Accounting CPU time usage on SJTU Pi cluster

The High Performance Computing Center (HPCC) at SJTU
http://hpc.sjtu.edu.cn

Updated at Dec 26, 2013

This document shows you how to how to check your CPU time usage on Pi and how to get a detailed view of historic status of you submitted jobs.

1 Check CPU time usage with `bacct`

LSF command `bacct` is used to check the CPU time consumed by finished jobs. Please notice that as baccount calculates the CPU time only for finished jobs, that for running ones is not included. So the baccnt result is for reference only. The accurate CPU time usage would be the ones displayed on the monthly bills send to you.

```
$ bacct -h
Usage:  bacct  [-h] [-V] [-b] [-l] [-w] [-d] [-e] [-x] [-f logfile]
               [-u 'userlist' | -u all ] [-m 'machinelist'][-M hostlistfile]
               [-q 'queuelist'] [-C time0,time1] [-N host_spec]
               [-S time0,time1] [-D time0,time1]
               [-P 'projectlist'] [-Lp 'licenseProjectList']
               [-U 'reservation ID list' | -U all ]
               [-app applicationProfileList]
               [-sla 'serviceClasslist'] [jobId | "jobId[index]" ...]
```

By default, `bacct` will diplay the total CPU time consumed on all compute queues. Please notice that the time units displayed in bacct is seconds..

```
$ bacct
Accounting information about jobs that are:
  - submitted by users xxx-xxxxx,
  - accounted on all projects.
  - completed normally or exited
  - executed on all hosts.
  - submitted to all queues.
  - accounted on all service classes.
------------------------------------------------------------------------------


SUMMARY:      ( time unit: second )
 Total number of done jobs:      66      Total number of exited jobs:    26
 Total CPU time consumed:   10260.4      Average CPU time consumed:   111.5
 Maximum CPU time of a job:  1813.4      Minimum CPU time of a job:     0.0
 Total wait time in queues: 158519.0
 Average wait time in queue: 1723.0
 Maximum wait time in queue:53149.0      Minimum wait time in queue:    2.0
 Average turnaround time:      1783 (seconds/job)
 Maximum turnaround time:     53164      Minimum turnaround time:         7
 Average hog factor of a job:  5.15 ( cpu time / turnaround time )
 Maximum hog factor of a job:  86.35      Minimum hog factor of a job:  0.00
 Total throughput:             0.04 (jobs/hour)  during 2398.04 hours
 Beginning time:       Sep  9 21:47      Ending time:          Dec 18 19:49
```

To check the CPU time usage consumed by the finished jobs which were submitted from Apr 1, 2013 to Nov 15, 2013, the option `-S` is helpful.

```
$ bacct -S 2013/04/01/00:00,2013/11/15/23:59
```

And option `-q is used to specify a compute queue, on which the time is queried.

```
$ bacct -q cpu
```

Or query on multiple queues

```
$ bacct -q "cpu gpu"
```

Given a job ID, bacct can display CPU time consumed on this job. Here the job with ID 100879 is taken as an exmaple. Option `-l` is used to display long detailed information.

```
$ bacct 100879
```

2 Review job process with `bhist`

LSF command `bhist` is used to review or check the running status of jobs.

```
$ bhist -h
Usage: bhist [-l] [-b] [-w] [-a] [-d] [-e] [-p] [-s] [-r]
             [-f logfile_name | -n num_logfiles | -n min_logfile, max_logfile]
             [-C time0,time1] [-S time0,time1] [-D time0,time1]
             [-N host_spec] [-P project_name]  [-Lp license_project]
             [-q queue_name][-G usergroup_name][-m "host_name ."][-J job_name]
[-Jd job_description]
             [-u user_name | -u all] [jobId | "jobId[index]" ...]
       bhist -t [-f logfile_name] [-T time0,time1]
       bhist [-h] [-V]
```

By default, bhist will show only pending, running and suspended job information during the past week. So more options are required to display the historical information of finished jobs.

The following command displays long (detailed) information about all jobs, which were submitted during 2013-04-01 and 2013-05-01. `-n 100` is commonly used option, which forces bhist to search more log files, up to 100.

```
$ bhist -a -S 2013/04/01/00:00,2013/05/01/23:59 -n 100
```

Option `-l` will make the output more detailed.

```
$ bhist -a -l -S 2013/04/01/00:00,2013/12/01/23:59 -n 100
```

If the job ID is given, only this specific job’ s information will be displayed.

```
$ bhist -a -l -S 2013/04/01/00:00,2013/05/01/23:59 -n 100 345189
Job <345189>, Job Name <HELLO_MPI_256>, User <xxx-xxxxxx>, Project <default>,
C
                     ommand <#BSUB -q cpu;#BSUB -J HELLO_MPI_256;#BSUB -L /bin/
                     bash;#BSUB -o job.out;#BSUB -e job.err;#BSUB -n 512;   MOD
                     ULEPATH=/lustre/utility/modulefiles:$MODULEPATH;module loa
                     d icc/13.1.1;module load impi/4.1.1.036; MPIRUN=`which mpi
                     run`;EXE="./mpihello"; cat /dev/null > nodelist; for host
                     in `echo $LSB_MCPU_HOSTS |sed -e 's/ /:/g'|  sed 's/:n/\nn
                     /g'`;do;echo $host >> nodelist;done; $MPIRUN -machinefile
                     nodelist $EXE>
Wed Dec 18 19:49:18: Submitted from host <mu05>, to Queue <cpu>, CWD <$HOME/wor
                     kspace/example-repo/mpi/hello>, Output File <job.out>, Err
                     or File <job.err>, Re-runnable, Checkpoint period 10 minut
                     e(s), Checkpoint directory </tmp/39086>, 512 Processors
Re
                     quested, Login Shell </bin/bash>;
Wed Dec 18 19:49:40: Dispatched to 512 Hosts/Processors <16*node054> <16*node20
                     1> <16*node202> <16*node008> <16*node067> <6*node297> <16*
                     node306> <16*node072> <13*node292> <16*node308> <16*node27
                     9> <16*node324> <16*node115> <16*node043> <16*node026> <16
                     *node312> <16*node164> <16*node182> <16*node174> <16*node0
                     90> <16*node301> <16*node023> <16*node124> <16*node151>
<1
                     6*node168> <16*node167> <16*node166> <15*node295> <13*node
                     103> <13*node219> <14*node328> <12*node138> <11*node258>
<
                     13*node092> <1*node303> <1*node027>;
Wed Dec 18 19:49:40: Starting (Pid 54488);
Wed Dec 18 19:49:40: Running with execution home </lustre/home/xxx-xxxxx>, Ex
                     ecution CWD </lustre/home/xxx-xxxxx/workspace/example-re
                     po/mpi/hello>, Execution Pid <54488>;
Wed Dec 18 19:49:54: Done successfully. The CPU time used is 604.9 seconds;
Wed Dec 18 19:49:54: Post job process done successfully;

Summary of time in seconds spent in various states by  Wed Dec 18 19:49:54
  PEND     PSUSP    RUN      USUSP    SSUSP    UNKWN    TOTAL
  22       0        14       0        0        0        36
```

3 Reference

- “LSF Manual: bacct” **http://www.cisl.ucar.edu/docs/LSF/7.0.3/command_reference/bacct.cmdref.html**
- “LSF Manual: bhist” **http://www.cisl.ucar.edu/docs/LSF/7.0.3/command_reference/bhist.cmdref.html**