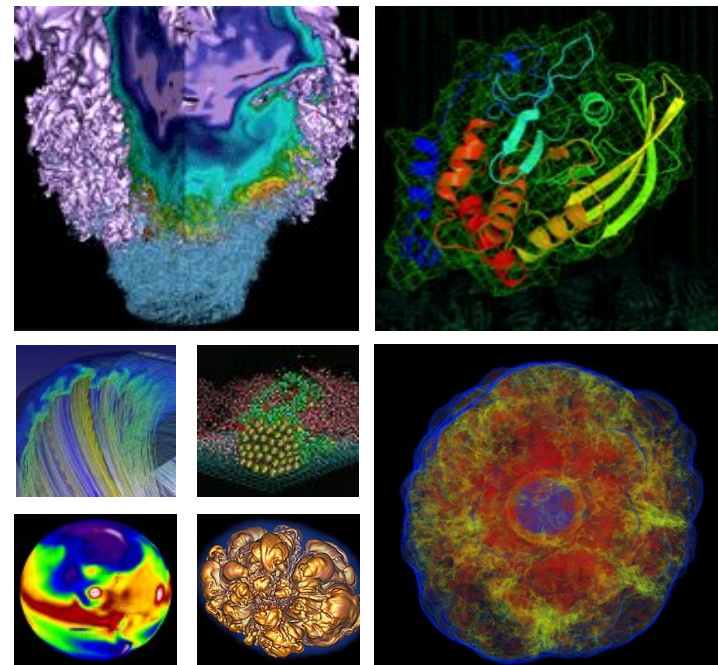


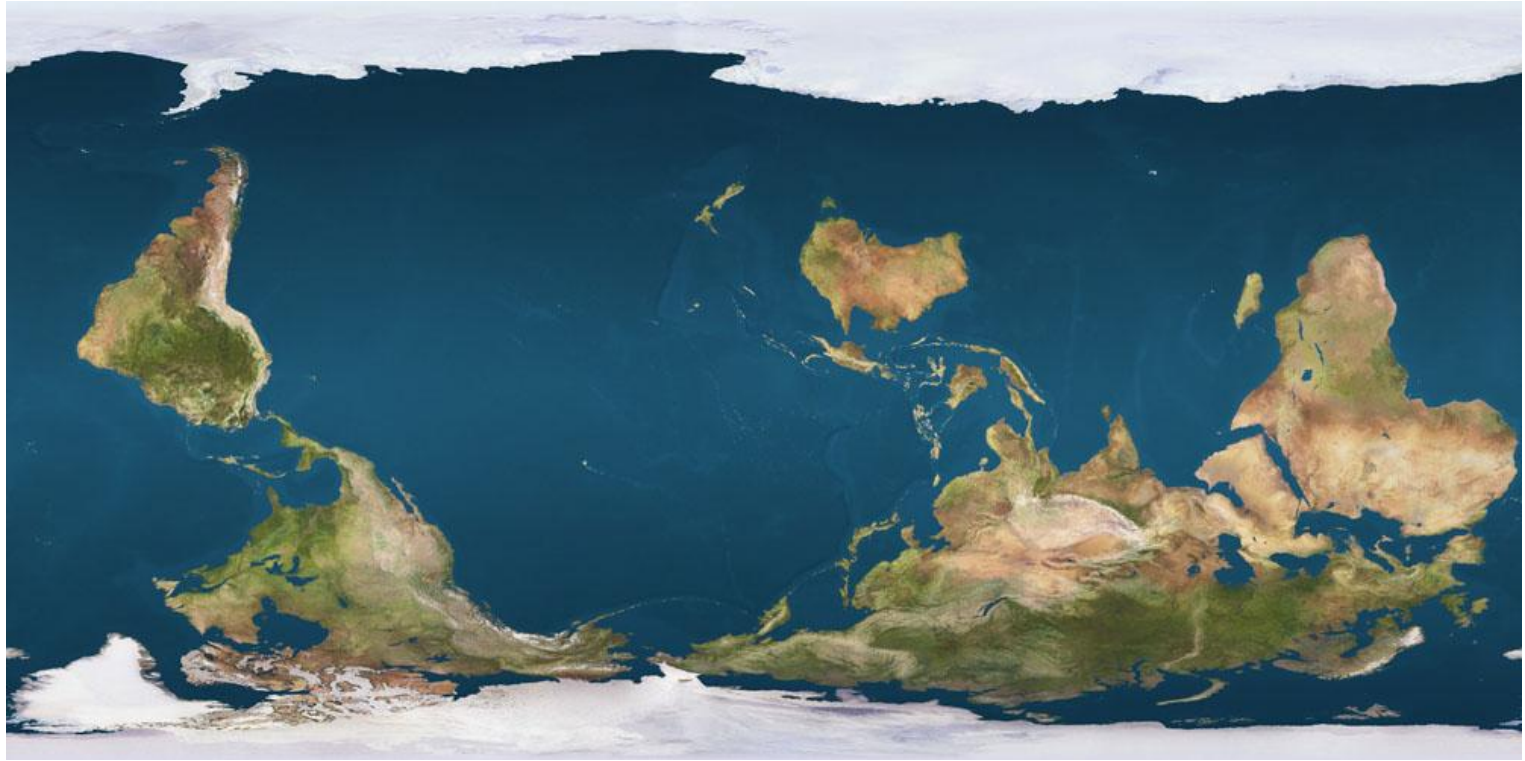
# NERSC Site Report



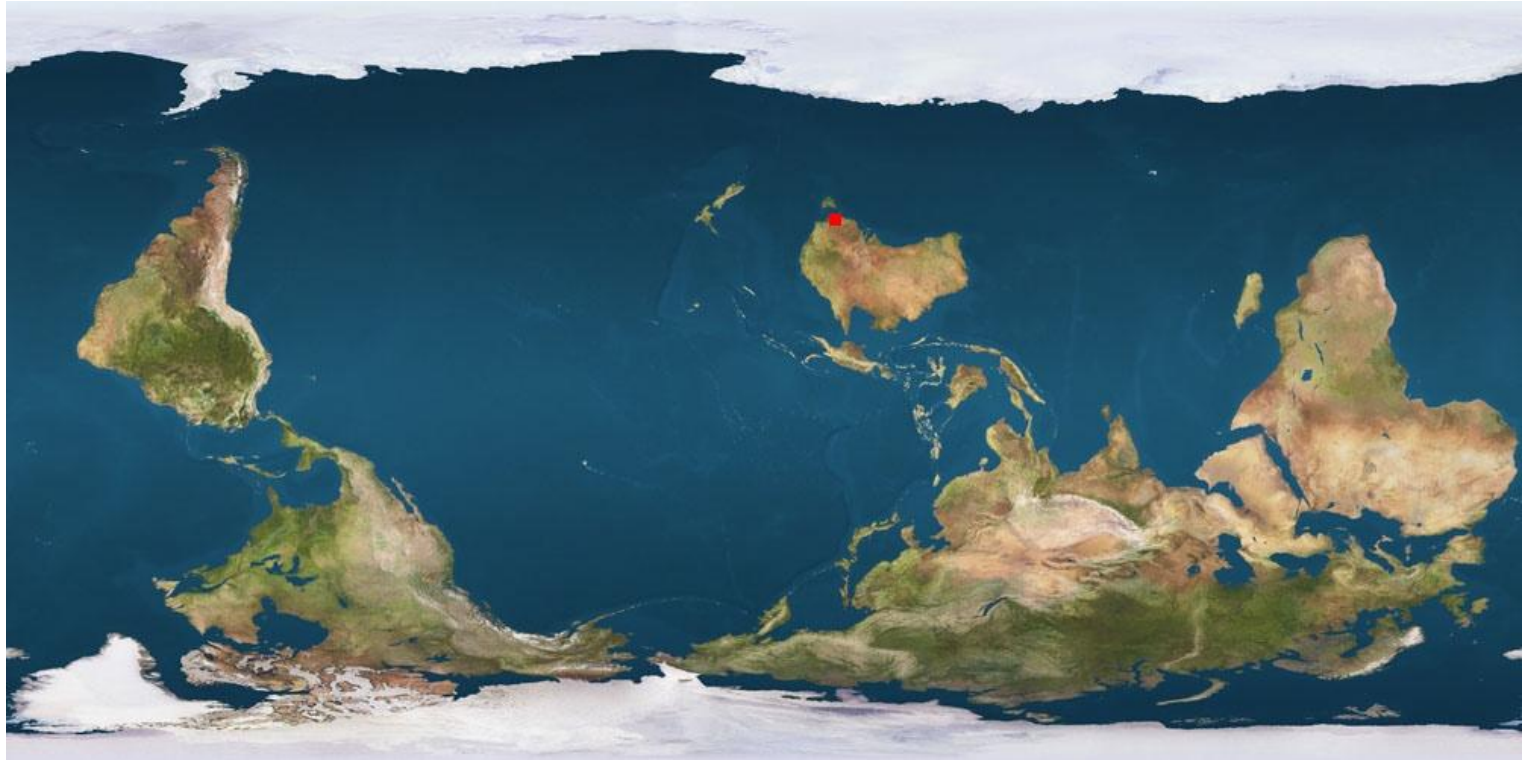
**Chris Samuel**

**Slurm User Group 2019**

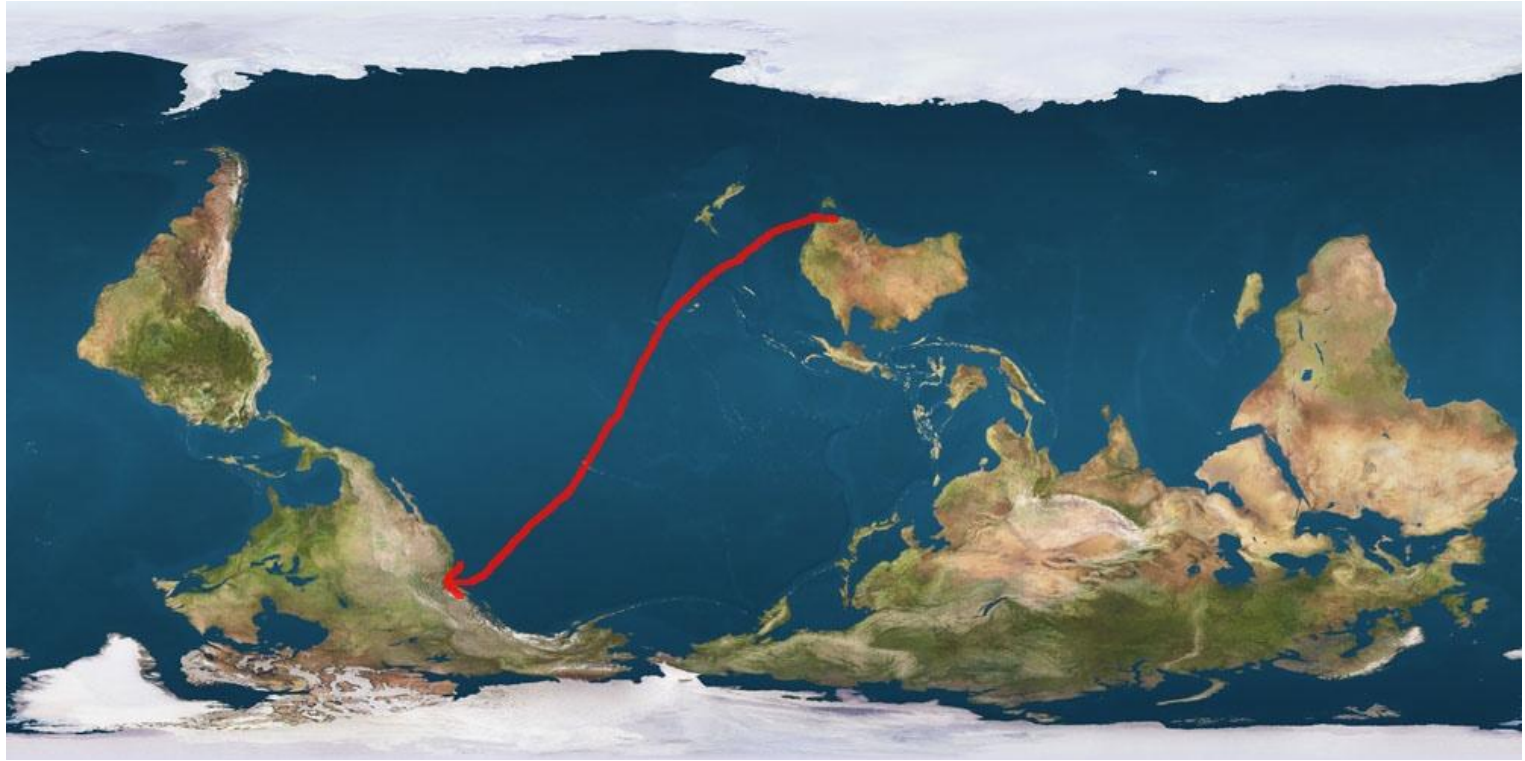
# Wait, weren't you in Australia?



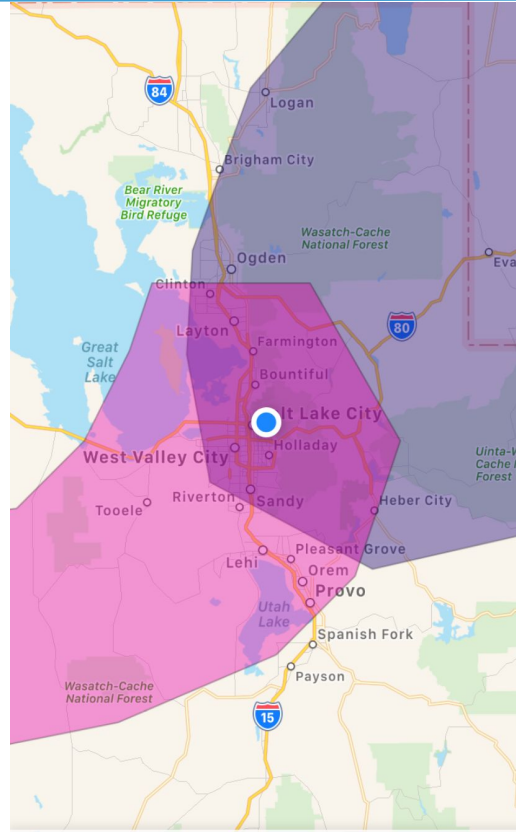
# Wait, weren't you in Australia?



# Wait, weren't you in Australia?



# Acknowledgement of Country



- [Eastern Shoshone \(Link\)](#)
- [Goshute \(Link\)](#)



# NERSC - Who/What/Where?



- National Energy Research Scientific Computing Center (NERSC)
  - Primary scientific computing facility for DoE Office of Science
  - Based at Lawrence Berkeley National Lab with an awesome view
- 45 years of supercomputing!
  - Started off with CDCs, first Cray-1 (serial #6) in 1978, also had IBM and Sun systems.
  - NERSC C90 #19 on first Top500 list in June 1993
- Cori (N8) - current system - Cray XC40
  - #5 Top500 Nov 2016, #14 June 2019
  - 2,388 HSW nodes + 9,688 KNL nodes
  - Slurm 19.05.2 (plus patches)
- Edison (N7) has been decommissioned - more later..

# #BestViewFromALab





# NERSC CSG approach to HPC



- Always have a test system
- All changes go through git and always on the test system first
  - Except in extenuating circumstances
  - Same git repos for all clusters
  - Repos for Cray ansible config, Slurm config and RPMs
  - User reframe for testing on TDS and Cori (thanks CSCS!)
- We pay for features in Slurm, so we tend to track releases/updates
  - Also not shy about applying extra patches
  - Always interested in performance improvements!
- We try to open source things we believe will be useful
  - Shifter for containers
  - Coming soon (by SC19) - submit filter library and CLI filter examples

# HPC at Scale - what does it mean?



- Cori has ~12,000 compute nodes, 300+ different filesystem related nodes (Lnet, DVS & datawarp), 250+ Lustre servers, 23 login nodes, 18 GPU nodes, 20 high memory nodes for Joint Genome Institute users.
- Two separate Slurm clusters, three sets of Slurm RPMs
- ~10K-20K+ jobs waiting to run at any one time
  - Backfill limited to 10 minutes, lots of work to reduce time spent here
  - Customised `cons_res` for exact scheduling (needed for large jobs)
  - Patch from 20.02 looks promising in reducing unnecessary testing
- ~15K-40K jobs a day ranging from 9,000+ nodes to a few cores.
  - Most compute nodes reserved for exclusive jobs
  - KNL only takes exclusive jobs
- Target 20% of KNL jobs should be at 1024 nodes or larger
- LDMS for monitoring Cori - it writes ~5TB of monitoring data a day

# HPC at Scale - If it can go wrong it wi



- FLOSS maxim - “with enough eyes all bugs become shallow”
- HPC maxim - “with enough compute nodes all bugs become manifest”
- Hardware peculiarities
  - What do do if your Cray HSN goes away to reroute for 50 seconds & you can’t talk to anyone? (Slurm knows how to spot this)
- Software idiosyncrasies
  - Have race condition(s) in 19.05 which seems to affect sbcast, message aggregation & prologs
  - Need to reboot nodes for (eg) hugepage recovery without disrupting scheduling for large jobs - added in 18.08.x release (NERSC funded)
- Network issues
  - We have load balancers in front of LDAP servers because when a 9,000 node job starts (more on that later)...

- Projects (“repos”) & their users have separate balances and can only submit or run whilst both are positive (managed outside of Slurm via submit filter & access to QOSs)
- Priority threshold before a job can get a forward reservation on resources
  - `bf_min_prio_reserve=69121`
  - Priority increments ~1 point per minute (`MaxJobsAccruePerUser=2`)
  - QOS sets initial priority at a set number of hours before threshold
    - Premium - immediate (but double the cost)
    - Large jobs - 6 hours short,  $\geq 1024$  KNL nodes, 50% discount
    - Regular - 3 days short
    - Low - 6 days short (but also get a discount)
    - Scavenger - 0 priority (backfill only) `--time-min <= 4` hours
- Jobs in reservations diverted to their own partition to avoid constraints

# Submit filter - gateway and guide



- NERSC makes extensive use of the submit filter
  - Framework written in Lua
  - Reads a YAML file with the rules/policy to enforce
  - Same YAML file also used to set rules on QOS's via cron job
- Gateway
  - Enforce architecture/time limit/node counts/etc for a QOS
  - Require `--time-min` for some QOS's
- Guide
  - User asks for the regular QOS with 2048 node KNL job
  - QOS gets rewritten to the appropriate internal QOS for discount
- Presented by Doug at SLUG 2018
  - [https://slurm.schedmd.com/SLUG18/lib\\_job\\_submit.pdf](https://slurm.schedmd.com/SLUG18/lib_job_submit.pdf)
  - Library intended to be released by SC19.

# #BestViewFromALab



# nss\_slurm - diverting the thundering herd



- ~12,000 compute nodes with ~25,000 jobs per day means an awful lot of LDAP queries, even with sssd/nslcd/etc caches
  - Large jobs generate a synchronised storm of queries
- NERSC funded work in 19.05 to create nss\_slurm (thanks Tim!)
- Documented here: [https://slurm.schedmd.com/nss\\_slurm.html](https://slurm.schedmd.com/nss_slurm.html)
- Provides an extra NSS module to query local slurmd for information
  - Requires modifying `/etc/nsswitch.conf` after checking config
- Needs “`LaunchParameters=enable_nss_slurm`” in `slurm.conf`
- NSS for a users job on the local node serviced quickly & reliably
- Reduces network traffic across the Cray HSN to the LDAP servers.
- Reduces strain on LDAP administrators :-)

# cli\_filter - filtering before the submit filter



- <Australian> How good is the submit filter! </Australian>
- But wouldn't it be nice to do checking in user commands first?
  - Catch problems before they hit slurmctld (reduce load on it)
  - Safely run longer running commands (like quota checks)
  - Modify/manipulate options before submission
  - Implement your own logging of user submissions
- Initially developed by Doug @ NERSC, integrated with refactoring work by SchedMD, included in 19.05, handy documents:
  - [https://slurm.schedmd.com/SLUG17/cli\\_filter-slug2017/cli\\_filter.html](https://slurm.schedmd.com/SLUG17/cli_filter-slug2017/cli_filter.html)
  - [https://slurm.schedmd.com/cli\\_filter\\_plugins.html](https://slurm.schedmd.com/cli_filter_plugins.html)
- Doug aiming to release sample code for Slurm contrib by SC19



# “Superfacility”: big science, right now



- Integrative HPC / HPC as a component of an experimental pipeline.
- Close coupling and support of HPC with experimental facilities. E.g.:
  - NCEM - want to stream ~360Gb/s for processing into Cori
  - SLAC - realtime processing of experiment data for operator feedback
  - DESI - ingest & process observations during daytime & provide feedback for next nights observing
- Requirements characterised as realtime response or deadline based
- Tricky balancing needs for utilisation & large jobs with these requirements
- Proposing NRE work for Slurm 20.02 to enable more flexible ways of nominating jobs for preemption that meshes with our existing maze of QOS's & partitions (112 QOS's & 14 partitions yesterday).

# This space intentionally left blank



# N9: “Perlmutter”

---



# N9: “Perlmutter”



# N9: “Perlmutter”



- Cray Shasta system (containers & kubernetes)
- Next generation AMD Milan CPUs and NVIDIA GPUs
- GPU nodes and CPU only nodes
- Cray Slingshot interconnect - low latency & ethernet compatible
- All Flash Lustre scratch filesystem
- CPU only cabinets to provide similar capability as Cori
- Will run Slurm

# We will be hiring...



Office of  
Science



# We have dinosaurs!





**Thank You**



U.S. DEPARTMENT OF  
**ENERGY**

Office of  
Science

