

Transition from NQE to PBS

Jim Glidewell

Boeing Shared Services Group

james.glidewell@boeing.com

Our Computing Environment

- Primary engineering datacenter for the Boeing Company
- Long time Cray shop (>20 years)
- Mix of vendors: Cray, IBM, Sun, SGI, STK
- Tend to make heavy use of enhanced OS features
 - ◆ DMF on both Cray and Origin
 - ◆ Job and project accounting
 - ◆ NQE/NQS

The Origin 3800

- This was our first SGI “HPC” system
- Origin 3800
 - ◆ Initially a 64 CPU, 64GB
 - ◆ Now 256 CPUs, 384GB
 - ◆ Mix of JBOD and RAID disk
- Job workmix for the Origin 3800
 - ◆ A few large multi-CPU aero jobs (32 - 64+ CPUs)
 - ◆ A small number of 4 - 16 CPU moderately parallel jobs
 - ◆ A fairly large number of single-CPU jobs

Origin 3800



Why PBS?

- NQE has been orphaned by SGI
- Shared roots and design philosophy with NQS
- Our Aero customers familiar with it from NAS
- Availability of source code
- Active user community
- Fairly flexible and customizable
- Supported on many platforms, including Cray as well as SGI

Why Pro?

- Professional support staff
- Active development
- Cpuset support
- Reasonable investment to ensure best possible utilization of the computing resource

Transition elements

- Initial evaluation and testing
- Overall transition plan
- User documentation
- Transition tools
- User beta test

Initial Evaluation

- Found several minor bugs
 - ◆ Most were fixed by the time we were ready to implement PBSPro
- Minor deficiencies were also remedied
- Single-CPU jobs in a cpuset environment
 - ◆ Jobs reserved a minimum of 1 full node (4 CPUs)
 - ◆ Requested enhancement to resolve the issue
- Overall, PBSPro appeared to be a solid product, well supported

Overall Transition Plan

- Required tools & conversion document
- Detailed planning
- Documenting or mitigating user impacts
- Project timeline

User Documentation

- Overview of PBS
- Common PBS commands
- PBS job structure and common qsub parameters
- Notes on Boeing's Origin 3800 PBS environment
 - ◆ Resource limits
 - ◆ Priorities
 - ◆ Large files & \$TMPDIR
- Conversion of existing NQS scripts
- Overview of the transition plan

Transition Tools

- nqs2pbs script conversion tool
 - ◆ Provides simple method of conversion
 - ◆ Scripts generated are usable under PBS and NQS
 - ◆ Useful for both users and admins
- PBS and NQS use common commands (qsub, qstat)
- “nqs” and “pbs” cover scripts
 - ◆ Allowed users to specify which batch subsystem they wanted
 - ◆ “pbs qsub my_job_script” forces use of PBS
 - ◆ “qsub my_job_script” uses the current default batch subsystem
 - ◆ Modifies both “PATH” and “MANPATH”

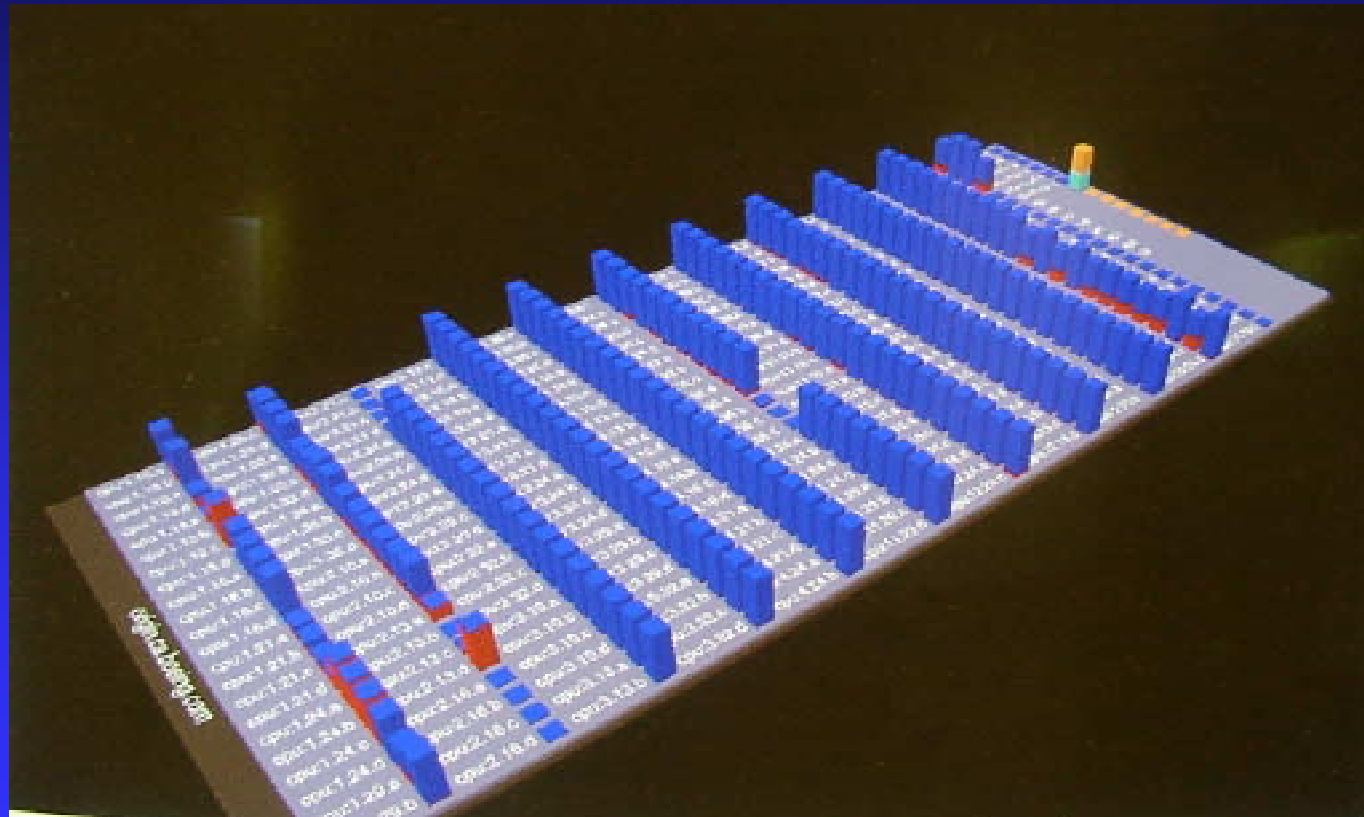
User Beta Test and Transition

- Used a phased approach, allowing users to convert at their own pace
- Initially installed PBS as secondary batch subsystem
 - ◆ Users could use the “pbs” cover script
 - ◆ Monitored usage of each batch subsystem, adjusting limits to avoid oversubscription
 - ◆ Users with large parallel jobs moved over immediately
- Switched default to PBS
 - ◆ Users who had not converted could still use the “nqs” cover script
- Stop NQS queues
- Remove NQS

Current Issues

- IRIX memory reporting
- Move from 5.1.3 to 5.2 was fairly rocky
- “Shared cpusets” feature was not quite ready
- Veridian has provided excellent support during this rough spot

Current Performance



Conclusions

- Overall the transition went fairly smoothly
- The decision to defer implementation of cpusets was a good one
- A decent transition guide is essential, both for users and support staff
- The ability to run both batch systems side-by-side allowed for a gradual user transition
- PBSPro appears to have been a good choice for our Origin 3800