



**75** **Los Alamos**  
NATIONAL LABORATORY  
— EST. 1943 —

# Los Alamos National Laboratory Site Report

One Year Post Migration

## Slurm Users Group 2018



Joseph 'Joshi' Fullop

26 September 2018



Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

# Current Systems

System	Node Count
Lightshow	16
Fog	32
Kodiak	66
ViewMaster II	92
Gadget	<100
Pinto	154
Woodchuck	190
Trinitite	200
Snow	368
Badger	372
Wolf	616
Fire	1,104
Ice	1,104
Grizzly	1,490
Trinity	20,000

# Migration to Slurm mid 2017

- **Transition Issues**

- Binding Affinity
  - Getting the combination of the job request and MPI to deliver the desired layouts
- Dependency Changes
  - Code rebuilds, Packaging, Configuration Management
- User Education
  - Training Classes
  - JIT Training scripts
    - NO wrapper scripts!

- **Integration Efforts**

- JobComp/script creation and Slurm modifications
- YETI: Multiple security zones
- SDBM: Synchronization of accounts/projects/allocations/etc. with LDAP

# Hurdles

- **Large number of node state changes flooded slurmdbd and crashed slurmctld**
- **DataWarp staging**
  - Needed implementation
  - SchedMD and Cray collaborative effort
- **OOM Message deliverance**
  - cgroup work
  - Also a collaboration of SchedMD and Cray (fun with hugepages)
- **Fairshare Edge Cases**
  - Accounts with lots of users and equal weights (1)
  - Fairshare = Parent helped
- **Jobs stuck in COMPLETING**
  - Potentially already fixed.
- **Jobs not being considered for launch**
  - Found during resiliency code of relaunching a node-failed job with a spare already in the allocation.
  - May be the same root cause as a rare occurrence that we were not able to reproduce.
  - Unstick the job by changing an attribute (update to set an attribute to the same thing does not work)

# User Feedback

- **Immediately High Utilization**
  - Good Training & Documentation
- **Efficient Scheduling**
  - We had to re-tune a number of things on the system
- **Very Fast Response Times**
  - Some patches were same day.
- **Syntax varies between command line tools**
  - (scontrol, srun, salloc) vs (sacct, sacctmgr, sreport)

# Upcoming Investigations

- **Federated Scheduling**
  - Special purpose clusters
- **Database Co-Location / Single slurmdbd**
- **File system aware scheduling**
- **PMIx**
- **Workflow Visualization/Monitoring**



# Upcoming features we are anxiously awaiting

- **True Heterogeneous jobs**
  - Trinity is half Haswell and half KNL architecture.
- **TRES scheduling**
- **Priority Accrual**
  - Limits on the number of jobs that grow in priority over time per user.
  - Some users flood the queue and overwhelm others with age points, even with caps

# Wishlist

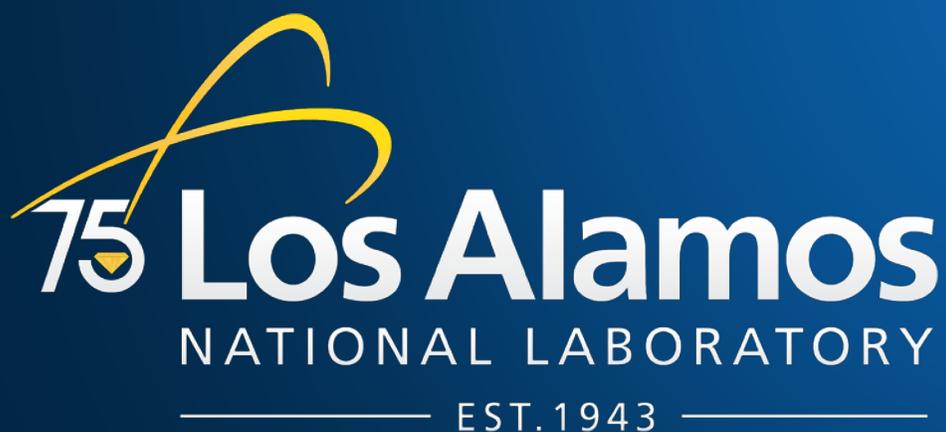
- **Account for which scheduling method a job was launched.**
  - Main vs Backfill vs Submit
  - Very valuable in mitigating users complaining about utilization of others
- **Accounting for last reason a job was blocked**
  - Allows us to monitor and appropriately modify policy
- **Accounting for which constraints a job required**
  - Searchable and available to jobcomp/script
- **Sortable SPRIO**
  
- **Job dependencies in federated scheduling**
  - For managing workflow utilizing specialized clusters

# Wishlist continued

- **Accrued Eligible Time**
  - How much time was a job eligible vs (submit – start).
- **scancel propagation for stage-out (DataWarp)**
  - Especially useful for failed/aborted jobs.
- **Current Backfill Opportunity**
  - “If I were to submit a job of X size, how long could it run?”
- **Job Pending state for nodes**
  - sinfo would show how many nodes are waiting on enough other nodes to launch the next priority job.
  - Disambiguate ‘idle’.
- **Auto sbcast of dependent files**
  - With obvious limitations.

# The End

## Slurm Users Group 2018



**Joseph 'Joshi' Fullop**

fullop@lanl.gov



Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA