Memory Usage Overview

Category: Memory Usage on Pleiades

Running jobs on cluster systems such as Pleiades requires more attention to the memory usage of a job than on shared memory systems. Below are a few factors that limit the amount of memory available to your running job:

- The total physical memory of a Pleiades compute node varies from 8 GB to 32 GB. A small amount of the physical memory is used by the system kernel; through PBS, a job can access up to about 7.6 GB of an 8-GB node (Harpertown), about 22.5 GB of a 24-GB node (Nehalem-EP and Westmere-EP), and about 31 GB of a 32-GB node (Sandy Bridge)
- The PBS prologue tries to clean up the memory used by the previous job that ran on the nodes of your current running job; if there is a delay in flushing the previous job's data from memory to disks (for example, due to Lustre issues), the actual amount of free memory available to your job will be less
- I/O uses buffer cache that also occupies memory; if your job does a large amount of I/O, the amount of memory left for your running processes will be less

If your job uses more than one node, beware that the memory usage reported in the PBS output file is not the total memory usage for your job: rather, it is the *memory used in the first node* of your job. To help you get a more accurate picture of the memory usage of your job, we provide a few in-house tools, listed below.

qtop.pl
: Invokes *top* on the compute nodes of a job, and provides a snapshot of the amount of used and free memory of the whole node and the amount used by each running process.
For more information, read the article Checking Memory Usage of a Batch Job Using qtop.pl.

qps
: Invokes *ps* on the compute nodes of a job, and provides a snapshot of the %mem used by its running processes.
For more information, read the article Checking Memory Usage of a Batch Job Using qps.

qsh.pl
: Can be used to invoke the command **`cat /proc/meminfo`** on the compute nodes to provide a snapshot of the total and free memory in each node.
For more information, read the article Checking Memory Usage of a Batch Job Using qsh.pl and "cat /proc/meminfo".

gm.x and gm_post.x
: Provides the memory high-water mark for each process of your job when the job finishes.

For more information, read the article Checking Memory Usage of a Batch Job Using qm.x.

These tools are installed under the directory **/u/scicon/tools/bin**. It is a good idea to include this directory in your path by modifying your shell startup script so that you don't have to provide the complete path name when using these tools. For example:

```
set path = ( $path /u/scicon/tools/bin )
```

If your job runs out of memory and is killed by the kernel, this event was probably recorded in system log files. Instructions on how to check whether this is the case are provided in the article Checking if a Job was Killed by the OOM Killer.

If your job needs more memory, read the article How to Get More Memory for your Job for possible approaches.

Article ID: 216
Last updated: 08 Aug, 2012
Computing at NAS -> Best Practices -> Memory Usage on Pleiades -> Memory Usage Overview
http://www.nas.nasa.gov/hecc/support/kb/entry/216/?ajax=1