PBS Works

Bill Nitzberg, Ph. D.
CTO, PBS Works
Altair Engineering, Inc.
December 2009
One-Slide Summary™

Portals to increase productivity

Scheduling to maximize value

Analytics to optimize ROI

Open architecture to gain agility
About Altair Engineering

A global software and technology company focused on:

**Product Development**

![HyperWorks Logo]

**Advanced Computing**

![PBS Works Logo]

**Enterprise Analytics**

![HiQube Logo]
Our Background

1980-90’s – NASA
- PBS for parallel computers
- MPI-2 I/O editor
- NASA Metacenter: one big computer
- Whitney Cluster project
- Grid Forum, founder & director

2000-10’s – Altair
- Commercialized PBS Professional
- NASA’s Information Power Grid
- PBS runs on all 7 continents!
- Open Grid Forum board member
- PBS Works suite

17 Countries – 1,300 Employees – All Industries
High Performance Computing

Before PBS

- Unused Resources
- Running Job
- Unused Resources

Job 1

Job 2

Job 3

Unused Resources
PBS Works

5 Strategic Pillars

• Easy to use
• Hard to break
• Do more (with less)
• Keep track and plan
• Open architecture
Easy to Use

Engineers should focus on engineering, not computer science

Designers should focus on design, not computer science

Scientists should focus on science, ...
Portals Increase Productivity

I push enter and my jobs run; the results come back when they’re done.

– Major Automotive OEM
Hard to Break

Bullet-proof reliability

Runs everywhere – all platforms, all workloads

Scales to 1000’s of users, 10,000’s of cpus, 1,000,000’s of jobs
Reduce Risk & Ensure Business Continuity

One of the ringing endorsements I can give PBS Professional ... it just works.

- TGen

We can load the system with a huge number of jobs and get a good response.

- TRW Tech Center
Do More (with Less)

Run the right job at the right time

Run applications as fast as possible

Zero waste – hardware, software licenses, energy, people’s time
Maximize Value

**Goal #1**
Enterprise Policy

*It allows us to keep our policies working at maximum levels.*

– CILEA

**Goal #2**
Optimal Use

*It's done a really good job of keeping the clusters utilized.*

– U. Florida
Keep Track and Plan

Show a detailed accounting of what happened

Monitor and control costs for computational projects

Plan infrastructure growth based on actual and predicted usage
Optimize ROI

Maximizing our license utilization means we don't have to buy a new license, set up another workstation, and hire another engineer to keep up with demand.

– Trelleborg
Open Architecture

Connect with any 3rd party middleware or application

Exploit unique hardware and OS-level services

Extend and customize for unique site policies and behavior
Gain Agility

Without its flexibility, it wouldn't have been able to handle what we're doing.

– NASA Ames
Products

Easy to use

Hard to break
Do more (with less)

Keep track and plan
Community Area – New!
Commercial-grade HPC workload and resource management
## A Few Examples of Large PBS Installations

<table>
<thead>
<tr>
<th>Customer</th>
<th>Hosts</th>
<th>Cores</th>
<th>Jobs per day</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMCAC</td>
<td>1,792</td>
<td>14,336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scripps</td>
<td>1,000s</td>
<td>3,725</td>
<td>50,000</td>
<td></td>
</tr>
<tr>
<td>Aerospace Company</td>
<td>1,000s</td>
<td>4,500</td>
<td>1,000</td>
<td></td>
</tr>
<tr>
<td>NASA Ames</td>
<td>10,000</td>
<td>70,000</td>
<td>1,000</td>
<td>1,000</td>
</tr>
<tr>
<td>LRZ</td>
<td>20</td>
<td>10,000</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>NIH</td>
<td>3,000</td>
<td>4,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CINES</td>
<td>2,500</td>
<td>20,000</td>
<td>500</td>
<td>~300</td>
</tr>
<tr>
<td>AIST</td>
<td>2,500</td>
<td>5,000</td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>
## PBS Professional – New

<table>
<thead>
<tr>
<th>Select &amp; Place</th>
<th>Hooks</th>
<th>Eligible Time</th>
<th>Fail-over</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fairshare</td>
<td>Policy-based scheduling</td>
<td>$restrict_user</td>
<td>Checkpoint / Restart</td>
</tr>
<tr>
<td>Scheduling Formula</td>
<td>Standing Reservations</td>
<td>Dynamic resources</td>
<td>Job Arrays</td>
</tr>
<tr>
<td>Backfill</td>
<td>Kerberos</td>
<td>Preemption</td>
<td>Job History (qstat –x)</td>
</tr>
<tr>
<td>High-availability Reservations</td>
<td>Age-based Scheduling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Dependencies</td>
<td>License Scheduling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24x7 On-line Community</td>
<td>Desktop harvesting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-demand Licensing</td>
<td><strong>GPU scheduling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Connect Heterogeneous Clusters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Green Provisioning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Over-subscription</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Meta-scheduling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Interactive Jobs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>User/Group VIP Limits</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Topology Placement Sets</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Multi-core</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Peer Scheduling</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Web Services</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Select & Place**
- **Hooks**
- **Eligible Time**
- **Fail-over**
Tunable Scheduling Formula – New!

Define any policy – including on-the-fly “exceptions”

Simple formulas are very simple (big jobs go first)

\[ \text{ncpus} \times \left( \frac{\text{walltime}}{3600.0} \right) \]

Complex formulas are pretty simple too… (adds priority accrual for smaller jobs, high-priority queue, deferred queue, “run this job next”)

\[ (\text{ncpus} \times \left( \frac{\text{walltime}}{3600.0} \right)) \times \text{Wsize} + \]
\[ (\frac{\text{eligible\_time}}{3600.0}) \times \text{Wwait} + \]
\[ \text{special\_p} \]
Green Provisioning™ – New!

Only Machines Running Jobs Use Power

- Automatic shutdown service
- Automatic wakeup service

Green Provisioning: ~20% operational power savings
Overall PBS Works ROI: ~12 months to break even

[Animation Customer, APAC]
Submission Filtering “Hooks” – New!

Change/augment capabilities in the field, on-the-fly, without source

```python
if e.job.Resource_List["walltime"] is None:
    e.reject("You must specify a walltime")
```

- Admission control – validate requests
- Allocation management
- On-the-fly tuning
- Custom logging, reporting, debugging, and even patches!
Standing Reservations – **New!**

**Guarantee resources for recurring needs**

```bash
pbs_rsub -R 0500 -E 0800 \
–r "FREQ=DAILY;BYDAY=MO,TU,WE,TH,FR;UNTIL=20091231" \n-1 select=200:ncpus=2 –l place=scatter:excl
```

- Run the weather simulations from 5-8am every weekday morning
- Reserve the computing lab for classes on MWF 14:00-16:00
- Block out time for maintenance the first weekend of every month
PBS Application Service – **New!**

**Software-as-a-Service (SaaS) Layer**

- **Web Services Interface**
  - Standards-based (leverages OGSA)

- **Application Definitions**
  - Definitions for major apps provided
  - Portable – platform and site neutral
  - Easy to add new applications
  - Trade via PBS On-line Community

- **Site-specific Configuration**
  - E.g., executable paths, is kept separate to simplify configuration

- **Clients and portals automatically recognize new applications**

- **Also**
  - Interim staging area for input and output files
  - Access working directory for running jobs
  - HPC Basic Profile Service
Job Status with History (via “qstat –x”) – New!

Never lose track of your jobs

Familiar qstat and qselect extended to include all job states
- Queued jobs and Running jobs
- Moved jobs (qmoved and Peer’d)
- Finished jobs (deleted, completed successfully, and exited with errors)

Submission arguments enable easy resubmission

Peer Scheduled jobs show destination

Additional information available
- Exit status, stage-out success, location job ran, resource usage

100% backward compatible
- Your existing scripts using qstat & qselect will still work

Job history duration is configurable
Supported Platforms

All major platforms and operating systems supported
Integration with HP CMU

Monitor jobs (with user groups) and PBS Green Provisioning™ state
PBS Analytics™

Visualize historical usage for optimized returns on HPC investments
PBS Analytics – New!

A web-based portal for visualizing HPC usage data
PBS Analytics Benefits

- Generate reports automatically “out-of-the-box”, then customize...
- Understand usage trends for capacity planning
- Verify project planning assumptions to meet deadlines
- Extract accounting data for chargeback
- Track software license use to optimize ROI
PBS Analytics Features

- Modern Web 2.0 technology
- Numerous graph types
- Secure, role-based access: users view their data; managers view group data
- Customize, save, and share reports or entire dashboards
- Drill-down to underlying data
- Publish via web and export to Excel
- Plugs-in to enterprise infrastructure
System is lightly loaded..

“Most jobs run right away. There was very little wait time.”

“Only in Oct08 users had to wait – well that’s due to the new project in Oct08!”
“Users are asking way more memory than needed.
This could have queued some jobs potentially..”
Power users

Company A: “Ensure they get good support as they are our top users.”
Company B: “Ensure they don’t consume all the resources, let’s set a limit…”
License Utilization & Denials – New!
PBS Catalyst™

Intuitive, application-aware job submission and management portal
PBS Catalyst – New!

Drag-and-drop input to submit
• Recognizes application and fills in default job parameters
• Output is automatically returned

Increase user productivity
• Monitor, manage, prioritize jobs
• Create profiles for common runs
• Connect to multiple PBS servers
PBS Catalyst
Trends

- GPUs
- Green
- Clouds
MIMOS Builds “Render Cloud” with PBS & HP

Les’s Copaque renders Malaysia’s first 3D animation feature with Altair’s PBS Works on HP equipment

Top movie for the first 2 weeks (Grossing RM 6 million in 1st month)

www.gengthemovie.com
Thank You!

- Increase productivity
- Reduce risk
- Maximize value
- Optimize ROI
- Gain agility

www.pbsworks.com