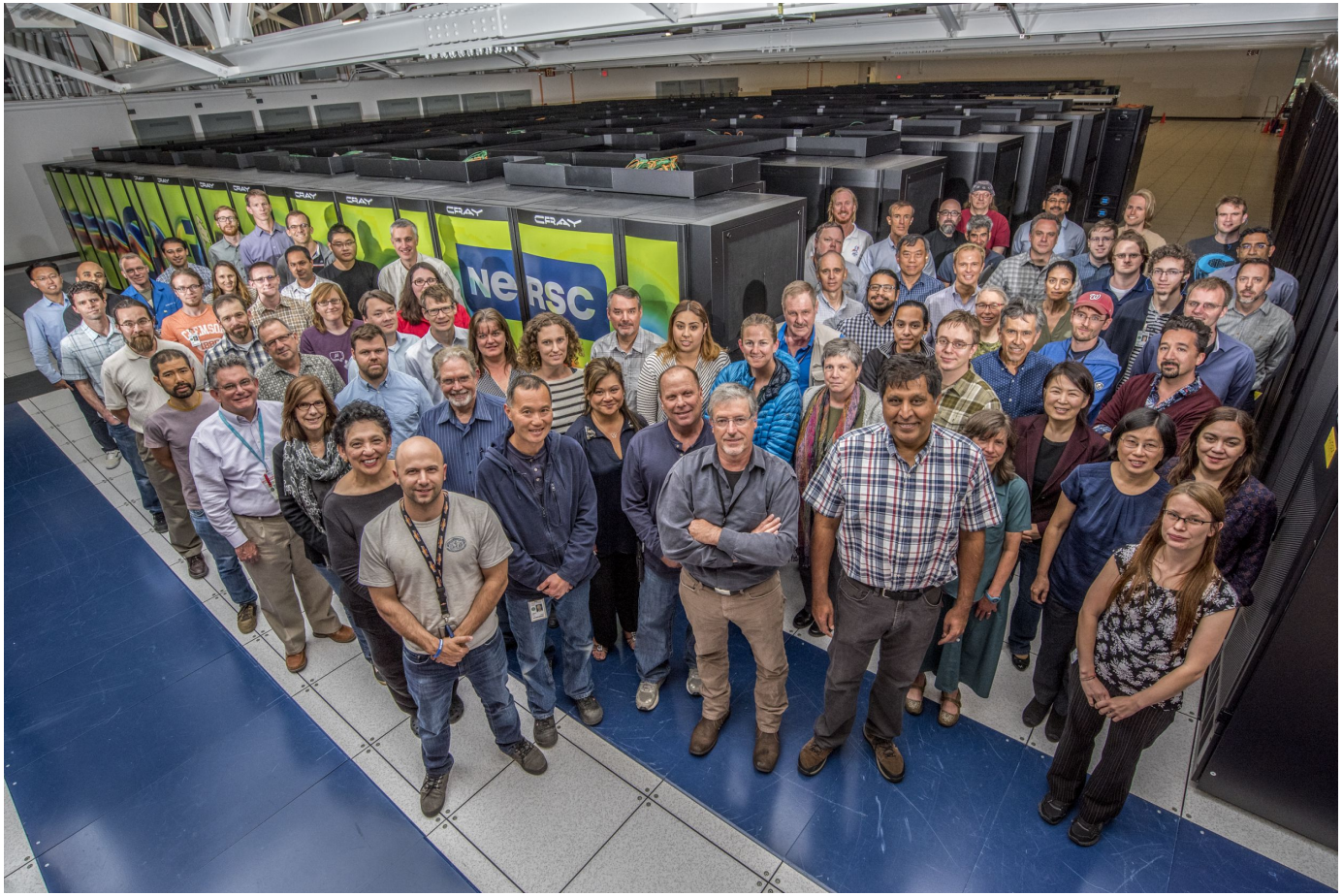


NERSC

# NERSC is proud to be transitioning Cori to production!

**The whole NERSC Center rallied around deploying, testing and accepting Cori**

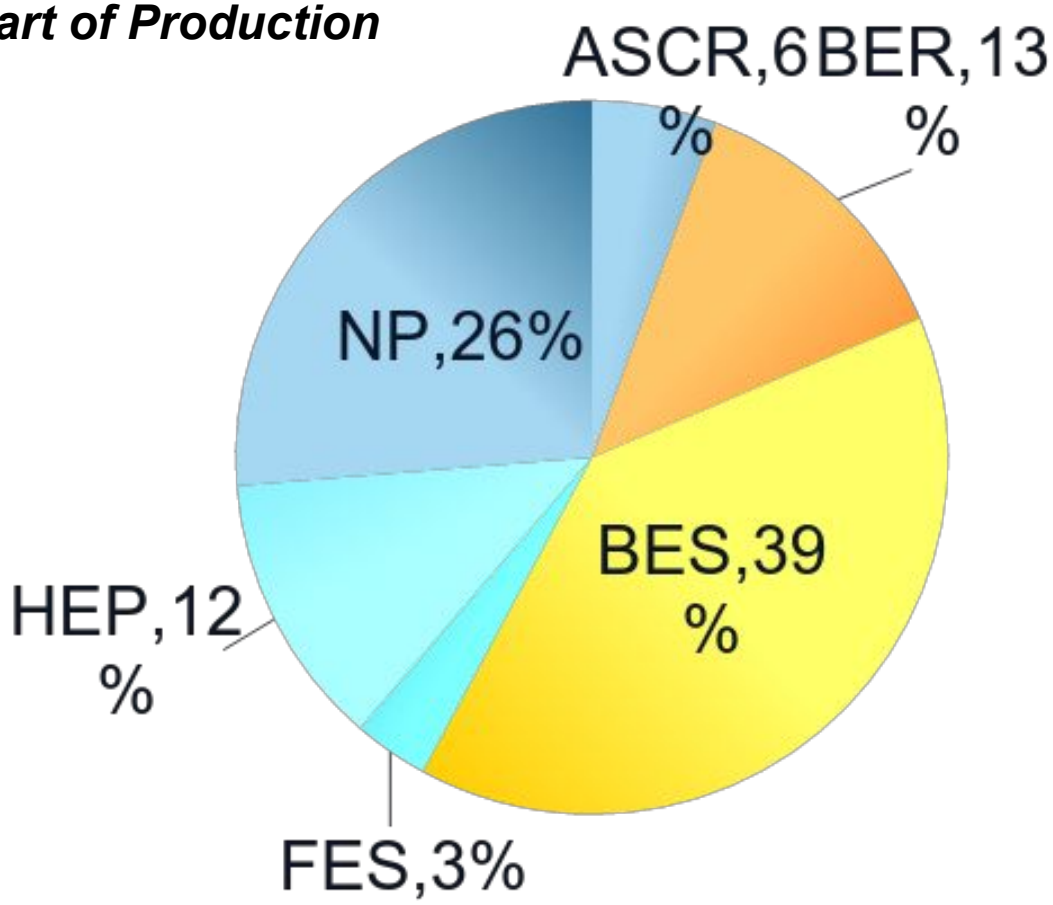


# Transitioning to Production

- **Cori will be entering production soon!  July 1st.**
- **We hope the early, uncharged time has been productive for your science team**
- **During this 'shake-out' period between system acceptance (Dec. 2016) and now we focused on:**
  - Improving KNL stability and reliability, BIOS upgrades on the processor
  - Many system patches, (improve reliability of network, mode changes, job launch)
- **Establish a collaboration with Cray and Intel to address cache-mode performance variability**
- **Testing system cooling with warmer water**
  - (Cori was accepted in winter.)

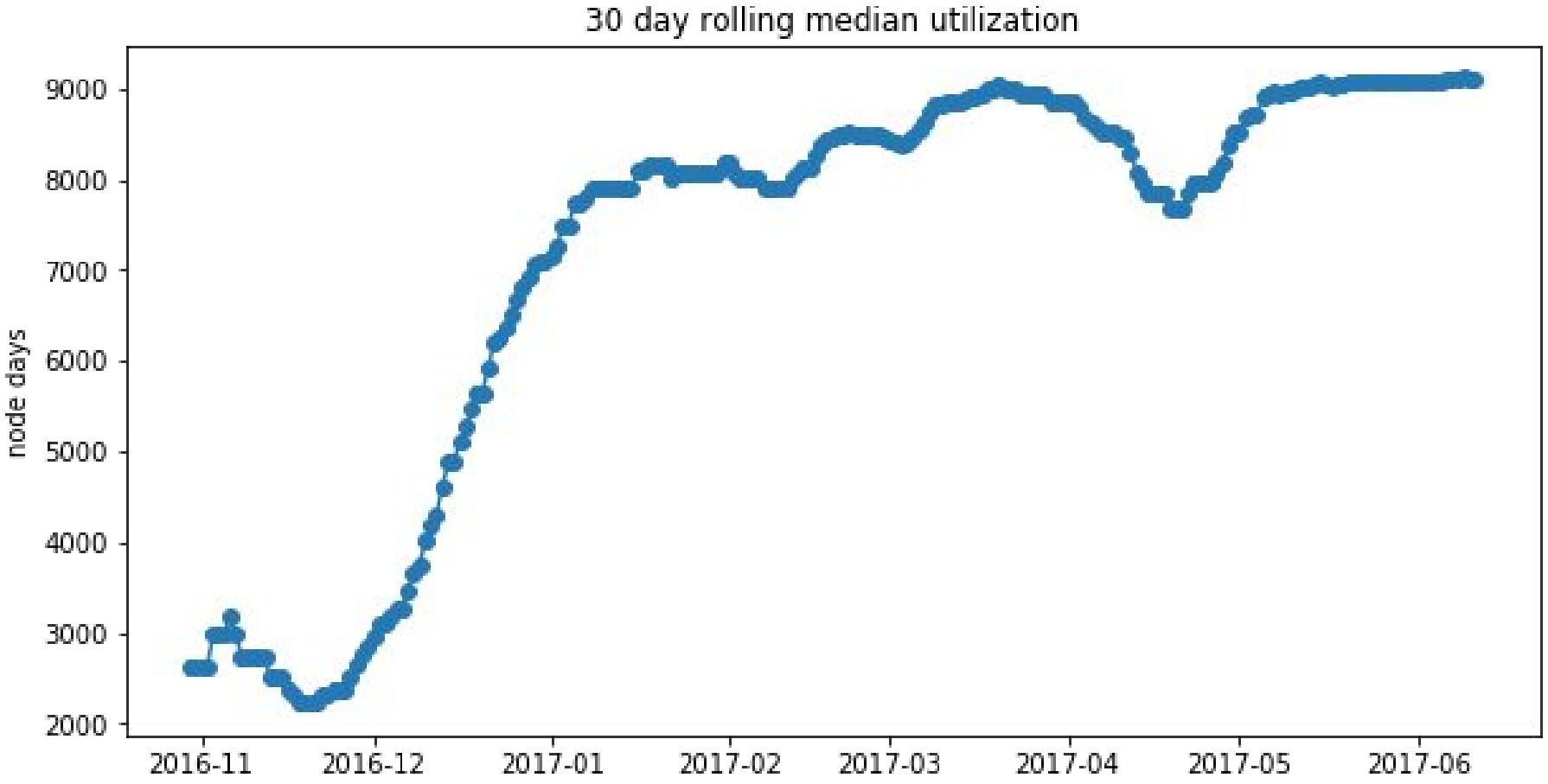
# Users are running science workloads on Cori!

***Breakdown of Early Usage on Cori***
***Acceptance – Start of Production***




**Since system acceptance on Dec. 22nd, NERSC users have used approximately 3B NERSC hours on Cori**

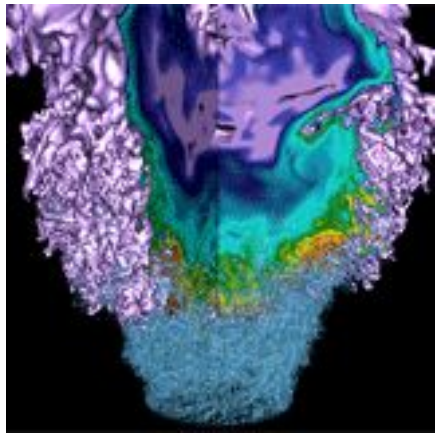
# Users ramped up quickly

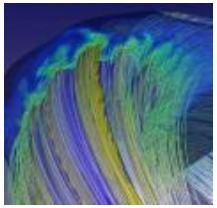


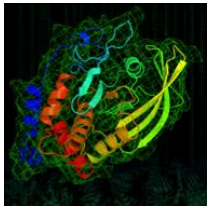
# What Happens on July 1

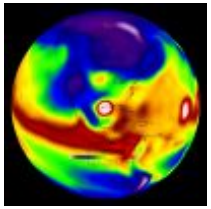
- Charging against repositories begins at 12:00:01 a.m. on July 1
- No disruption of queued or running jobs
- Your jobs will be charged against a repository allocation for all time used after 12:00:01 a.m. on July 1
- If you have jobs in the Cori KNL queue and don't want them to be charged, remove them before July 1
- Allocation augmentation for Cori KNL: see next set of slides

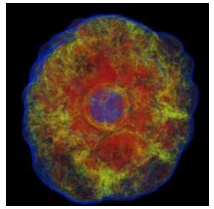
# KNL allocations and allocations outlook

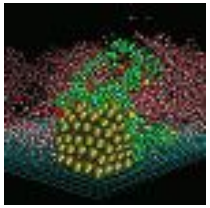












# Allocation Augmentations

- **Charging on Cori KNL nodes begins July 1**
  - 2.4 B additional NERSC Hours to be distributed by DOE program managers
  - Must be used on KNL; Edison & Cori Haswell already fully allocated for 2017
- **Request additional hours**
  - Request form: https://www.nersc.gov/users/accounts/allocations/2017-cori-knl/request/
  - Codes known to run on KNL: https://www.nersc.gov/users/accounts/allocations/2017-cori-knl/knl-codes/
  - Deadline Monday, June 19 (multiple notices have been sent to PIs/Managers)
- **Allocation decisions posted on NERSC web site on June 26**
- **NERSC will monitor usage and take action if newly allocated hours are used too much on Edison/Cori Haswell**

# NERSC Hours Calculation

The currency for charging at NERSC is the “NERSC Hour”.

1 NERSC Hour ~ 1 core hour on Hopper (retired 2015)

The charge ratios reflect the average per node performance ratio on a broad range of applications

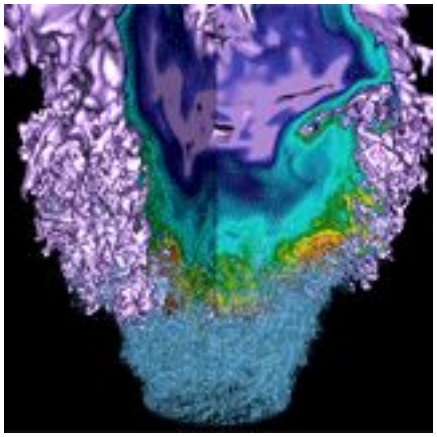
| Node Type | Cores per Node | Charge per Node Hour |
|---|---|---|
| Cori KNL | 68 | 96 |
| Cori Haswell | 32 | 80 |
| Edison | 24 | 48 |
| Hopper (retired) | 24 | 24 |

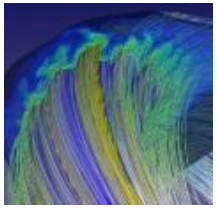
NESAP unoptimized codes on KNL: 1.2X performance vs. Edison per node
NESAP optimized codes on KNL: 3.5X performance vs. Edison per node

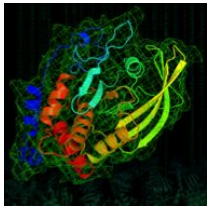
# Allocations Forecast

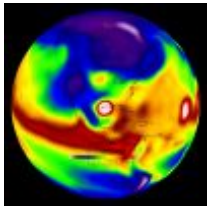
| Year | DOE Production (M NERSC Hrs) | ALCC | Director's Reserve |
|---|---|---|---|
| 2014 | 2,400 | 300 | 300 |
| 2015 | 2,400 | 300 | 300 |
| 2016 | 2,400 | 300 | 300 |
| 2017 | 4,800 | 662.5 | 600 |
| January | 2,400 | 150 | 300 |
| July For KNL | 2,400 | 512.5 | 300 |
| 2018 | 6,800 | 962.5 | 850 |
| 2019 | 6,800 | 850 | TBD (edison retirement date) |

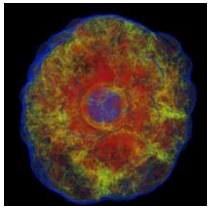
# NERSC 9 plans and status

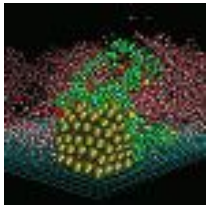












U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

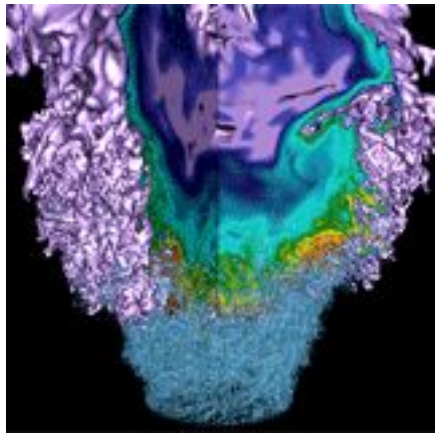
# NERSC-9 Plans

- **NERSC-9 will be the next iteration of our large systems**
- **Expect to have delivery in the 2020 timeframe**
- **A pre-exascale system, designed to support both traditional and data intensive workloads well**
- **Targeting a substantial performance increase over Edison and Cori**
- **Requires facility work at B59 to install and run the system**
  - Will require removal of NERSC-7 (Edison) in the 2019 timeframe

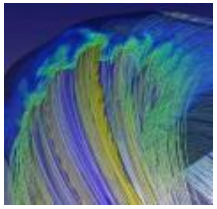
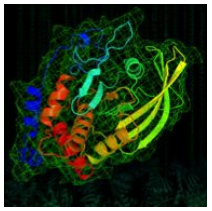
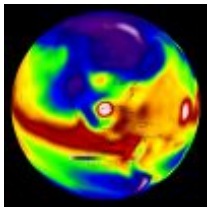
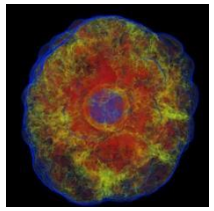
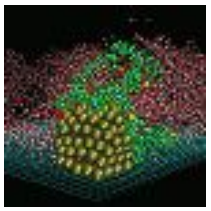
# NERSC-9 Status

- **NERSC-9 RFP was released late 2016 (in collaboration with LANL)**
  - http://www.lanl.gov/projects/apex/request-for-proposal.php
- **System selection is underway**
- **System details to be announced in 2018 - in time for NESAP-2 efforts towards application readiness**

# Storage 2020



NERSC

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

# Storage 2020 Background

NERSC commissioned a working group to determine the future storage roadmap and storage strategy at NERSC.

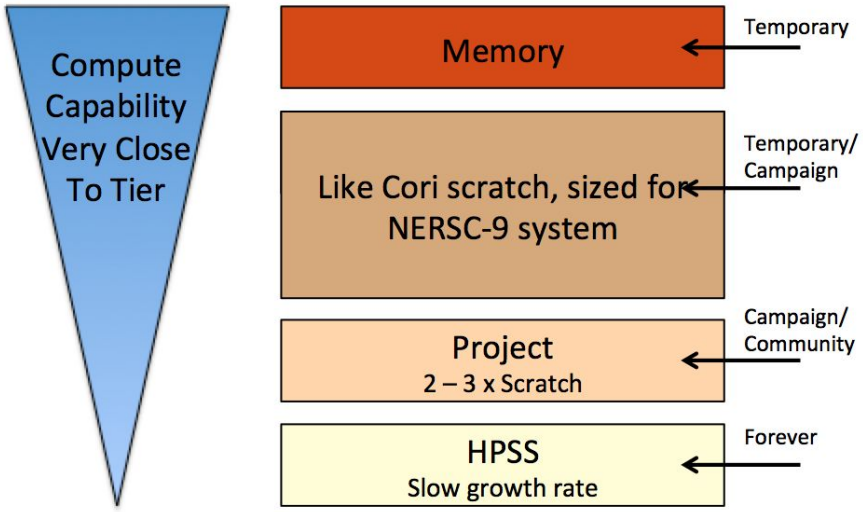
Process:

- Educate ourselves on current and upcoming storage technologies
- Define requirements and a set of recommendations for a center-wide storage infrastructure
- Create a plan to transition our current storage infrastructure to the 2020 plan
- Group of ~20 with expertise in user services, applications, compute and storage systems management

# Findings

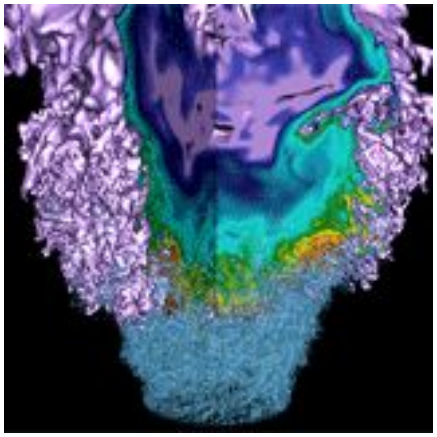
- Complex workflows are a very common mode of operation at NERSC
- Users need high BW and medium capacity when computing
- Users need medium- and long-term storage and place to share data with their local group and wider community
- Complex workflows drive a need for data management tools
- Archived data is not  write-once, read-never at NERSC. Reads account for 40% of HPSS I/O

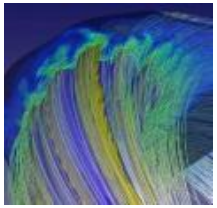
# Where is the Hierarchy Going?

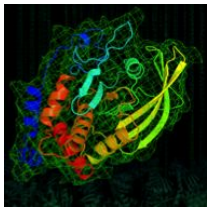


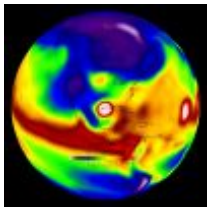

- Ensure that the N9 storage system meets RFP and integrates with other center resources
- Transition today's project tier to provide larger capacity relative to Cori and NERSC-9 tiers, but at lower bandwidth
- Enact policies and quotas that will decelerate the growth of the tape-based archive
- Fully utilize the data management features present in today's project and archive tiers (GHI, ILM, etc)

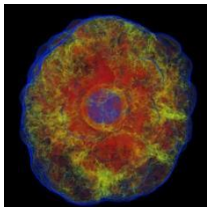
# NESAP next steps

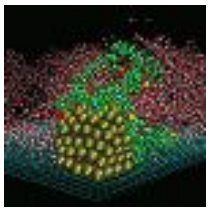




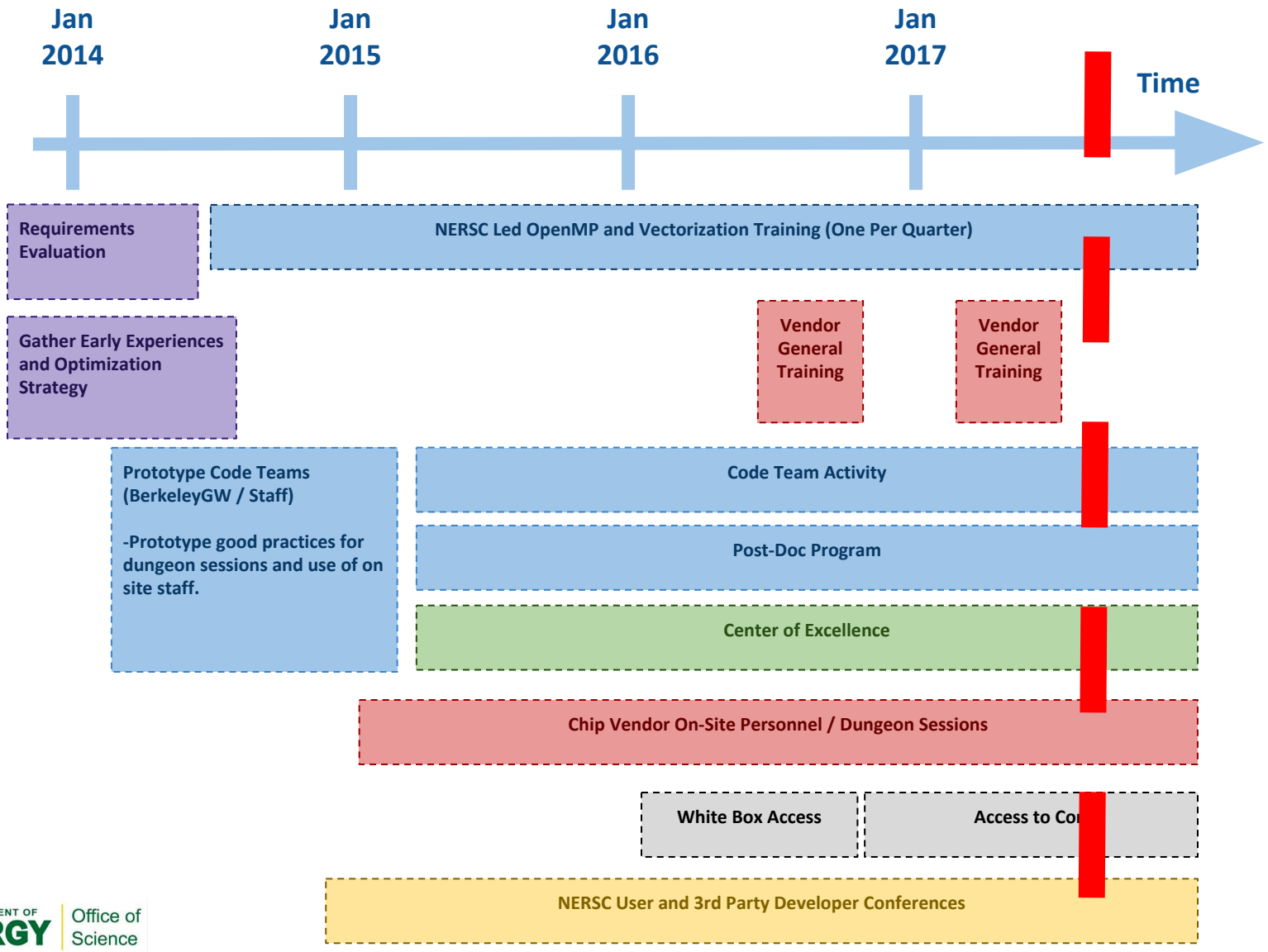




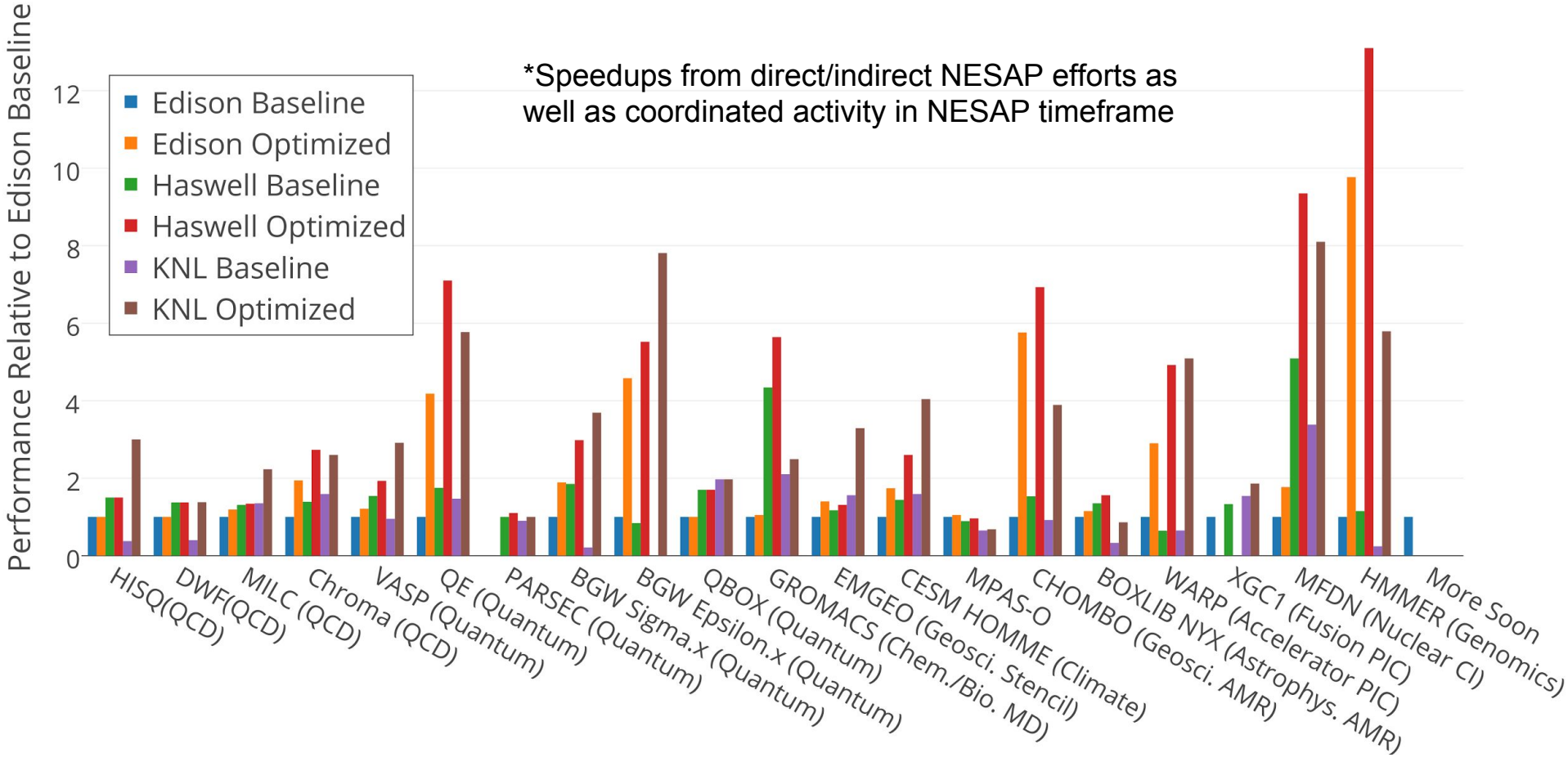




# Timeline


Jan 2014
Jan 2015
Jan 2016
Jan 2017
Time
Requirements Evaluation
NERSC Led OpenMP and Vectorization Training (One Per Quarter)
Gather Early Experiences and Optimization Strategy
Vendor General Training
Vendor General Training
Prototype Code Teams (BerkeleyGW / Staff)
-Prototype good practices for dungeon sessions and use of on site staff.
Code Team Activity
Post-Doc Program
Center of Excellence
Chip Vendor On-Site Personnel / Dungeon Sessions
White Box Access
Access to Co
NERSC User and 3rd Party Developer Conferences

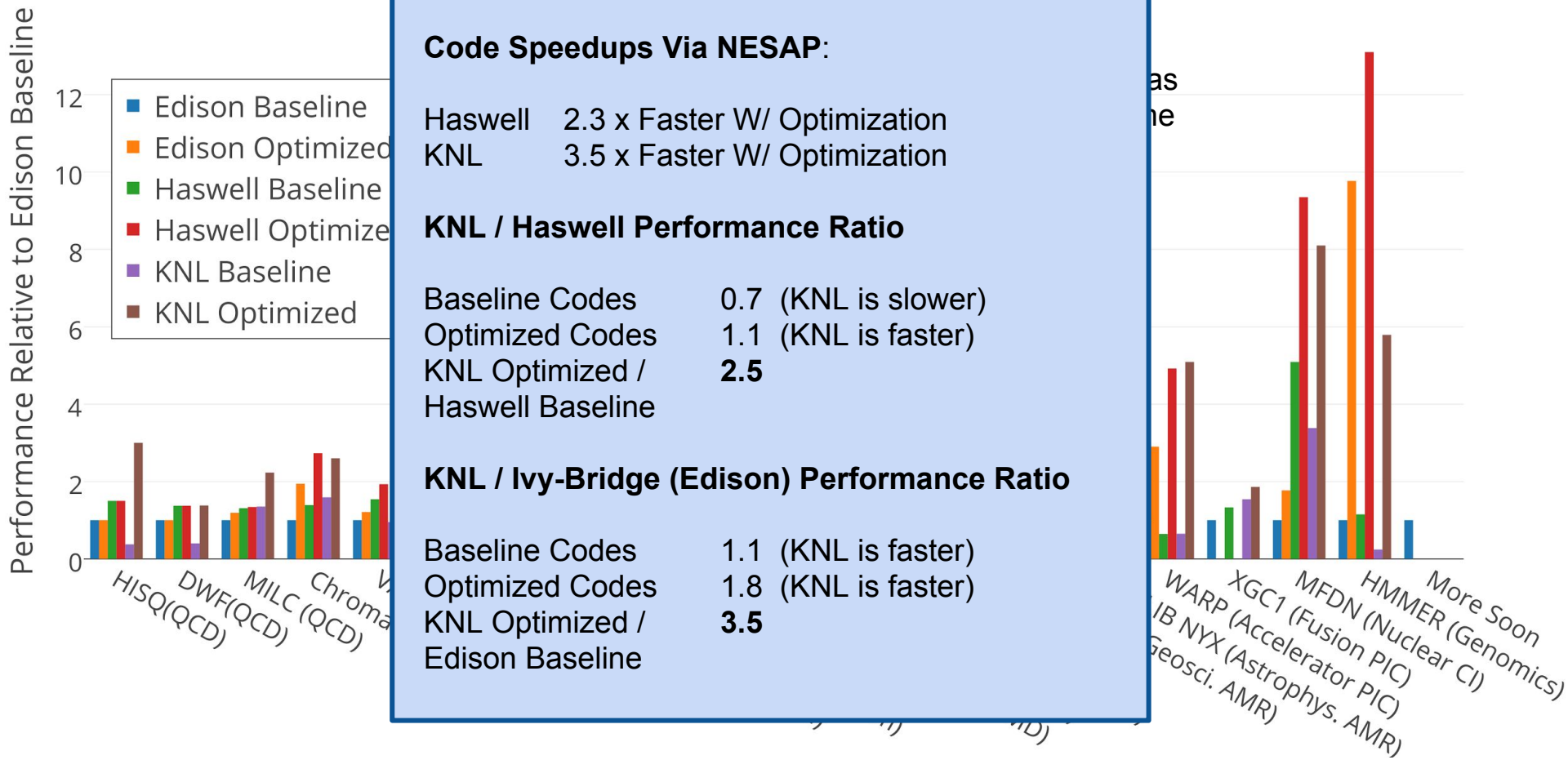
# NESAP Progress




U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB

# NESAP Progress




*PRELMINARY*

**Code Speedups Via NESAP**:

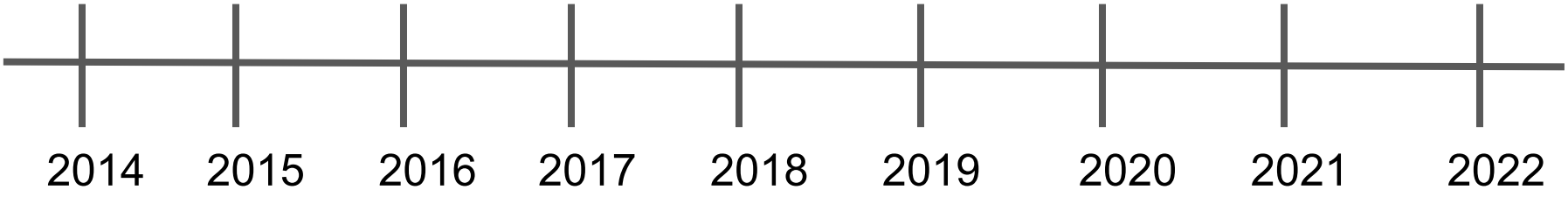
| | |
|---|---|
| Haswell | 2.3 x Faster W/ Optimization |
| KNL | 3.5 x Faster W/ Optimization |

**KNL / Haswell Performance Ratio**

| | |
|---|---|
| Baseline Codes | 0.7 (KNL is slower) |
| Optimized Codes | 1.1 (KNL is faster) |
| KNL Optimized / Haswell Baseline | **2.5** |

**KNL / Ivy-Bridge (Edison) Performance Ratio**

| | |
|---|---|
| Baseline Codes | 1.1 (KNL is faster) |
| Optimized Codes | 1.8 (KNL is faster) |
| KNL Optimized / Edison Baseline | **3.5** |

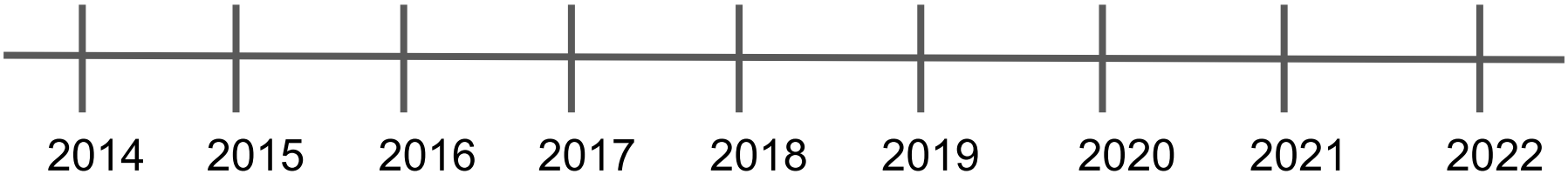



**NESAP 1**

?

**NESAP 2**

# NESAP Roadmap




**NESAP 1**

**NESAP For Data**

✓ **Initiated**

**NESAP++ Cori Large Science Run Support**

✓ **Initiated**

**NESAP++ Programming Models + Portability**

✓ **Initiated**

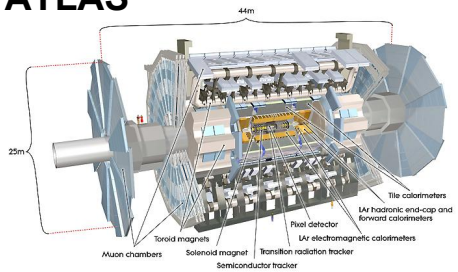
**NESAP 2**

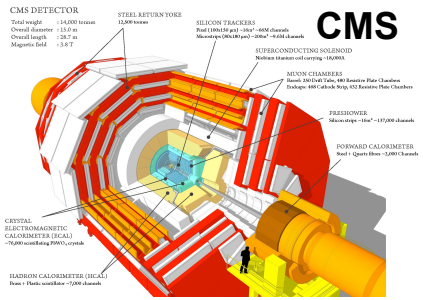
# NESAP for Data

- Dramatic advances in detectors, compute, storage, & networks are fueling growth in data-intensive DOE science.
- **Applications that process and analyze big data streams from experiments and instrumentation supported by DOE need help to prepare for exascale.**
- Teams get early/expanded access, vendor collaboration, extra support from NERSC.
- Proposal process for code teams:
  - CfP in October.
  - 6 selections in December (pictured).
- NESAP postdocs:
  - 1 postdoc hired at NERSC.
  - Interviewing for 2 more now.
- Code teams gathering initial performance data on KNL now.

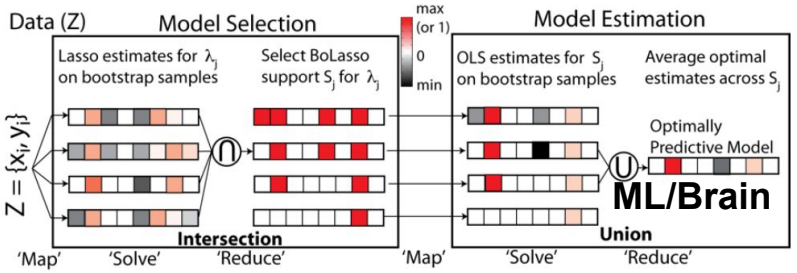


ATLAS








U.S. DEPARTMENT OF ENERGY

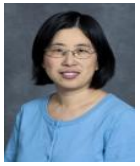
BERKELEY LAB
Lawrence Berkeley National Laboratory

# Getting a Jump Start on Exascale with NESAP++

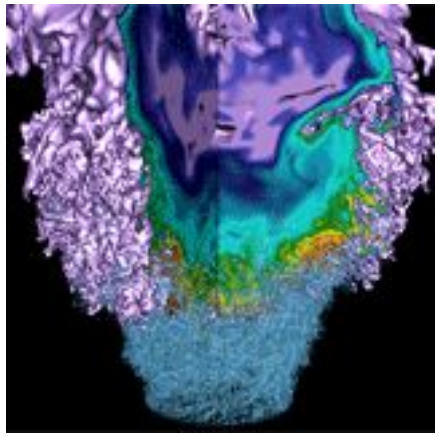
- Call released  (Jan 2017 and June 2017) to support large scale science runs on Cori. Partnering NERSC staff with NERSC apps that require running at the largest scales. Brandon Cook leading.
  - Supported 7 GB submissions and a couple other large jobs so far.
- Cross Lab Portability Effort Under way. Developing case studies with OpenMP 4.5, Kokkos, HPX etc.
- Planning partnership with Subset of NESAP teams to explore “exascale” programming models. Chroma - Balint Joo - will be a prototype for this effort. Summer 2017.
- **Are you looking into advanced programming models/libraries for performance and portability? We want to hear from you! Email jrdeslippe@lbl.gov**
- We are reaching out / making a difference on standards bodies:
  - Drafting/championing changes
  - Holding workshops for standards bodies
    - Advanced OpenMP Training & C++ 2 Day Workshop (Collaboration with Bryce CRD). Including Intel, Nvidia, Charm++, UPC++, Kokkos, HPX training sessions



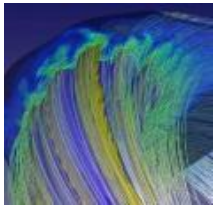
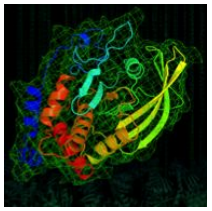
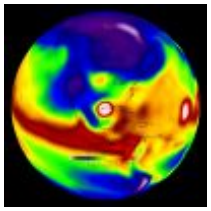
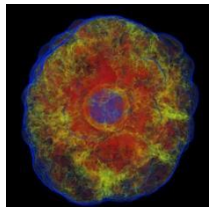
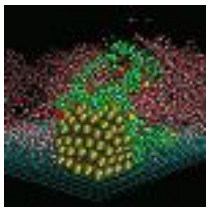



# Science at Scale CFP



NERSC

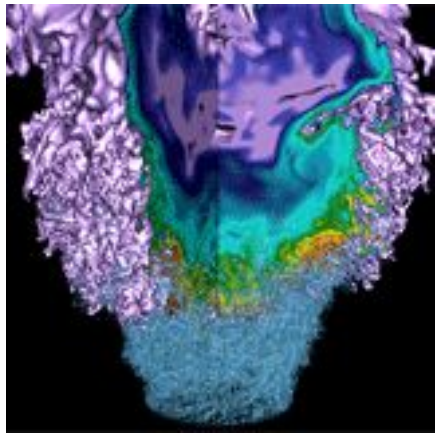
U.S. DEPARTMENT OF ENERGY | Office of Science

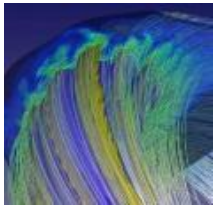
BERKELEY LAB
Lawrence Berkeley National Laboratory

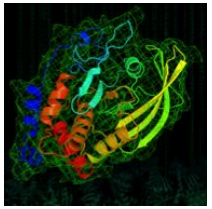
# Science at Scale CFP

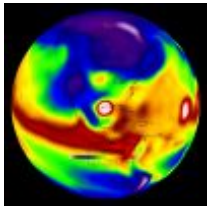
- Jan 2017: First call
  - Supported 7 Gordon Bell submissions
  - Several other large jobs
- June 2017: Updated with 400M NERSC hours
  - Rolling reviews/ acceptance
- Partnership with NERSC staff important for success
- Preference for proposals that require and utilize
  - large fraction of Cori (full machine) jobs
  - burst buffer
- https://goo.gl/forms/8BAUbafJJ6gu99YE3

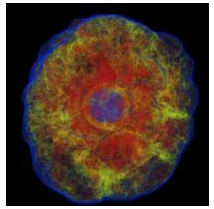
# Data Day and NUG Meeting plans

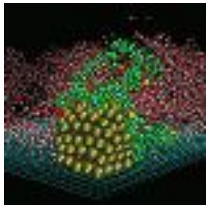












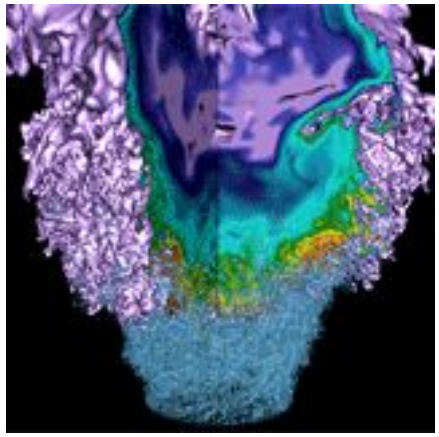
U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

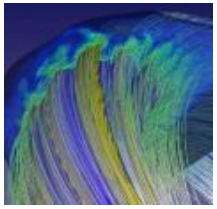
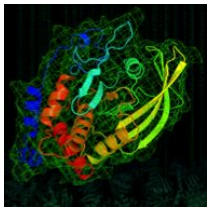
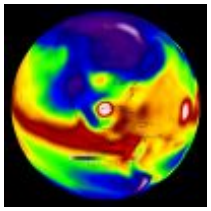
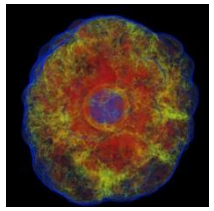
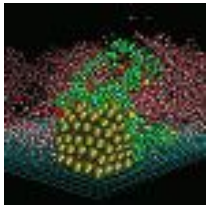
# Data Day + NUG17: Sept 19-21

- **Two popular NERSC annual events back-to-back**
- **Data Day**
  - 1.5 days covering Data-centric topics
    - Machine learning, workflows, data management, visualization…
    - 1 day of talks from scientists and demos from NERSC staff
    - ½ day "hackathon" of guided tutorials and general hands-on fun
- **NERSC Users Group Annual Meeting: NUG17**
  - The future of NERSC and HPC
  - Breakthrough science at NERSC
  - Getting the most out of Cori
- **Save the dates!**
- https://www.nersc.gov/users/NUG/annual-meetings/nersc-data-day-and-nug2017/

# ECP



NERSC

U.S. DEPARTMENT OF ENERGY | Office of Science

BERKELEY LAB
Lawrence Berkeley National Laboratory

# Exascale Computing Project

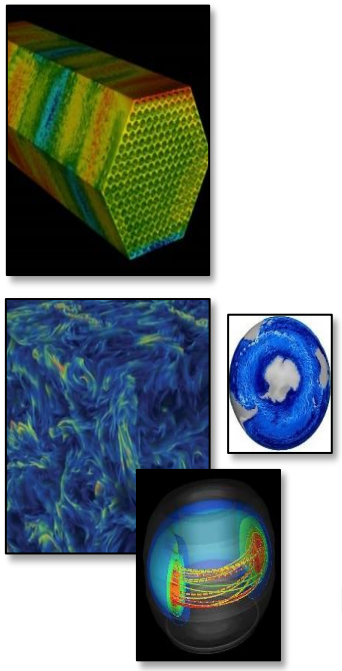
- As part of President Obama's National Strategic Computing initiative, ECP was established to accelerate delivery of a capable exascale computing system that integrates hardware and software capability to deliver approximately 50 times more performance than today's petaflop machines.
- ECP's work encompasses applications, system software, hardware technologies and architectures, and workforce development to meet the scientific and national security mission needs of DOE.
- About a $4 B 7-year project
  - Does not include system acquisition

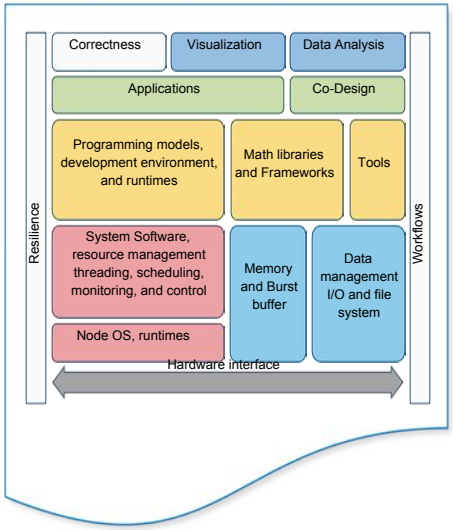
# Exascale Computing Project Goals

| | | | |
|---|---|---|---|
| Develop scientific, engineering, and large-data applications that exploit the emerging, exascale-era computational trends caused by the end of Dennard scaling and Moore's law | Create software that makes exascale systems usable by a wide variety of scientists and engineers across a range of applications | Enable by 2021 and 2023 at least two diverse computing platforms with up to 50× more computational capability than today's 20 PF systems, within a similar size, cost, and power footprint | Help ensure continued American leadership in architecture, software and applications to support scientific discovery, energy assurance, stockpile stewardship, and nonproliferation programs and policies |
| Foster application development | Ease of use | Rich exascale ecosystem | US HPC leadership |

# ECP has formulated a holistic approach that uses co-design and integration to achieve capable exascale

| Application Development | Software Technology | Hardware Technology | Exascale Systems |
|---|---|---|---|
| Science and mission applications | Scalable and productive software stack | Hardware technology elements | Integrated exascale supercomputers |






ECP's work encompasses applications, system software, hardware technologies and architectures, and workforce development

# What is a *capable* exascale computing system?

A capable exascale computing system requires an entire computational ecosystem that:

- Delivers 50× the performance of today's 20 PF systems, supporting applications that deliver high-fidelity solutions in less time and address problems of greater complexity
- Operates in a power envelope of 20–30 MW
- Is sufficiently resilient (perceived fault rate: ≤1/week)
- Includes a software stack that supports a broad spectrum of applications and workloads

This ecosystem will be developed using a co-design approach to deliver new software, applications, platforms, and computational science capabilities at heretofore unseen scale

# The ECP Plan of Record

- A 7-year project that follows the *holistic/co-design* approach, which runs through 2023 (including 12 months of schedule contingency)
  - To meet the ECP goals
- Enable an initial exascale system based on *advanced architecture* and delivered in 2021
- Enable *capable exascale* systems, based on ECP R&D, delivered in 2022 and deployed in 2023 as part of an NNSA and SC facility upgrades
- Acquisition of the exascale systems is outside of the ECP scope, will be carried out by DOE-SC and NNSA-ASC facilities

# NERSC Involved in all ECP areas

ECP will benefit NERSC users by developing technologies and methods for next-generation systems

- **Application Development**
  - Natural synergy with NESAP
  - Allocations support for developers
  - Joint training activities
- **Software Technologies**
  - Defining and building the support software for exascale
  - Allocations and infrastructure support for R&D
- **Exascale Systems**
  - ECP RFI
  - Technologies applicable to NERSC 10
- **Hardware Technologies**
  - Collaborations with vendors
  - Design space evaluations


NERSC