From Moab to Slurm: 12 HPC Systems in 2 Months

Peltz, Fullop, Jennings, Senator, Grunau

Tuesday, 26 September 2017
Where we started

- Multiple systems with various operating systems and architectures
- Moab as the primary scheduler
- Resource Managers included:
  - PBS
  - PBS/Alps
  - Slurm (version 2.3) – Dictated by TOSS
Preparations

- Test & Development systems
  - Kit, Cub, VMs
- Scale Profiling on decommissioned cluster – Mustang
- Trip to NERSC
- Early Hybrid system - Woodchuck
# Systems List

<table>
<thead>
<tr>
<th>System</th>
<th>OS</th>
<th>Zone</th>
<th>Date</th>
<th>Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mustang</td>
<td>TOSS2</td>
<td>Open</td>
<td>Jan 12</td>
<td>1,500</td>
</tr>
<tr>
<td>Woodchuck</td>
<td>TOSS3</td>
<td>Open</td>
<td>Mar 8</td>
<td>226</td>
</tr>
<tr>
<td>Pinto</td>
<td>TOSS2</td>
<td>Open</td>
<td>Apr 10</td>
<td>154</td>
</tr>
<tr>
<td>Trinitite</td>
<td>CLE</td>
<td>Open</td>
<td>Apr 21</td>
<td>100</td>
</tr>
<tr>
<td>Snow</td>
<td>TOSS3</td>
<td>Open</td>
<td>Apr 24</td>
<td>368</td>
</tr>
<tr>
<td>Moonlight</td>
<td>TOSS2</td>
<td>Open</td>
<td>May 1</td>
<td>308</td>
</tr>
<tr>
<td>Ice</td>
<td>TOSS3</td>
<td>Secure</td>
<td>May 8</td>
<td>1,104</td>
</tr>
<tr>
<td>Grizzly</td>
<td>TOSS3</td>
<td>Open</td>
<td>May 15</td>
<td>1,490</td>
</tr>
<tr>
<td>Wolf</td>
<td>TOSS3</td>
<td>Open</td>
<td>May 17</td>
<td>616</td>
</tr>
<tr>
<td>Lightshow</td>
<td>TOSS2</td>
<td>Open</td>
<td>May 22</td>
<td>16</td>
</tr>
<tr>
<td>Fire</td>
<td>TOSS3</td>
<td>Secure</td>
<td>May 24</td>
<td>1,104</td>
</tr>
<tr>
<td>Luna</td>
<td>TOSS2</td>
<td>Secure</td>
<td>May 30</td>
<td>1,540</td>
</tr>
<tr>
<td>Viewmaster II</td>
<td>TOSS2</td>
<td>Secure</td>
<td>May 31</td>
<td>60</td>
</tr>
<tr>
<td>Trinity</td>
<td>CLE</td>
<td>Secure</td>
<td>June 1</td>
<td>19,364</td>
</tr>
</tbody>
</table>
Accounting Integration

- **Allocations: LDAP to Slurm**
  - `sacctmgr load file...`
    - Additive only
    - Adds may fail and derail
  - Finding discrepancies & issuing `sacctmgr` commands
    - May fail silently

- **Accounting: Slurm to HPCStats**
  - Join `slurm_jobcomp` with `slurm_acct_db` tables to generate csv
  - Import to HPCStats
Implementation

- **Ansible**
  - Developed playbooks
    - Removal of Moab/PBS
    - Installation of Slurm & databases

- **Interstate – Tool for managing Dedicated Service Time (DST)**
  - Stop Slurm services
  - Generate maintenance reservation(s)
  - Revert above for Return To Service

- Test Rollout/Rollback during regular DSTs
User Education

- Self-presented training for users
- Official SchedMD training for Admins and Users
- Just In Time training scripts
  - Created and used these instead of wrapper scripts
  - Tells users to use the equivalent command instead
- Consulting Staff
General Hurdles

- Firewall configs
  - Functioned, but did not scale well.
  - SYN Flood protection
- FastSchedule / Memory Specification
  - Opted to not use Memory as part of CR selection type
  - Exclusive use allocations
- Bind Affinity
  - Fun with hyper-threads and NUMA domains
- Job Accounting
Architecture on Cray

- Planning for Scale
  - Slurmctld on a unused service node rather than SDB
  - Backup slurm controller/slurmd/mysql on external node
  - Test boots into slurm on Trinity
    - Find bottlenecks and scaling issues
    - Debug level too high
    - Slurmctld log file needs to be on local file system to slurmctld node
    - Cray’s NHC was extremely slow at cleaning up jobs
    - Various timeout settings
Cray Scaling Issues

- Parallel File System
- PMI
- bcast
  - Temporary solution to help synchronize launches
Cray Customizations & Tuning

- MCDRAM
- TCP Timeouts
- Parallel srun
  - Specify GRES:craynetwork=0
- VM Overcommit
- DataWarp
- Re-provisioning of KNL nodes
Problems Encountered Specific to Cray

- `ncmd`
  - TCP/IP communications break down
- `aeld`
- `capmc`
  - Slurm does not see correct node features
  - Leads to job denial after submit script supplies default features
- `mapn`
  - No parallel `sruns`
Post-Migration Cleanup

- Config Consolidation
  - CR_Select type
  - Host Definitions