Welcome and NERSC Overview

New User Training and Updated Best Practices on Perlmutter
June 12, 2024

Rebecca Hartman-Baker, Lipi Gupta
User Engagement Group
Agenda

- Introduction to NERSC
  - Hardware
  - Software
- Interacting with NERSC Staff
- User Responsibilities & Expectations
Introduction to NERSC
About NERSC

● National Energy Research Scientific Computing Center
  ○ Established 1974, first unclassified supercomputer center
    ■ Original mission: to enable computational science as complement to magnetically controlled plasma experiment
  ○ 2024: NERSC 50th anniversary!
● Today’s mission: *Accelerate scientific discovery at the DOE Office of Science through High-Performance Computing and Extreme Data Analysis*
● NERSC is a national user facility
● NERSC is part of Berkeley Lab
NERSC: Mission HPC for DOE Office of Science Research

Largest funder of physical science research in U.S.

Allocations primarily controlled by DOE

- 80% DOE Annual production awards (ERCAP)
  - O(100)-O(10000) hour awards
  - Proposal-based, chosen by DOE program managers
- 10% DOE ASCR Leadership Computing Challenge
- 10% NERSC reserve
NERSC by the Numbers

NERSC USERS ACROSS US AND WORLD

50
States,
Washington D.C.
& Puerto Rico

53
Countries

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

32% Graduate Students
19% Postdoctoral Fellows
15% Staff Scientists
13% University Faculty
8% Undergraduate Students
5% Professional Staff
60% Universities
29% DOE Labs
5% Other Government Labs
4% Industry
1% Small Businesses<br><1% Private Labs

- ~1000 codes
- hundreds of users log in each day

2,500 Refereed Publications per Year
NERSC Users Produce Groundbreaking Science

**Materials Science**
Revealing Reclusive Mechanisms for Solar Cells  
NERSC PI: C. Van de Walle, UC Santa Barbara, *ACS Energy Letters*

**Earth Sciences**
Simulations Probe Antarctic Ice Vulnerability  
NERSC PIs: D. Martin, Berkeley Lab; E. Ng, Berkeley Lab; S. Price, LANL. *Geophysical Research Letters*

**High Energy Physics**
Shedding Light on Luminous Blue Variables  
NERSC PI: Yan-Fei Jiang, UC Santa Barbara. *Nature*

**Advanced Computing**
Scalable Machine Learning in HPC  
NERSC PI: L. Oliker, Berkeley Lab, 21st *International Conference on AI and Statistics*

**Plasma Physics**
Plasma Propulsion Systems for Satellites  
NERSC PI: I. Kaganovich, Princeton Plasma Physics Lab, *Physics of Plasmas*

**Nuclear Physics**
Enabling Science Discovery for STAR  
Nobel-Prize Winning Users

2013 Chemistry
Martin Karplus
for the development of multiscale models for complex chemical systems

2011 Physics
Saul Perlmutter
for the discovery of the accelerating expansion of the Universe through observations of distant supernovae

2006 Physics
George Smoot
for the discovery of the blackbody form and anisotropy of the cosmic microwave background radiation

2007 Peace
Warren Washington
for their efforts to build up and disseminate greater knowledge about man-made climate change

2017 Chemistry
Joachim Frank
for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution

2015 Physics
SNO Collaboration
for the discovery of neutrino oscillations, which shows that neutrinos have mass

NERSC

NERSC

BERKELEY LAB
Bringing Science Solutions to the World

U.S. DEPARTMENT OF ENERGY
Office of Science
Hardware
NERSC Systems Roadmap

NERSC-7: Edison
- Multicore CPU
2013

NERSC-8: Cori
- Manycore CPU
- NESAP Launched: transition applications to advanced architectures
2016

NERSC-9: CPU and GPU nodes
- Continued transition of applications and support for complex workflows
2021

NERSC-10: Exa system
2026

NERSC-11: Beyond Moore
2030+

Increasingly energy-efficient architectures
• Shared 2011 Nobel Prize in Physics for discovery of the accelerating expansion of the universe.

• Supernova Cosmology Project, lead by Perlmutter, was a pioneer in using NERSC supercomputers combine large scale simulations with experimental data analysis

• Login “saul.nersc.gov”
2 types of “blades”

Stack blades sideways into server racks
2 types of “blades”
The System is a Sum of Many Parts!

35 PB All-Flash Scratch

HPE Slingshot 11 interconnect
4 NICs/GPU node,
1 NIC/CPU node

>5 TB/s
The System is a Sum of Many Parts!

35 PB All-Flash Scratch

>5 TB/s

HPE Slingshot 11 interconnect
4 NICs/GPU node,
1 NIC/CPU node

3,072 CPU-only nodes
2 AMD “Milan” CPUs
1,536 TB CPU memory

memory memory

CPU CPU
The System is a Sum of Many Parts!

- 35 PB All-Flash Scratch
- >5 TB/s
- HPE Slingshot 11 interconnect
- 1,792 GPU-accelerated nodes
  - 4 NVIDIA A100 GPUs + 1 AMD “Milan” CPU
  - 448 TB (CPU) + 320 TB (GPU) memory
- 3,072 CPU-only nodes
  - 2 AMD “Milan” CPUs
  - 1,536 TB CPU memory
The System is a Sum of Many Parts!

- **35 PB All-Flash Scratch**
- **>5 TB/s**
- **1,792 GPU-accelerated nodes**
  - 4 NVIDIA A100 GPUs + 1 AMD “Milan” CPU
  - 448 TB (CPU) + 320 TB (GPU) memory
- **1,792 GPU-accelerated nodes**
  - 4 NVIDIA A100 GPUs + 1 AMD “Milan” CPU
  - 448 TB (CPU) + 320 TB (GPU) memory
- **3,072 CPU-only nodes**
  - 2 AMD “Milan” CPUs
  - 1,536 TB CPU memory
- **HPE Slingshot 11 interconnect**
  - 4 NICs/GPU node,
  - 1 NIC/CPU node
- **1.6 TB/s**
- **Ethernet LAN**
  - 2 x 400 Gb/s
  - 2 x 100 Gb/s
The System is a Sum of Many Parts!

- **35 PB All-Flash Scratch**
- **HPE Slingshot 11 interconnect**
  - 4 NICs/GPU node,
  - 1 NIC/CPU node

**3,072 CPU-only nodes**
- 2 AMD “Milan” CPUs
- 1,536 TB CPU memory

**1,792 GPU-accelerated nodes**
- 4 NVIDIA A100 GPUs + 1 AMD “Milan” CPU
- 448 TB (CPU) + 320 TB (GPU) memory

**Ethernet LAN**
- 2 x 400 Gb/s
- 2 x 100 Gb/s

**Off-Platform Storage**
- HPSS Tape Archive ~300 PB
- Common File System 130 PB
- /home 450 TB
- DTNs, Gateways
- edge services
Any Questions So Far?
The System is a Sum of Many Parts!

- **HPE Slingshot 11 interconnect**
  - >5 TB/s
  - All-Flash Scratch

- **3,072 CPU-only nodes**
  - 2 AMD “Milan” CPUs
  - 1,536 TB CPU memory

- **1,792 GPU-accelerated nodes**
  - 4 NVIDIA A100 GPUs + 1 AMD “Milan” CPU
  - 448 TB (CPU) + 320 TB (GPU) memory

**Off-Platform Storage**
- HPSS Tape Archive ~300 PB
- Common File System 130 PB
- /home 450 TB
- DTNs, Gateways
- edge services

**Ethernet LAN**
- 2 x 400 Gb/s
- 2 x 100 Gb/s

**NERSC**
- NERSC
- Energies Sciences Network

**BERKELEY LAB**
- Berkeley Lab
- Bringing Science Solutions to the World

**U.S. DEPARTMENT OF ENERGY**
- Office of Science
The System is a Sum of Many Parts!

35 PB All-Flash Scratch

HPE Slingshot 11 interconnect
4 NICs/GPU node,
1 NIC/CPU node

3,072 CPU-only nodes
2 AMD “Milan” CPUs
1,536 TB CPU memory

1,792 GPU-accelerated nodes
4 NVIDIA A100 GPUs+1 AMD “Milan” CPU
448 TB (CPU) + 320 TB (GPU) memory

Ethernet LAN
2 x 400 Gb/s
2 x 100 Gb/s

Off-Platform Storage

- HPSS Tape Archive ~300 PB
- Common File System 130 PB
- /home 450 TB

DTNs, Gateways

deltext edge services
Simplified NERSC File Systems

Memory (256-512 GB)

Performance

Capacity
Simplified NERSC File Systems

Performance

Memory (256-512 GB)

Scratch

Capacity

local file system
Simplified NERSC File Systems

- Performance
  - Memory (256-512 GB)
  - Scratch
  - Community File System

- Capacity
  - local file system
  - global file system
Simplified NERSC File Systems

- **Memory (256-512 GB)**
- **Scratch**
- **Community File System**
- **HPSS (Tape Archive)**

**Performance**

**Capacity**

- local file system
- global file system
- archival storage
Simplified NERSC File Systems

- Performance
  - Memory (256-512 GB)
  - Scratch
  - Community File System
  - HPSS (Tape Archive)

- Capacity
  - local file system
  - global file system
  - archival storage

Global Common
  - shared software

Global Home
  - individual home directories
Using NERSC File Systems

● Analogy:
  ○ Computing = baking
  ○ Input = baking ingredients
  ○ Output = cake

● NERSC is gigantic shared kitchen space with all the latest kitchen gadgets
  ○ Computers = ovens
  ○ Home, CFS = pantry, fridge
  ○ HPSS = freezer
  ○ Scratch = kitchen counter
Using NERSC File Systems

- When baking, stage ingredients from pantry and fridge (plus maybe rarely used ingredients from freezer) onto kitchen counter
  - Likewise, stage data and executable onto scratch file system

New Mexico. Mrs. Fidel Romero proudly exhibits her canned food, 1946 US National Archives NWDNS-33-S-12785
Using NERSC File Systems

- After baking, clean up after yourself!
- It’s okay to let cake cool on kitchen counter, but need to leave space clean for next user
  - After a while, we will clean up if you don’t, but not like you would want
  - We will throw all your materials in the trash (even your cake!)

Queen cakes cooling on a wire rack by James Petts
https://www.flickr.com/photos/14730981@N08/1347333725/
Software
Software

Operating System: manages the computer’s memory, process as well as hardware and software

Perlmutter OS is a version of Linux, offered by Cray (vendor) for their supercomputers

Some of your applications may not run out-of-the-box!
Software

Operating System: manages the computer’s memory, process as well as hardware and software

Perlmutter OS is a version of Linux, offered by Cray (vendor) for their supercomputers

Some of your applications may not run out-of-the-box!

Solution:

- Use NERSC-provided compilers and libraries
- Use the provided application (via modules)
- Use Extreme-scale Scientific Software Stack (E4S) via the Spack package manager
More than 13.5 million lines of source code Compiled, Optimized, and Tested
NERSC has a rich data ecosystem!

- **data transfer and access**
  - Globus Online
  - Jupyter

- **data analytics**
  - R
  - Julia

- **data management**
  - MongoDB
  - MySQL

- **machine learning**
  - scikit-learn
  - PyTorch

- **visualization**
  - VisIt
  - ParaView

- **containers**
  - SHIFTER
  - Spin

- **workflows**
  - FireWorks

- **GNU parallel**

- **machine learning**

- **visualization**

- **containers**

- **workflows**

- **data analytics**

- **data management**

- **data transfer and access**
Perlmutter Supports Multiple Compilers and Every GPU Programming Model

<table>
<thead>
<tr>
<th></th>
<th>Fortran/C/C++</th>
<th>CUDA</th>
<th>OpenACC 2.x</th>
<th>OpenMP 5.x</th>
<th>CUDA Fortran</th>
<th>Kokkos / Raja</th>
<th>MPI</th>
<th>HIP</th>
<th>DPC++ / SYCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVIDIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNU</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLVM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vendor Supported

NERSC Supported
Interacting with NERSC

● NERSC User Engagement Group
  ○ User Engagement, outreach
  ○ User Training and Documentation

● NERSC Consulting & Account Support
  ○ User Tickets
  ○ User Appointments

● NERSC Operations
User Engagement Group - Our People

Lisa Claus
Kevin Gott
Lipi Gupta
Rebecca Hartman-Baker
UES Group Lead

Helen He
Charles Lively
Kelly Rowland
Kadidia Konate
Woo-Sun Yang

Alumni:
Tiffany Connors
Zhengji Zhao
Steve Leak
Erik Palmer
Justin Cook
Shahzeb Siddiqui
NERSC User Group (NUG)

- Community of NERSC users
- Source of advice and feedback for NERSC (we listen!)
- Regular teleconferences hosted by NERSC
- Join the NUG Slack: https://www.nersc.gov/users/NUG/nersc-users-slack/
- Join us NUG Annual Meeting (Oct 22-24, 2024) for our 50th anniversary!
NERSC User Training

● NERSC provides a robust training program for users of all skill levels, interests, and personas
  ○ All trainings are recorded, professionally captioned, & posted to NERSC YouTube channel
  ○ Slides posted to training event webpage

● For more information on upcoming and past events, see https://www.nersc.gov/users/training/events/

● Collection of Categorized Training Materials
● Training Events Archive
Interacting with NERSC

- NERSC User Engagement Group
  - User Engagement, outreach
  - User Training and Documentation

- NERSC Consulting & Account Support
  - User Tickets
  - User Appointments

- NERSC Operations
Consulting & Account Support Team
What is a “Help Ticket”?

https://help.nersc.gov
What is a “Help Ticket”?

https://help.nersc.gov
NERSC Consulting: Tips & Tricks

● Help us help you!
● Provide specifics:
  ○ What is the problem?
  ○ What machine? Which node? Which file path?
  ○ When did it happen?
  ○ What modules were loaded?
  ○ How did you try to fix or work around it?
  ○ How can I reproduce the problem?
● Tips for filing a good ticket:
  https://docs.nersc.gov/getting-started/#how-to-file-a-good-ticket
NERSC Consulting: Expectations

- Our first response will be within four business-hours
  - Monday - Friday from 8am - 5pm. Weekend tickets answered on Monday morning
- We will help you resolve your problem, and keep you apprised of progress
- We will attempt to accommodate user needs that don’t fit within our operating structure
- We welcome user feedback and constructive criticism
NERSC User Appointments

- In 2018, we began offering “office hours”
  - Open Zoom meeting which users could join to get help with a particular topic, e.g., MFA, KNL Optimization, ERCAP, etc.
  - Shortcoming: long periods with no participants, then many jump on simultaneously
- Appointments: more efficient use of everyone’s time
- 30-minute appointments offered on a variety of topics:
  - GPU basics, Optimization, File Systems, Using GPUs in Python, Containers, NERSC 101, Checkpoint/Restart jobs with MANA, Spin, Appentra Codee
- Schedule an appointment: nersc.as.me
Interacting with NERSC

- NERSC User Engagement Group
  - User Engagement, outreach
  - User Training and Documentation
- NERSC Consulting & Account Support
  - User Tickets
  - User Appointments
- NERSC Operations
NERSC Operations

- Operations staff are on site 24/7/365 to supervise operation of the machine room.
- Operations know the health of the machines and can help users with some tasks (killing jobs, changes to running reservation, etc.).
- **NERSC MOTD** (message of the day, live status)
- Please avoid contacting Operations except in urgent cases.
Any Questions So Far?
User Expectations
User Responsibilities & Expectations

- Be kind to your neighbor users
  - Don’t abuse the shared resources!
User Responsibilities & Expectations

● Be kind to your neighbor users
  ○ Don’t abuse the shared resources!

● Use your allocation smartly
  ○ Pick the right resource for your job and your data
User Responsibilities & Expectations

- Be kind to your neighbor users
  - Don’t abuse the shared resources!
- Use your allocation smartly
  - Pick the right resource for your job and your data
- Back your stuff up
  - Especially from scratch, which has a purge policy
User Responsibilities & Expectations

● Be kind to your neighbor users
  ○ Don’t abuse the shared resources!

● Use your allocation smartly
  ○ Pick the right resource for your job and your data

● Back your stuff up
  ○ Especially from scratch, which has a purge policy

● Acknowledge NERSC in your papers
  ○ Acknowledge us so we can stay in business!
User Responsibilities & Expectations

● Be kind to your neighbor users
  ○ Don’t abuse the shared resources!

● Use your allocation smartly
  ○ Pick the right resource for your job and your data

● Back your stuff up
  ○ Especially from scratch, which has a purge policy

● Acknowledge NERSC in your papers
  ○ Acknowledge us so we can stay in business!

● Pay attention to security
  ○ Don’t share your account with others!
Thank You and Welcome to NERSC!