LSF at SLAC

Using the SLAC/LCLS Offline Batch Cluster

Neal Adams
SLAC National Accelerator Laboratory
neal@slac.stanford.edu

What is LSF?

- Load Sharing Facility (LSF) product by Platform Computing Corporation.
- Allows queuing and scheduling of batch jobs.
- Provides scheduling of jobs based on load conditions and resource requirements specified by the user.

What is a batch job?

- "A unit of work run in the LSF system."
- A batch job can be a script, command or program.

  Example: bsub *hostname*

LCLS Offline LSF Servers

- LSF commands for querying and job submission can only be performed from licensed LSF hosts.

- LCLS Offline interactive servers licensed for LSF.
  psexport
  pslogin

- LCLS Offline LSF batch servers.
  psana1101-1120
  Psana1201-1210
  psana1301-1320
  Psana1401-1420

The LCLS Offline Cluster

LCLS Batch Server Farm

- 80 Supermicro Blade servers each with Intel(R) Xeon(R) CPU @ 3.07GHz; 24GB memory
- 960 cores (job slots)

Experimental Hall	Queue	Nodes	Data	Comment
NEH	psnehq	psana11xx, psana12xx	ana01, ana02	Jobs <= 6 cores
	psnehmpiq	psana11xx, psana12xx	ana01, ana02	OpenMPI jobs > 6 cores, preemptable
FEH	psfehq	psana13xx, psana14xx	ana11, ana12	Jobs <= 6 cores
	psfehmpiq	psana13xx, psana14xx	ana11, ana12	OpenMPI jobs > 6 cores, preemptable

Useful LSF Commands

bsub	submit a batch job to LSF
bjobs	display batch job information
bkill	kill batch job
bmod	modify job submission options
bqueues	display batch queue information
busers	displays information about batch users
lshosts	display LSF host information
lsload	display LSF host load information

For more details use: *man <command_name>*.

Useful LSF Commands

- bqueues

```
[neal@pslogin01 ~]$ bqueues
QUEUE_NAME      PRIO STATUS          MAX JL/U JL/P JL/H NJOBS  PEND   RUN  SUSP
[…]
psnehq          115  Open:Active       -    -    -    -     0     0     0     0
psfehq          115  Open:Active       -    -    -    -    10     0    10     0
psnehmpiq       112  Open:Active       -    -    -    -    64     0    64     0
psfehmpiq       112  Open:Active       -    -    -    -    12     0    12     0
[…]
```

- busers

```
[neal@pslogin01 ~]$ busers
USER/GROUP          JL/P    MAX  NJOBS   PEND    RUN  SSUSP  USUSP    RSV
neal                   -      -      0      0      0      0      0      0

[neal@pslogin01 ~]$ busers perazzo
USER/GROUP          JL/P    MAX  NJOBS   PEND    RUN  SSUSP  USUSP    RSV
perazzo                -      -    384      0    384      0      0      0
```

Useful LSF Commands

- lshosts

```
[neal@pslogin01 ~]$ lshosts –R psana
HOST_NAME      type    model  cpuf ncpus maxmem maxswp server RESOURCES
psana1201     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1202     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1203     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1204     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1205     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1206     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1207     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1208     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1209     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
psana1210     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
[…]

[neal@pslogin01 ~]$ lshosts psana1320
HOST_NAME      type    model  cpuf ncpus maxmem maxswp server RESOURCES
psana1320     LINUX INTEL_29  14.6    16 24098M  4095M    Yes (bs linux linux64 rhel50 psana)
```

Using bsub

- To submit batch jobs to the SLAC/LCLS LSF cluster use the *bsub* command.

  bsub [*bsub options*] *command* [*arguments*]

  For example:

  bsub -o *outputfilename* date -u

Using bsub

Example of a simple bsub:

```
[neal@pslogin01 ~]$ bsub –q psnehq hostname
Job <945166> is submitted to queue <psnehq>.

[neal@pslogin01 ~]$ bjobs
JOBID   USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
945166  neal    PEND  psnehq      pslogin01               hostname   Jun 28 15:12

[neal@pslogin01 ~]$ bjobs
JOBID   USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
945166  neal    RUN   psnehq      pslogin01  psana1202   hostname   Jun 28 15:13

[neal@pslogin01 ~]$ bjobs 945166
JOBID   USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
945166  neal    DONE  psnehq      pslogin01  psana1202   hostname   Jun 28 15:13
```

Using bsub

Output from my simple batch job:

```
Job <hostname> was submitted from host <pslogin01> by user <neal>.
Job was executed on host(s) <psana1202>, in queue <psnehq>, as user <neal>.
</reg/neh/home/neal> was used as the home directory.
</reg/neh/home/neal> was used as the working directory.
Started at Mon Jun 28 15:13:27 2010
Results reported at Mon Jun 28 19:13:32 2010
Your job looked like:

------------------------------------------------------------
# LSBATCH: User input
hostname
------------------------------------------------------------

Successfully completed.
Resource usage summary:
CPU time   :      0.06 sec.
Max Memory :         2 MB
Max Swap   :        16 MB
Max Processes  :         1
Max Threads    :         1

The output (if any) follows:
psana1202
```

A few useful bsub options.

- Submit with a CPU limit (normalized): bsub -c

  example: bsub –q psnehq -c 24:00 date

- Submit with a RUN limit (wallclock): bsub -W

  example: bsub –q psnehq -W 24:00 date

- Submit with a jobname: bsub -J "job_name"

  example: bsub –q psnehq -J "Date_job" date

- Submit a job array: bsub –J "job_name[array-elements]"

  example: bsub –q psnehq –J "amedeo[1-100]" my_script

Good Practice

- Specify output files for batch job output. (bsub with -o or -oo options).
  Make sure the file path exists and that you have the appropriate permissions.
- Before submitting 100s of jobs to LSF, please try submitting a smaller number to ensure that you get the expected results.
- Everything required by the batch job (incl. binary) needs to be visible from the batch nodes.
- Use local disk space on the LSF servers for job files and output files for better performance and copy files to your user or group space at job completion.
- LSF can handle tens of thousands of jobs. However we would prefer that not all of them are yours.

Batch Job Exit Codes

- Job exit codes 1-128 are from whatever the user is running while those exceeding 128 are the signal values modulo 128.

  Example:

  A job exit code of 137 would indicate that the job was sent SIGKILL (137-128=9) or kill signal 9.

  A job exit code of 152 would indicate that the job was sent SIGXCPU (152-128=24) or kill signal 24.

- To determine the signal name and number use *man*.

  Linux: `man 7 signal`

Is LSF having problems?

- You may see the following messages in response to your LSF batch commands (bjobs, bsub, etc). These can occur briefly when we have initiated an LSF reconfiguration for administrative purposes.

```
batch system daemon not responding ... still trying
batch system daemon not responding ... still trying
batch system daemon not responding ... still trying
```

  This does not effect jobs already running or pending in the LSF cluster.
  It only affects LSF's ability to talk to you. The commands will eventually complete.

- If you see these messages Monday through Thursday around 19:35 (7:35PM) we automatically run an LSF reconfiguration during those times.

- Scheduled outages of the LSF cluster are normally announced via the SLAC Computing Outages web page https://www-internal.slac.stanford.edu/comp-out.

LSF Documentation

- SLAC specific LSF documentation.

  http://www.slac.stanford.edu/comp/unix
  Click on "High Performance"

- Platform LSF documentation (available only on the SLAC network).

  http://www.slac.stanford.edu/comp/unix/package/slaconly/lsf/currdoc/html

  http://www.slac.stanford.edu/comp/unix/package/slaconly/lsf/currdoc/pdf/manuals

Problem Reporting

Send email to:

pcds-help@slac.stanford.edu