High Performance Computing at SLAC

Overview

The batch system at SLAC uses the IBM Platform Load Sharing Facility, LSF, and is made up of

- a general farm of batch servers that is open to all SLAC users
- an mpi farm which requires that you request access (send email to unix-admin) and run only mpi jobs
- various other farms of machines that belong to individual scientific computing groups.

Various scientific groups at SLAC contribute to the purchase of the general farm batch systems and jobs submitted to run in the general farm are scheduled according to a fairshare priority structure. If you are not part of a computing group that has a fairshare, your jobs will run according to the priority of the group called AllUsers. More information about fairshare scheduling at Slac can be found here https://confluence.slac.stanford.edu/display/SCSPub/Fairshare+Scheduling

Submitting Jobs to the General Farm

To submit jobs to the general farm you will need to

1. get a SLAC unix account: https://slacprod.servicenowservices.com/it_services?id=sc_cat_item&sys_id=17176b676ff12100aae0c6012e3ee4f7&sysparm_category=d65827c46fd921009c4235af1e3ee434
2. login to a SLAC public machine using ssh

The public and general farm batch machines are described here http://www.slac.stanford.edu/comp/unix/public-machines.html . The simplest command to submit a job to the general farm is

bsub <your job>

This command will submit your job to the first available general farm machine and will be able to run for 20 minutes of wall clock time. It is recommended to run your job specifying a wall clock time since LSF will use that information to make use of windows of opportunity on systems that are accumulating cores for mpi jobs that are not yet running, and your jobs will schedule more quickly. To specify a wall clock time:

bsub -W <time in minutes> <your job>

bsub argument	requirement	description
-q <queue_name>	optional	Specifies a job submission queue. Not required for running jobs in the general shared clusters.
-R "rhel60" or -R "centos7"	optional	The general queues are a mix of rhel6 and centos7 hosts. Use one of these -R options to restrict your job to the indicated OS.
-W <[hours]:minutes>	required	Wallclock runtime limit that is not normalized for CPU differences. Essential for efficient job scheduling. Jobs will be terminated if they exceed this runlimit
-We <[hours]:minutes>	optional	Wallclock estimated runtime limit that is not normalized for CPU differences. The system will consider this estimate for scheduling purposes. Jobs may be not terminated immediately if they exceed this limit
-c <[hours]:minutes>	optional	cputime limit that is normalized by the CPU factor of the assigned core. Intended to prevent runaway jobs but it is not used to schedule time on cores.

General Queues

Queue	Default Runlimit	Maximun Runlimit	Priority
EXPRESS	4 min	4 min	200
SHORT	20 min	60 min	185
MEDIUM	30 min	2 days	180

LONG	30 min	5 days	175
IDLE	12 hours	12 hours	5
bulletmpi *	15 min	7 days	187
bulletmpi-large *	15 min	1 day	187

*The bulletmpi and bulletmpi-large queues are for mpi jobs only. Access can be requested by email to unix-admin.

Fermi/GLAST and batch

Information about using batch at SLAC for the Fermi/GLAST user can be found here: http://www.slac.stanford.edu/exp/glast/wb/prod/pages/installingOfflineSW/usingSlacBatchFarm.htm

Documentation

There are manpages available for the various batch commands and there is full documentation for LSF here https://www.ibm.com/support/knowledgecenter/SSWRJV_10.1.0/lsf_welcome/lsf_welcome.html .