

UNCLASSIFIED

LANL Computing Environment for PSAAP Partners

Robert Cunningham

rtc@lanl.gov

HPC Systems Group (HPC-3)

October 2009

LANL Resources Available To Alliance Users

- Coyote is old, has a tiny allocation
 - Accounts available for porting, data transfer, small scale runs
 - Bproc Linux cluster, difficult port, LSF scheduler
 - 2,600 Opterons with IB, 13.5 TF, max of 510 procs/job
- Lobo is current workhorse -- TLCC platform
 - Simple Linux OS, Moab job scheduler, Panasas filesystem
 - 4,352 compute CPUs; max 2,144 procs/job
- Small Roadrunner platform available: Cerrillos
 - 1,440 hybrid pairs (AMD/cell), max 360 procs (x2) / job

LANL Lobo Cluster

- Standard DoE Lab cookie-cutter cluster from Appro, Intl.
- 75% allocation for ASC, primarily for PSAAP
- 2 Connected Units (CUs) combined include:
 - 4,352 x AMD Opteron @ 2.2 GHz cores on compute nodes
 - Voltaire InfiniBand interconnect
 - 8.7 TeraBytes RAM
 - Theoretical peak of ~17.4 teraflops



LANL Cerrillos Cluster

- Hybrid architecture: Opteron+Cell, from IBM.
- 25% allocation for ASC, primarily for PSAAP
- 2 Connected Units (CUs) combined include:
 - 1,440 x AMD Opteron @ 1.8 GHz cores on compute nodes
 - 1,440 Cell Broadband Engines
 - Voltaire InfiniBand interconnect
 - 11.8 TeraBytes RAM
 - Theoretical peak of ~152 teraflops



LANL HPC Environment

- Obtain an account, acquire cryptocard (foreign nationals start early!)
- Access HPC platforms in Open Collaborative (Turquoise) Network via firewall: `ssh wtrw.lanl.gov`
- After connecting, `ssh` into front-end: `lo1-fe`, `lo2-fe`; `ce-fe1`, `ce-fe2`
- Moab + [Slurm | Torque] schedules nodes, batch or interactive
 - `msub scriptname -or- msub -I` (hopefully)
 - `checkjob`, `showstate`, `showq`, `mjobctl`
 - `mpirun` to run parallel jobs across nodes
 - Fairshare scheduling to deliver pre-determined allocation
- Unique security model
 - no connecting (`ssh`) between platforms
 - No kerberos tickets; keep cryptocard handy
 - `scp` or `sftp` using File Transfer Agents (FTAs), `turq-fta1.lanl.gov` and `turq-fta2.lanl.gov`

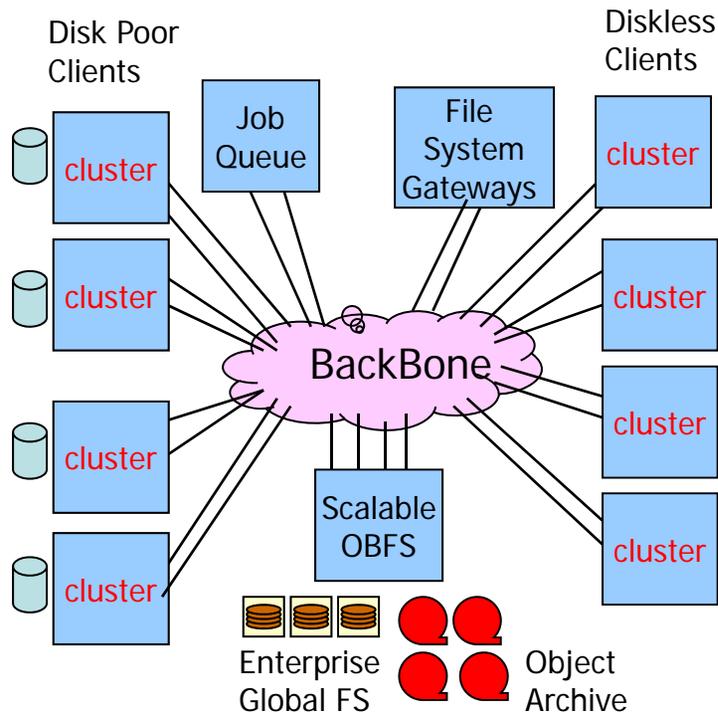
LANL HPC Environment

- Usage model
 - Front-end: text editing, job script set-up, pathname arrangements, compilations, linking, and job launching
 - Compute nodes: run applications
 - I/O nodes: out to local disk, no direct user access
 - Possible intra-network File Transfer Agent (FTA) arrangement in future
- Modulefiles to establish compilers, libraries, tools in \$PATH, \$LD_LIBRARY_PATH
- Compilers: Intel, PGI, Pathscale. MPI: OpenMPI, Mvapich
- Tools: TotalView, IDL, python, perl, Subversion, Atlas

LANL Turquoise HPC Storage

- Tiny home directories, not shared between clusters (security)
- Larger NFS-based “workspace” in `/usr/projects/proj_name`
- Big parallel, globally-accessible filesystem: `/scratchN`
 - Cross-mounted to all HPC nodes in Turquoise
 - ~800 TB total space
 - Fast for parallel I/O, slower than NFS for serial transfers
 - No automated back-ups
 - Purged weekly, 30-days or older!
- Archival storage available via Tivoli (TSM), NFS-mounted as `/archive/proj_name/moniker`
- File Transfers in/out of Turquoise are *S L O W*, we will be addressing this in the next fiscal year, may need your help.

Parallel Scalable Back Bone (PaScaIBB)



- Relieve the compute nodes
- Multiple clusters sharing large, global namespace parallel I/O subcluster
 - Includes Cerrillos/Lobo/Coyote
- Network is combination of HPC Interconnect + commodity networking bridge
- Panasas is storage vendor
- I/O through a set of fileserver nodes over IB; nodes serve as interconnect<->GigE routers.

3 LANL Web Sites You Can't Live Without

- User Docs

`http://bear.lanl.gov/drupal/?q=tlcc_home`

(or) `https://bear.lanl.gov/drupal`

(or call 505-665-4444 option 3, consult@lanl.gov)

- Calendar `http://ccp.lanl.gov`

- HPC Training

`http://int.lanl.gov/projects/asci/training`

UNCLASSIFIED

HPC Accounts



Don't forget Photo Op!

UNCLASSIFIED

Questions?