

ALICE Grid Activities in US





ALICE-USA Computing Project



- ALICE-USA Collaboration formed to focus on the ALICE EMCal project
 - Construction, installation, testing and integration
- 11 participating institutions
 - 3 National Labs: LBNL, LLNL, ORNL
 - 7 Universities: Creighton, Houston, Purdue, Tennessee, Texas, Wayne State, Yale
- ALICE-USA Computing Project was proposed to build a facility to:
 - Enable US ALICE Scientist to pursue research goals
 - Meet ALICE-USA computing obligation to provide its share of computing resources for ALICE data analysis and simulations
- Proposal settled on 2 Department of Energy (DOE) labs as primary sites
 - Lawrence Livermore National Lab (LLNL)
 - Lawrence Berkeley National Lab (LBNL) – National Energy Research Scientific Computing (NERSC)



Sites: LLNL- Livermore Computing



- **LLNL Computing**
 - Institutional based High Performance Computing
 - Support local Science & Engineering activities (some classified)
- **Developed Very Cost-Effective Procurement & Operations model**
 - Very large purchases of scalable units
 - In-house managed & optimized Redhat OS and other software
- **Security requirements make external collaborations a challenge**
 - LLNL Computing actively pursued adding more open computing facilities:
Green Linux Computing Cluster
 - Green Data Oasis serves Climate modeling data to research community
 - Grid Activities are often a problem, but the ALICE workflow, which allows for no direct user access, fit the Security plan

Sites: LBNL-NERSC



- DOE Office of Science Flagship High Performance Scientific Computing Facility for Scientific Research
 - Available to DOE Office of Science sponsored research
 - Allocations via competitive awards
- Computing for Scientific Research
 - Two Cray Systems: Hopper & Franklin: ~200k cores
 - Special Clusters: HENP, Visualization, Analytics
 - NERSC Global File System (GPFS): >1PB
 - Archival Storage (HPSS) : ~40PB
 - Data Transfer, Science Gateways & Grid Services
 - Evaluation Systems: GPU & Cloud Clusters
- Large number of user support services
 - Programming & benchmarking
 - Workflow & Gateway hosting
 - Visualization
 - Tutorials



- | | | |
|-------------|-----------------|-----------------|
| ■ Physics | ■ Math + CS | ■ Astrophysics |
| ■ Chemistry | ■ Climate | ■ Combustion |
| ■ Fusion | ■ Lattice Gauge | ■ Life Sciences |
| ■ Materials | ■ Other | |



Sites LBNL-NERSC/PDSF



- PDSF: High Energy Nuclear Physics Computing Facility
 - Operated by NERSC since mid-90s for HENP community
 - Supports both production and user analysis
- Nuclear Science Groups:
 - Tier 1 facility for the STAR experiment
 - Neutrino research: SNO, IceCube, KamLAND
- Physics/Astrophysics Groups:
 - ATLAS
 - Nearby Supernova Factory, Planck
 - DayaBay
- First Grid-enabled facility within NERSC:
 - PPDG
 - Open Science Grid
- HENP Data store → NERSC/HPSS Archival Storage
 - STAR near top user group of NERSC/HPSS: ~1PB stored
 - ALICE allocation is comparable to the STAR allocation

PDSF at a Glance

Interactive Nodes

| | |
|---|---------------------------------------|
| 4 | pdsf.nersc.gov pdsf[1-4].nersc.gov |
|---|---------------------------------------|

Compute Nodes

| | |
|-----|------------|
| 205 | 1100 Cores |
|-----|------------|

GPFS Filesystems

| | |
|-------|-------------|
| 641TB | Eliza[1-18] |
|-------|-------------|

Local Disk

| |
|-------|
| 450TB |
|-------|

Batch System

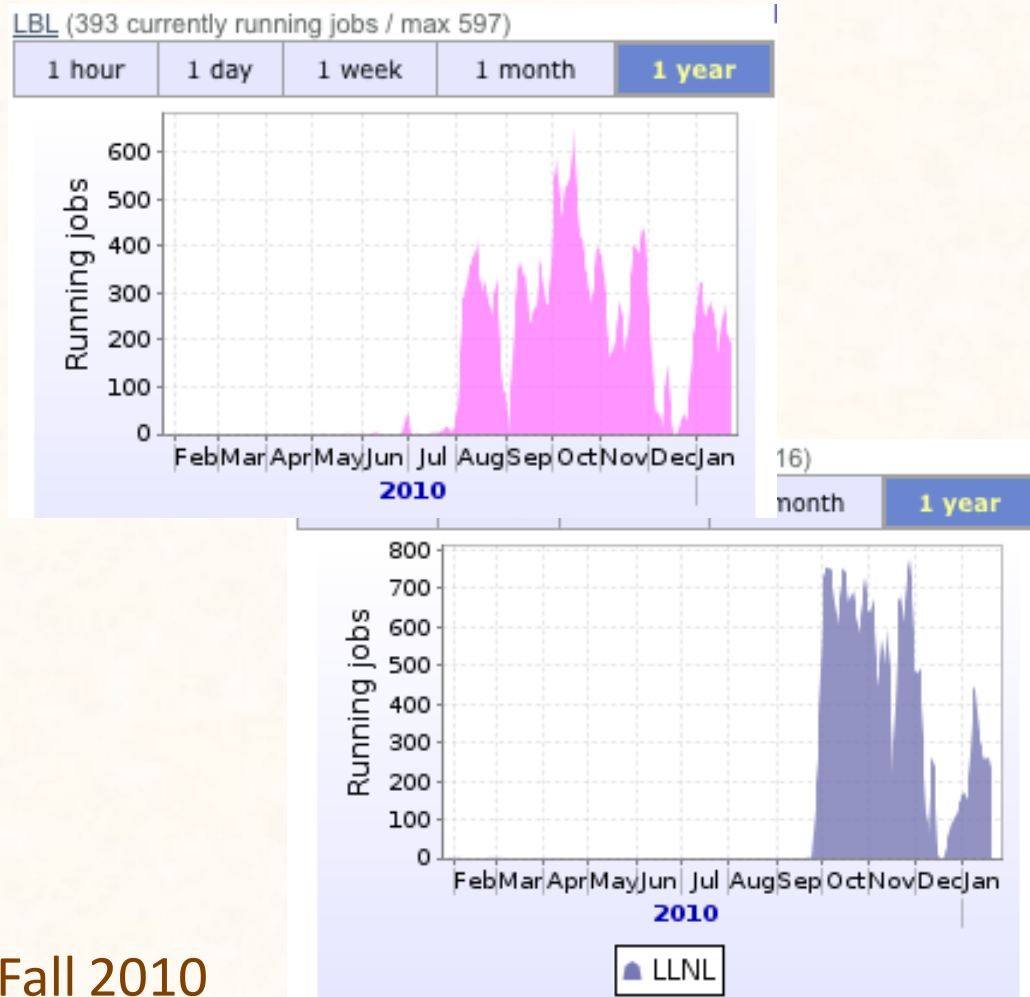
| | |
|-----|------------|
| SGE | <u>Sun</u> |
|-----|------------|



ALICE-USA Computing Project Timeline



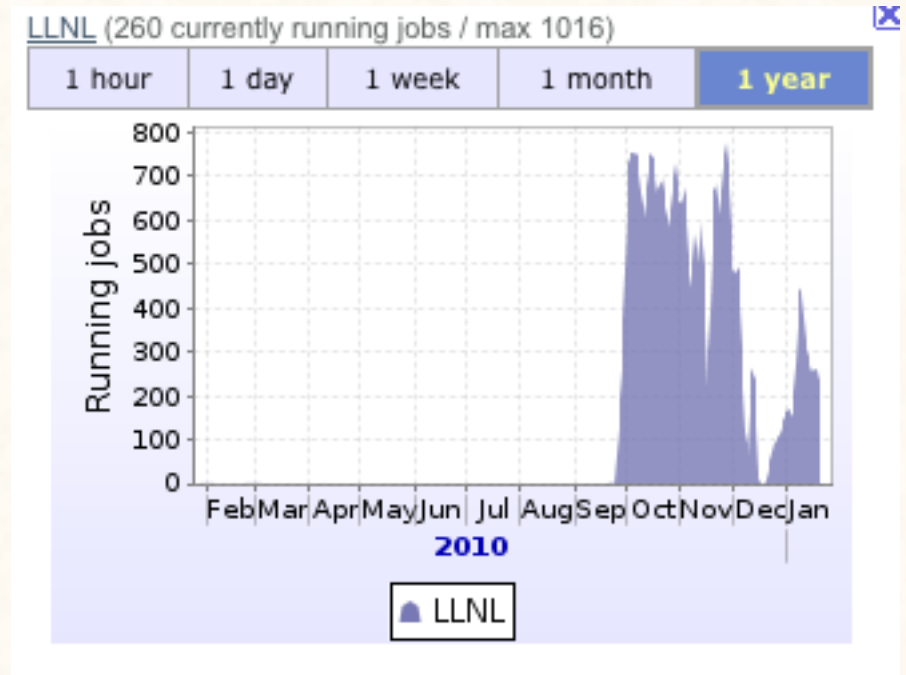
- Proposal reviewed in 2008
- Refined during 2009
 - Specific & yearly HW targets
 - Operational milestones
 - Grid-Enabled CPU & Storage
 - Stress-test services
 - Update security plans
 - Develop resource reporting tools
 - Integrate NERSC HPSS allocations
 - Independent site reviews
- Plan Steps:
 - Establish operation
 - Register Both as Tier 2 Facilities
 - Transition NERSC Site to Tier 1
- Plan accepted by DOE 02/10
- HW procurements begin 04/10
- Sites begin operation Summer/Fall 2010



Facilities: LLNL/LC



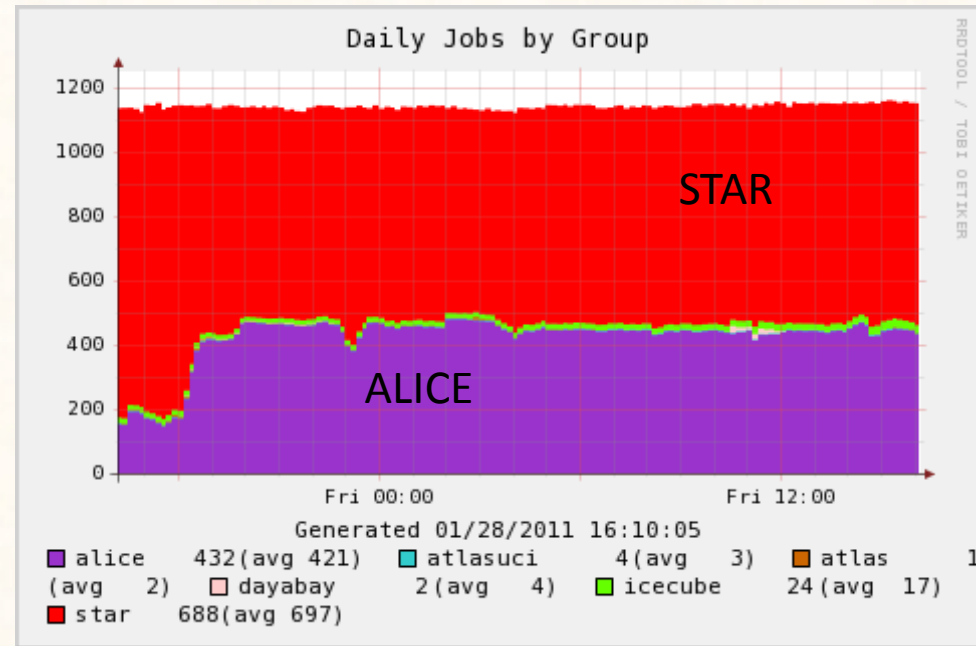
- Single-use facility
 - Grid-only
- ALICE Grid Facility
 - Site → ALICE::LLNL
 - >~700 job-slots (cores)
 - 13.5 kHEP06
 - 650TB Disk storage
 - Resources fixed for 3 year life-cycle
- No user accounts for non-LLNL persons



Facilities: LBNL NERSC/PDSF



- Multi-group facility:
 - ALICE, STAR, ATLAS, ...
- ALICE Grid Facility:
 - Site → ALICE::LBL
 - 300-400 jobs slots (cores)
 - ~4.0 kHEP06
 - 400 TB disk storage
 - PB-scale tape: NERSC/HPSS
 - Resources grow each year by:
 - ~500 cores
 - ~500 TB disk storage
- ALICE User facility
 - Local user accounts are allowed
 - Used as a grid job-submission site
 - Allows direct batch submission use of unused grid cycles



Open Science Grid (OSG)



- US project & consortium: national, distributed computing grid for data-intensive research

- ~30 Participating Virtual Organization (VO) comr
- ~80 independent self-managed computing centre

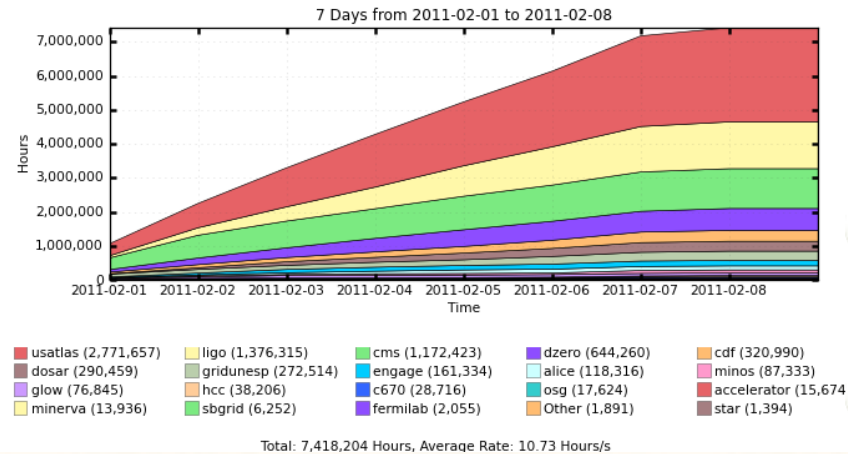
- OSG provides pre-built software bundles

- Compute Element packages
 - GRAM-2, GRAM-4 → GRAM5
 - Working to provide CREAM option
- Storage Element packages
 - dCache, BeStMan, Hadoop, xrootd
- Client Software
 - Submit host (Condor-G ...)
 - WN-client

- Authorization/Authentication Tools
 - Grid Certificate request & retrieval tools
 - VO-discovery maps, VOMS, GUMS

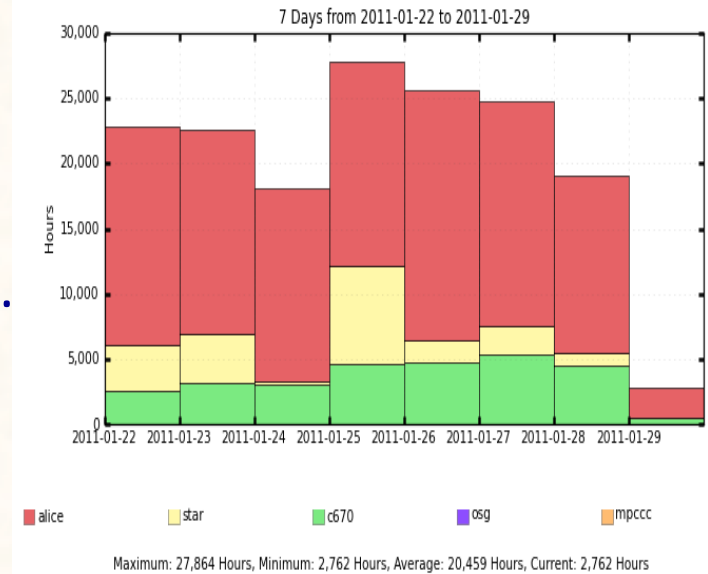
- OSG Grid Operations Center (GOC)

- Site registration and availability
- Resource discovery repositories
- Accounting/Job statistic reports
- trouble ticket management



- **Open Science Grid & WLCG**
 - US ATLAS & CMS major stakeholders
 - Works with WLCG for VOs
 - Provides usage & capacity reports
 - DOEGrids CA partner: OSG Registration Auth.
- **ALICE joined OSG as a VO**
 - NERSC/PDSF was already an OSG site
 - LLNL Job Accounting is reported to OSG
 - Full LLNL integration is in the ALICE-USA Computing Project Plan
- Accounting records for ALICE jobs at both sites are being reported to the WLCG via OSG standard reports

OSG Accounts from NERSC & LLNL



OSG – Software Deployment Model



- Software is packaged by VDT
- OSG Validation Test Bed (VTB)
 - 2-3 small sites (1-3 nodes each)
 - Test software deployment & config
- OSG Integration Test Bed (ITB)
 - Small clusters (4-8 nodes) operated at larger sites
 - Admins deploy software updates & test functionality
 - VOs test workflows on new/updated software
 - Allows New VOs to test building distributed workflows
- OSG is working to provide a CREAM-CE option
 - LBNL is a participating OSG-ITB site
 - ALICE will run site at LBNL-ITB as part of that integration work

OSG Integration Test bed



OSG Interoperability & Outreach

- OSG active interoperability efforts

- EGEE/EGI/WLCG
- US National Grids (TeraGrid)
- Campus & Regional Grids
- non-US National Grid Initiatives

- National Grids in the Americas

- OSG members collaborate directly
 - osg-americas@opensciencegrid.org
- Previous workshops & Grid Schools: Costa Rica, Columbia, & Brasil
- CLCAR 2010 Panel discussion

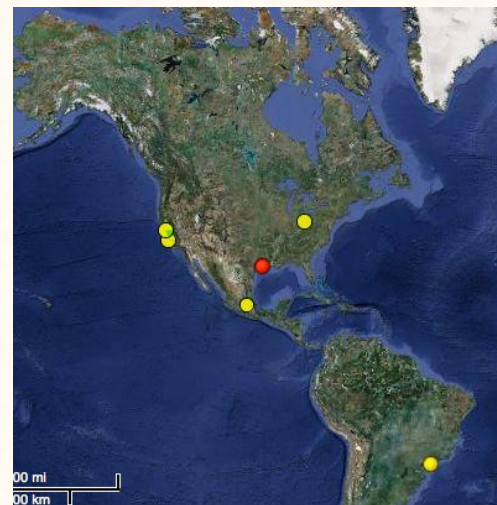
- Support National Grid Initiatives with goal of providing International Grid Infrastructure



Other US ALICE Grid Sites & Activities



- Ohio Supercomputing Center, Glenn Cluster
 - ALICE-Ohio State U. Group
 - ~150 cores
 - 30TB Disk Storage
- University of Houston
 - Planned return to operation



- ALICE-US Data Transfer Group
 - ESNnet, NERSC, LLNL, OSC, OARNet, CERN, ALICE
 - Previous efforts to optimize targeted networks
 - ESNnet-KISTI-BNL
 - NERSC-ANL-ORNL, NERSC data transfer group
 - <http://fasterdata.es.net>

