

### KEY FACTS

#### Overview:

ORNL's Cray XT3 23,416 core supercomputer powers the DOE's INCITE program, comprised of a wide variety of users.



#### Challenge:

Due to the diversity of users in the INCITE program, ORNL found there were limitations in the workload management of their system.

#### Solution:

Moab Workload Manager® is selected to resolve the limitation issues caused by the diverse user base.

#### Results:

Increased system utilization; decreased downtime; more control over resources; and an enhanced end-user experience that does not require changes to submission habits.

#### Contact:

[info@clusterresources.com](mailto:info@clusterresources.com)

US Headquarters

+1 (801) 717-3700

+1 (888) 221-2008

EMEA Headquarters

+44 (1223) 437134

[www.clusterresources.com](http://www.clusterresources.com)



## ORNL Selects Moab for Cray XT3 Jaguar

### ABOUT OAK RIDGE NATIONAL LABORATORY

Oak Ridge National Laboratory is the Department of Energy's largest science and energy laboratory. Managed since April 2000 by a partnership of the University of Tennessee and Battelle, ORNL was established in 1943 as a part of the secret Manhattan Project.

Jaguar, Oak Ridge National Laboratory's 23,416 core Cray XT3 super-computer, ranked as the 10<sup>th</sup> fastest system in the world in the November 2006 Top500 list. Jaguar is the largest system in the Department of Energy's Office of Science and is the major computing resource for DOE's Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. The system is available to all scientific researchers and research organizations, including industry, through an annual call for proposals. Three of the four companies awarded INCITE grants in 2006 do work at ORNL—Boeing, DreamWorks Animation and General Atomics.

### CHALLENGE

Due to the diversity of users in the INCITE program, ORNL found there were severe limitations in the way their system handled workload management; these issues needed to be overcome. Limitations included inflexible policies that limited system performance, simplistic prioritization that could not weigh all the factors for job priority, and limited reporting and information.

### THE SOLUTION

ORNL chose Cluster Resources' Moab Workload Manager® to resolve the limitation issues. Through its ability to abstract resource management from policy management layers, Moab is able to intelligently optimize and manage almost any type of resource and workload. Moab additionally provides advanced optimization and quality of service features together with highly flexible resource ownership and resource management policies.

Moab can import basic job information from the existing PBS-based resource manager using a native resource manager interface. The native resource manager interface was preferred because it provided more flexibility in job monitoring and job control, avoiding any potential reporting errors and/or limitations.

Once data import was achieved, the next step was analysis of the larger environment. Moab's advanced diagnostics features made this process exceptionally fast. With Moab, issues are most commonly found by running a number of diagnostic commands using the 'mdiag' routine rather than by reviewing logs. This command will execute a detailed analysis of the current and historical states, and will report any unexpected conditions such as unexpected values, misconfigurations, internal table overflows or corruption, externally detected failures, network issues, peer service failures, policy conflicts, etc.

With all diagnostic feedback addressed, Moab was allowed to execute in normal production mode and was evaluated in a matter of days. Performance was good and exceeded even the most optimistic expectations. Shortly thereafter, at ORNL's request,

the evaluation period was reduced by an additional 10 days. Moab went into full production operation after a design, development, proof-of-concept, and testing cycle of a little more than three weeks.

## RESULTS

After an aggressive evaluation, ORNL deployed Moab in production mode on Jaguar and experienced resolutions to their workload management problems including:

- More flexible resource management capabilities
- More advanced and robust policies
- Tunable prioritization
- Provided administrators a “run-this-job-next” ability
- Evaluation of new policies without disruption
- Resolution of SSH X11 forwarding issues
- More in-depth system and workload information
- Effectively mitigated failures, decreasing downtime
- Decreased administration workload
- Provided centralized policies
- Improved reporting and resource utilization information

*“We are pleased with how well Moab has worked on top of the Cray XT3 cluster to give administrators fine-tuned control. ORNL appears also to be pleased, as they are now running Moab on their leadership-class computers.”*

— David Jackson  
CTO, Cluster Resources

Additionally, Moab succeeded in increasing system utilization, decreasing downtime, and allowing more control over resources. Due to Moab’s integration and support of the underlying resource manager, users have not experienced any changes in the interface to the batch system on NLCF machines. The Moab scheduler is also able to provide a start-time estimate for a job by utilizing the ‘showstart’ command. Moab has built-in diagnostics to provide answers to many questions about job status quickly and easily, including answers to why jobs are not running and any unexpected scheduler behavior.

## BUILDING ON SUCCESS

After the successful rollout of Moab on the Cray XT3 helped ORNL reduce administrative staff time and increase overall system utilization, ORNL chose to install Moab on additional systems such as "Phoenix", a 1,024 multi-streaming vector processor Cray X1E system, "Ram", a large SGI Altix system, and a SLURM-based Linux cluster. Additionally, Moab may serve as the underlying architecture for a management and reporting grid or a workload grid across ORNL’s Cray and non-Cray hardware systems.

In March 2007, ORNL upgraded Jaguar to the Cray XT4; recent expansion of the system has doubled performance of the supercomputer at ORNL to 119 teraflops of peak performance (119 trillion mathematical calculations per second). It is now the most powerful open scientific computing system in the world.



© 2007 Cluster Resources, Inc. All Rights Reserved. Moab Workload Manager® is a registered trademark of Cluster Resources, Inc. All third party trademarks are the property of their respective owners.



Cluster Resources, Inc.

www.clusterresources.com • info@clusterresources.com  
+1 (801) 717-3700 • +44 (1223) 437134