

Torque

How to get all cgroup cpuset cores with singlejob policy job submissions (without specifying procs)

Issue:

After upgrading to Torque 6.0.x with cgroups enabled, jobs submitted with "-l nodes=1,naccesspolicy=singlejob" get restricted to one core because that's what gets assigned to the cpuset, despite the job being the only one on the node.

Previously, jobs submitted with singlejob or -n (node-exclusive) could access all resources on the compute node. We understand that this is the purpose of cgroups, but we still need a way to enable support for the previous behavior (for legacy scripts).

Solution:

Upgrade to Moab 9.0.2 and Torque 6.0.2, and set the following parameter in moab.cfg:

```
RMCFG[<your Torque name>] FLAGS=MigrateAllJobAttributes
```

This will enable full assignment of cores to the cgroup's cpuset for any jobs submitted with the following syntaxes (these all apply for both qsub and msub):

- -l naccesspolicy=singlejob
- -W x=naccesspolicy=singlejob
- -n
- msub and MWS jobs that get singlejob applied via policy, template, etc.

Notes:

* Jobs submitted with -n will automatically get all node cores assigned to the cgroup cpuset(s), regardless of requested ppn. Jobs that specify singlejob, but do NOT include -n, will get the number of cores stated on the ppn.

* This method has been provided as a bridge. Future scripts should use the -L NUMA Resource Request syntax, described here:

<http://docs.adaptivecomputing.com/torque/6-0-1/help.htm#topics/torque/11-UsingNUMA-Lresource.htm>

Example: qsub -L tasks=1:lprocs=all:place=node

* This method supplants the deprecated "allprocs" job submission flag, which Moab does not support for full cpuset core assignment. Until Adaptive Computing completely discontinues support for the "allprocs" flag, it will still work to assign all cores to the job, as far as Moab's tracking of the node's resources; however, it is,

Torque

effectively, incompatible with systems that have configured Torque with the --enable-cgroups option.

Unique solution ID: #1113

Author: Rick McKay

Last update: 2016-07-12 18:59