

Lawrence Livermore National Laboratory

Operations in the US



Jeff Cunningham
GLCC Project Lead

Lawrence Livermore National Laboratory, P. O. Box 808, Livermore, CA 94551

This work performed under the auspices of the U.S. Department of Energy by
Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.



Linux HPC at LLNL - Livermore Computing

- Very Large Center: ~14k nodes in ~30 clusters, ~2PFLOP/s
- All run same software stack based on RHEL 5 (soon RHEL 6)
- All worker nodes are diskless
 - Requires less hardware maintenance
 - Consistent OS image across whole cluster
 - Reboot returns node to pristine state
- All configuration comes from central subversion server using CFEngine
 - Consistent configuration across clusters/centers
 - Setting up new cluster is streamlined
- All files outside of configuration live in packages (RPM/deb)
- LC proper = ~125 staff (admin, ops and dev)
 - Standard practices and good documentation are essential.



GLCC at Livermore

- Running the CREAMCE/PBS for AliEn grid services
- Also on the Open Science Grid w/ PBS scheduler, using same PBS partition
- Worker nodes are running diskless
 - Robust Network and NFS infrastructure
 - Requires large amount of Physical RAM, since no swap.
- Security profile only allows LLNL employee logins from on-site.
 - No remote gsissh logins, so can only be managed by local staff.

- Compute Statistics
 - 2.8 Ghz Westmeres
 - 48GB RAM/node
 - 13,500 HEP SPEC Score
 - 72 Nodes, 12 Cores/Node
- LLNL::SE
 - 670TB usable (44% used)
 - 12 servers and 1 redirector providing xrootd storage



Monitoring

- AliEn Only -> <http://alimonitor.cern.ch>
- OSG only RSV -> <https://myosg.grid.iu.edu>
- Locally monitored by LC staff/operations

NERSC - PDSF at LBNL

- Also on Open Science Grid
 - VOBox submits to OSG CE, which submits to SGE scheduler
- Multi-group of ~200 nodes with internal disks, ~1800 cores
- Scheduler limit of 4.5GB of memory per job.
 - Not all nodes have this much RAM/core, so many jobs requesting max memory will result in fewer jobs.
- ALICE purchases HW as a share of the facility
 - Currently ~700 cores, but often opportunistically use more
- LBL::SE ~ 700TB --> 18% utilization
- OSG services at both sites:
 - RSV service monitors reliability/availability, reports them to OSG, which then forwards to WLCG
 - Gratia service monitors accounting, also OSG -> WLCG



Grid Tools Requests

- All files in packages.
 - RPM/deb are nice, but any will do.
 - Standard file locations: /usr/bin, /usr/sbin, /var/log, etc.
- Use of standard System V init scripts for all services.
- Prefer use of system supplied standard tools over custom built
 - perl, httpd, etc.
- Ability to start services or check proxy lifetime without providing authentication.
- Better WAN diagnostics, esp. for non-european sites.
- MonaLisa Requests:
 - Number of running jobs should better reflect local scheduler numbers.
 - SE monitoring at individual xrootd server level (not just SE).

