



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

# **Purple Milestone Report System Software and Scalability**

Donald A. Lipari and Morris A. Jette Jr.

**December 5, 2006**

A report on items 40 through 44 of the Purple Milestone Report

## **Disclaimer**

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

## **Auspices Statement**

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

# Purple Milestone Report

## System Software and Scalability

By Don Lipari and Morris Jette  
December 5, 2006

### **Item 40 - LCRM/SLURM collect and report computing resource usage statistics for every job run on Purple**

The following resource usage statistics are collected for every job run on Purple:

- User
- Bank
- Execution Host
- Node Partition
- CPU\*seconds Used
- Memory Integral Used

Evidence: Users can invoke the `lrmusage` tool on the SCF to retrieve this data.

### **Item 41 - Users invoke the same commands and follow the same conventions for running LCRM/SLURM on Purple as they do for all other machines in the Center**

Normal LCRM and SLURM commands are used to submit, monitor, and cancel jobs. Task launch is performed using IBM's `poe` command rather than SLURM's `srun` command. `poe` is needed to interface with IBM's Federation switch and IBM's version of MPI.

Evidence: Users invoke the following LCRM commands to run batch jobs on purple:

- `psub` to submit their jobs to purple
- `palter` to alter their purple job attributes
- `pstat` to see their purple job statistics
- `prm` to cancel or terminate their purple jobs

Users invoke the following SLURM commands to run interactive (in the viz partition) jobs on purple:

- `poe` to submit their job
- `scancel` to cancel or terminate their job
- `squeue` to status their jobs

Users man pages for all the above are installed on purple.

**Item 42 - Job launch times are under a minute or two (excluding factors outside the control of LCRM/SLURM)**

Evidence: Test results show typical time for 12000 task job launch on idle system (by component):

5 sec LCRM  
30 sec SLURM  
50 sec POE  
65 sec MPI  
150 sec Total (sum of above)  
35 sec Total time attributed to LCRM/SLURM

**Item 43 - LCRM/SLURM enforce the computing resource usage policy defined by the CCCs**

Evidence: Bank administrators use LCRM's *lrmmgr* tool to assign shares (proportional usage targets) to each of the CCC banks. The CCC banks are allocated 85% of purple. The ASC Exec banks are allocated 5% to each lab.

LCRM prioritizes jobs to run in an order that achieves the target usage: jobs of users who have used less of their share are granted higher priority.

LCRM administrators use LCRM's *lrmmgr* tool to configure job limits that LCRM imposes:

- Maximum number of jobs that purple will run at any given time
- Maximum number of jobs a user can run on purple at any given time
- Maximum number of nodes a user can request for a job
- Maximum wall clock time duration for a job

**Item 44 - Jobs terminate cleanly (without leaving orphaned processes) in less than one or two minutes (again, excluding factors outside the control of LCRM/SLURM)**

Evidence: Typical time for SLURM to terminate and clean-up a 12000 task job: 12 sec. Typical time for LCRM to process a job termination is under 30 seconds. In addition, LCRM employs a mechanism to detect and terminate job processes once the job completes or is removed.