

# Torque Resource Manager 6.1.1 Release Notes

The release notes file contains the following sections:

- [New Features](#)
- [Differences](#)
- [Known Issues](#)
- [Resolved Issues](#)
- [Key Documentation Changes](#)

Document build: 03/29/2017 15:08 UTC-06

## New Features

This topic contains a summary of key new features in Torque Resource Manager.

`momctl -u` updates `pbs_server` with a hardware configuration from the MOM.

### 6.1.1

#### *Additional Options for Email Notifications*

Additional % format specifiers are available for including information in email subject lines and messages. For example, %o inserts the name of the job owner and %u inserts a summary of the resources used by the job.

#### *pbs\_mom Switch to Force the Server to Accept a Hardware Update*

`pbs_mom -f` forces the server to accept an update of the hardware on the node. You should use this switch the first time you start the MOM after an upgrade.

#### *CUDA 8 Support*

You can now set a `default_gpu_mode` parameter to control the default GPU mode for jobs that request GPUs.

### 6.1.0

#### *Added Presetup Prologue for MOMs*

The `pres_setup_prologue` executes before the MOMs attempt to set up jobs on the node. This prologue can be used, for example, to set up Kerberos.

#### *Added default\_gpu\_mode Server Parameter*

The `default_gpu_mode` server parameters sets the default value GPU mode. This should facilitate using CUDA 8.

#### *Simplified Torque RPM Installation*

The RPM installation no longer writes the server name into `pbs_mom`'s config file, eliminating the need to configure the MOM config file in the default case.

#### *Simplified Installation Process*

"make install" and the package installer now install service/unit files, set up the path to the Torque libraries, and establish a shell path to binaries for sh and csh shells.

### *Resource Plug-In*

There is now an API for creating resource plug-ins that report custom varattrs, generic resources, generic metrics, and features. Additionally, jobs can report custom resources through the plug-in.

### *-L Submission Recorded in the Accounting Log*

Jobs submitted with the -L request syntax will now have the -L submission recorded in the accounting log.

### *NUMA-Aware cgroup Creation by Per Task or Per Job*

A new Torque server parameter "cgroup\_per\_task" is available to let you specify whether cgroups are created per task or per job. The default is FALSE, meaning jobs submitted with the -L syntax will have *one* cgroup created per host; this behavior is similar to the pre-6.0 cpuset implementation.

The qsub/msub -L syntax is also modified to let you specify whether the cgroup is per task or per job at the job submission time.

**i** Some MPI implementations are not compatible with using one cgroup per task.

### *Change hostname for a Torque Job to Match the Corresponding Moab Job ID*

When jobs are submitted using qsub -J, the job's hostname is set to the job ID of the Moab job.

### *Job Arrays Now Only Partially Instantiated By Default*

Job arrays are now only partially instantiated by default; additional idle jobs are added as the array is executed. This is controlled by the new idle\_slot\_limit server parameter.

A new qsub -i option has also been added to set the slot limit for the job array being submitted. If set for a non-array job, it will be rejected. If the user requests an idle slot limit that exceeds the server parameter's default, the job will be rejected.

### *Recover Array Subjobs*

The new server parameter "ghost\_array\_recover" is added. pbs\_server will now recover array subjobs even when the array (.AR file) couldn't be recovered. This parameter is set to TRUE by default.

### *Improved Job Cleanup Performance*

The number of communications involved in job cleanup has been reduced, resulting in improved performance.

### *Epilogue Script Runs, Even if Output Files Cannot Be Appended*

The epilogue script will now run when `spool_as_final_name` is configured, even if error and output files are not available and cannot be appended.

### *Cray-enabled Torque May Also Be Configured With cgroups*

Cray-enabled Torque may also be configured with cgroups. On the login node, each job will have all of the CPUs and all of the memory controllers in its cgroup.

### *Job Script Path is an Argument to Prologue and Epilogue Scripts*

A new positional parameter contains the full path of a job's job script to the job's prologue and epilogue scripts when the Torque pbs\_mom "mother superior" launches the scripts. The prologue script is the new 8th positional parameter. The epilogue script is the new 11th positional parameter.

### *RPM Install for SLES 12*

You can now use RPM to install and upgrade Torque on SLES 12.

### *qrun Can Interpret the -L Syntax*

You can run jobs that request resources using `qrun -L` without a hostlist.

### *User May Request No Emails be Sent, Even on Failure*

Specifying `-m p` at `qsub` time makes it so no emails are sent for the job, even on failure.

### *User settable kill\_delay Through qsub -K Option*

Added a user settable, per-job kill delay, called `kill_delay`. It is settable via the new `qsub -K` option.

### *Torque Commands Appear in User's Login Shell Path*

RPMs created via `build-torque` now create `/etc/profile.d/torque.sh` (and `torque.csh`) files so that Torque commands appear in a user's login shell path.

### *Option to Email User on All Non-zero Exit Codes*

The `qsub -m f` option sends an email if a job has a non-zero exit code. The `qsub -m f` option can be used with the `a`, `b`, and `e` options, but not with `n` or `p`.

### *Added an Option for When to Send Mail Notifications*

Added a `-m f` option to `qsub` to send mail when a job terminates with a non-zero exit code.

### *Option to Disable Reading of RUR Information*

Added `$cray_check_rur` configure option to disable reading of Resource Utilization Reporting (RUR) energy usage for Cray login nodes. If set to false, login MOMs will not look at the energy resource information used for each job. Disabling this may improve performance.

### *Improved Job Submission and Start Time Performance*

Transactions per job submitted have been reduced from 11-13 to 3, improving job submission time by an average of 11-13%.

### *Added `ghost_queue` Queue Attribute*

If `pbs_server` restarts and recovers a job but cannot find that job's queue, a new queue with the 'ghost\_queue' attribute is created. That job is added and then run on the new queue. Jobs cannot be added to ghost queues. Once an administrator reviews and corrects the queue's settings, the 'ghost\_queue' attribute can be removed, allowing the new queue to function normally.

### *New "`email_batch_seconds`" Server Parameter*

The new server parameter "`email_batch_seconds`" lets you control at what frequency a batch of emails are sent to each user.

### *MOM Parameter to Overwrite Output Files*

The `$force_overwrite` MOM parameter, if set to true, forces the output files to be overwritten each time a job is started.

### *Option to Turn Off `trqauthd` Logging*

Added a `-n` option to `trqauthd` that disables `trqauthd` from logging anything.

## Differences

This section contains differences in previously existing features that require a change in configuration or routine.

### 6.1.1

#### *\$CUDA\_VISIBLE\_DEVICES is not set by default*

\$CUDA\_VISIBLE\_DEVICES is not set by default if you're using cgroups. Also, this makes its contents compatible with varying CUDA versions.

### 6.1.0

#### *Use --enable-cgroups Recommended; --enable-cpusets is Deprecated*

--enable-cgroups is the recommended configuration option. When cgroups are supported, cpusets are handled by the cgroup cpuset subsystem.

--enable-cpusets is deprecated; meaning that the functionality is still supported; however, no new features will be added.

#### *pbs\_server Does Not Shut Down When trqauthd is Stopped*

The *trqauthd* service is no longer a requirement of *pbs\_server*, so *pbs\_server* no longer shuts down when *trqauthd* is stopped.

#### *TorqueConfigure Script Does Not Enable the GUI Component*

When the *devel-tk* and *devel-tcl* libraries have been installed, the *Torqueconfigure* script does not enable the GUI component. To build the GUI component, specify *--enable-gui* when building from source.

#### *Default Value for keep\_completed Parameter Changed*

The default value for the Torque *keep\_completed* parameter has been changed to 300 seconds. The *keep\_completed* parameter specifies the number of seconds jobs are kept in the job execution queue after they are completed. Keeping completed jobs in the queue enables Torque to report on the status of the jobs.

#### *legacy\_vmem Server Parameter Affects Behavior of the -l vmem Option*

*legacy\_vmem* is a new server parameter that affects the behavior of the *-l vmem* option. When set to true, the *vmem* request will be the amount of memory requested for each node of the job. When it is unset or false, *vmem* will be the amount of memory for the entire job and will be divided accordingly.

### *Queue Support for Both resources\_default.\* and req\_information\_default.\* Settings*

When queues have both resources\_default.\* and req\_information\_default.\* set then they are applied according to their resource request type. resources\_default.\* settings are applied to jobs that do not explicitly use the -L syntax, while req\_information\_default.\* settings are applied only to jobs that explicitly use the -L resource syntax.

### *Prohibited Mode Not Allowed for User Jobs*

Setting the compute mode of an NVIDIA GPU to prohibited makes it so the GPU cannot be used at all. In previous versions of Torque users were allowed to set a GPU to prohibited mode. But then it could not set the mode to anything else since the GPU was now prohibited. This change went into effect for version 5.1.3, 6.0.2, and later.

### *Cray-Enabled Torque May Be Configured with cgroups*

Support is added for Cray-enabled Torque configured with cgroups.

- On the login node, each job will have all of the cpus and all of the memory controllers in it's cgroup.

### *Support for Single Job Dependencies and Array Dependencies at the Same Time*

Jobs can depend on single job dependencies and array dependencies at the same time.

### *Added a Way to Exclude Compute Nodes From allow\_node\_submit.*

Set 'allow\_node\_submit=true' in qmgr, and then add a remote mom to node\_submit\_exceptions with 'set server node\_submit\_exceptions = <nodename>' in qmgr. This setting will disable job submissions on the <nodename> specified by the user. Previously, the process of disallowing compute nodes from submitting jobs would require making a separate entry in qmgr for each individual compute node.

### *Reduced the Number of Logging Statements*

Reduced the number of logging statements when a node isn't up and therefore can't receive the mom hierarchy.

### *Added tcp\_incoming\_timeout Server Parameter*

tcp\_incoming\_timeout specifies the number of seconds before incoming connections timeout. tcp\_timeout now specifies the timeout for outgoing connections or connections initiated by pbs\_server. tcp\_incoming\_timeout functions exactly the same as tcp\_timeout, but governs incoming connections while tcp\_timeout governs only outgoing connections (or connections initiated by pbs\_server).

### *pbs\_server Enhancement for Very Large Number of Jobs.*

pbs\_server has been enhanced to better handle a very large number of jobs (several hundred thousand or more) by enabling an alternate way for it to store job-related files in the directories \$PBS\_HOME/server\_priv/jobs and \$PBS\_HOME/server\_priv/arrays.

A new boolean server attribute, `use_jobs_subdirs`, lets an administrator direct the way pbs\_server will store its job-related files. When `use_jobs_subdirs` is unset (or set to false), job and job array files will be stored directly under \$PBS\_HOME/server\_priv/jobs and \$PBS\_HOME/server\_priv/arrays. This is the default behavior and the way the server has stored these files in the past. When `use_job_subdirs` is set to true, job and job array files will be distributed over 10 subdirectories under their respective parent directories. This method helps to keep a smaller number of files in a given directory.

If an administrator wishes to change the `use_jobs_subdirs` attribute from its previous value (or when setting it to true when it has not previously been set), it is highly recommended that Torque be drained of all jobs. Failing to take this action may result in the loss of existing jobs.

## Known Issues

This topic lists known issues in Torque Resource Manager. Following each issue description is an associated issue number in parentheses. Known issues are aggregated and grouped by the release version for which they were first reported.

### Torque Resource Manager

#### 6.1.1

- Updating pbs\_server when GPU jobs are running can result in job issues. Make sure no GPU jobs are running before updating pbs\_server. (DOC-3132)

#### 6.1.0

- qstat -f accounting logs sometimes report "resources\_used.vmem=0". This is due to how Linux reports memory usage. (TRQ-3760)
- Torque will not compile when the tk-devel and tcl-devel packages are installed on the host. To work around this, disable building of the gui component by using --disable-gui when executing configure. (TRQ-3723)



## Resolved Issues

Resolved issues are aggregated and grouped by the release version in which they were resolved. When applicable, each resolved issue has the associated issue number in parentheses.

### 6.1.1

- qsub did not properly validate the working directory (-w) path when VALIDATEPATH was unset or set to FALSE. (TRQ-3929)
- Jobs were sometimes lost during pbs\_mom systemd unit file shutdown. (TRQ-3927)
- The shutdown function for pbs\_server in the init.d script was sometimes failing. (TRQ-3924)
- A large number of acl\_users could deadlock pbs\_server. (TRQ-3910)
- Command-line directives did not override submit filter directives for interactive jobs. (TRQ-3862)
- GPU models k40s and k80s were not detected correctly when using GPUs and cgroups. (TRQ-3861)
- pbs\_server crashed when loading jobs after a server shutdown. (TRQ-3839)
- stdout was not properly delivered on NFS file systems when Torque was configured with --disable-spool. (TRQ-3792)
- Torque was not allocating resources correctly for cgroup jobs. (TRQ-3790)
- Torque had communication issues when IPv6 was configured but not used. (TRQ-3773)
- Torque was not terminating all processes in a cgroup before removing it. (TRQ-3753)
- Job holds were not updated when a slot limit was changed for a job array. (TRQ-2360)

### 6.1.0

- Systemd unit files do not correctly start/stop Torque daemons. A new -F (don't fork) switch has been added to pbs\_server, pbs\_mom, and trqauthd. (TRQ-3795)
- pbs\_server was not starting when an array dependency was cleared before pbs\_server started. (TRQ-3774)
- Torque was not waiting for a compute node to do a copy out process. (TRQ-3762)

- Fixed node locking issues that were causing pbs\_server to abort when built with newer versions of the threading library. (TRQ-3755)
- Server build was failing with --disable-spool configure option. (TRQ-3751)
- Forced requeue command (qrerun -f) was not obeying kill\_delay settings. (TRQ-3737)
- When adding a dynamic node, the server was initially putting the node in free state, but not reporting this state to the server. (TRQ-3733)
- The memory cgroup was not being set correctly for jobs requesting pmem and procs. (TRQ-3728)
- Jobs were not being removed from node\_usage files after ending. (TRQ-3727)
- Jobs were becoming stuck after not starting correctly. (TRQ-3720)
- pbs\_server aborted under ghost array recovery when job array files were missing. (TRQ-3719)
- momctl was not displaying the \$varattr script or the correct time and date. (TRQ-3712) (TRQ-3617)
- pbs\_mom crashed when started on the reporter node without \$apbasil\_protocol set in the config. (TRQ-3715)
- Torque was not allocating enough memory controllers to satisfy memory requests. (TRQ-3681)
- pbs\_server was not being properly shut down when in HA mode. (TRQ-3670)
- pbs\_server was not detecting and updating total\_threads when a node's hyperthreading was enabled. (TRQ-3662)
- pbs\_server was not properly restarted when running "service pbs\_server restart" during installation. (TRQ-3657)
- Memory and swap limits were not set in cgroup. For information on how memory and swap options are used, see 1.1 -L NUMA Resource Request in the *Torque Resource Manager Administrator Guide*. (TRQ-3656)
- Jobs submitted with -l option with exclusive access to the node were not receiving all CPUs and memory controllers in the cgroup. (TRQ-3649)
- Corrected logging to only log that a signal is sent to a process when it is actually issued. (TRQ-3638)
- pmem was not getting set correctly. With cgroups enabled, pmem is the amount of resident memory allocated per process where are process is given by the value of ppn. For example: `qsub -l nodes=1:ppn=2,pmem=250mb` will allocate a total of 500 MB on the node where the job is run, 250 MB per ppn. (TRQ-3628)

- NUMA -L syntax defaulted to override user-specified parameters. (TRQ-3623)
- `qstat -x` returned nothing (instead of an empty XML document) when there are not jobs queued. (TRQ-3622)
- Jobs in which a task required more than one socket could not be started using NUMA -L syntax. (TRQ-3618)
- `$usecp` parameter was ignored when specifying which directories should be staged. (TRQ-3613)
- Server deadlocked when `job_save()` failed. (TRQ-3605)
- Tasks' memory usage was sometimes not reported. (TRQ-3601)
- Interactive jobs skipped submit filter directives if the first line was not `#PBS`. (TRQ-3585)
- Client commands were not attempting to contact the fallback server when the primary server was down. (TRQ-3582)
- Crash/infinite loop when loading certain node usage files. (TRQ-3576)
- Issue reported with alps login nodes. Updated cpusets for alps login nodes so that all of the cpus are in the job's cpuset. (TRQ-3568)
- Torque crashed intermittently when using the -L syntax. (TRQ-3566)
- Torque was not reporting nodes in the correct order. (TRQ-3559)
- Torque returned non-specific network failure messages to Moab. (TRQ-3539)
- Completed jobs were still reported in `pbsnodes`. (TRQ-3525)
- A deadlock occurred when handling job dependencies. (TRQ-3519)
- `cgroup` directories were not removed when jobs were completed. (TRQ-3515)
- `drmaa` unable to link with Torque. (TRQ-3511)
- Memory calculation issues reported when `cgroups` enabled and `-l vmem|pmem|mem` are used. (TRQ-3499)
- Epilogue not showing up in `momctl -d3` output. (TRQ-3495)
- Job dependencies were not being cleared with High Availability server. (TRQ-3477)
- Array subjobs did not have a queued entry in the accounting log. (TRQ-3470)
- Segfault in `create_alps_subnode` with `node_note` populated. (TRQ-3445)
- Problems with clearing a node note. Removed length restriction on a node note. (TRQ-3439)

- Jobs that never ran were receiving end records. (TRQ-3432)
- Array templates were being reported as jobs. (TRQ-3405)
- A shell escape in pbs\_mom's config file when specifying GRES did not show up in pbsnodes or Moab. (TRQ-3393)
- Resources\_used.walltime changed to seconds from HH:MM:SS in accounting logs. (TRQ-3385)
- libtorque.so was not being created. (TRQ-3374)
- pbs\_server timed out connection to pbs\_mom. Added load balancing to login nodes when they start to get busy. (TRQ-3367)
- pbs\_mom would hang when sending status from a child. Added a timeout for node health check scripts so that they cannot make the mom daemon hang. (TRQ-3364)
- pbs\_mom hangs on restart with init script. Ensured that necessary services have been brought up before starting the Torque daemons and that the Torque daemons are shutdown before their required services are shutdown. (TRQ-3345)
- Numbered directories in server\_priv/jobs (and arrays) were missing when the server attribute use\_jobs\_subdirs was set to TRUE. (TRQ-3185/)
- qrls gave no response and logged no problem when a failure occurred due to a slot limit restriction. (TRQ-3328)
- Fixed a memory leak when jobs were being started asynchronously. (TRQ-3326)
- qsub -W stage-in was not working. Fixed failures where the group name showed up in the log as the problem but the user did not belong to the group name given in the error. (TRQ-3312)
- Multiple moms sent invalid destroy\_alps\_reservation/req\_delete\_reservation. Only allows one kill orphaned reservation request per reservation at one time. (TRQ-3299)
- Problems building RPMs on Red Hat 6/CentOS 6 systems. (TRQ-3283)
- Jobs with square brackets in the name were aborted on restart if they weren't array subjobs. An issue was fixed with jobs getting aborted if they are named with "[" in the name but aren't Torque array jobs. (TRQ-3214)
- Logs filled with messages about not sending hierarchy to mom. Failures are only logged the first time it can't send the hierarchy to a mom. (TRQ-3156)
- Jobs started even if mother superior could not resolve the hostname for a sister node. (TRQ-3134)

- qsig was not working correctly when `display_job_server_suffix = false`. (TRQ-3102)
- Error condition where the mom's port would be inserted into the .JB file name. (TRQ-3090)
- Torque was not able to release holds on job arrays. Running `qrls` on an array subjob allows `pbs_server` to correct slot limit holds for the array to which it belongs. (TRQ-3088)
- Down/offline nodes caused TORQUE to not online elastic nodes. `pbs_server` is now able to bring up new nodes even when there are nodes in the system that are down or offline. (TRQ-3066)
- Completed jobs were not getting cleaned up. Fixed various issues when restarting dependency jobs, including them not getting removed even after completion. (TRQ-3175)
- Several log messages were unclear. (TRQ-2860)
- `qsub -v` was not processing environment variables with no value correctly. (TRQ-2699)
- Node recovers when behind processing requests. `pbs_server` now detects when a node is failing too frequently and mark it down temporarily if this happens. Once a node is marked down, it will be marked up again if either two consecutive communications from `pbs_server` to the node receive successful replies, or after five minutes of staying offline (whichever comes first). A node is considered to be failing too frequently if it has three failures to reply to a server request without have two consecutive successes in between. (TRQ-2517)
- `qstat -f` was not displaying `init_work_dir`. (TRQ-2459)
- Job holds were not updated when the slot limit was changed for a job array. (TRQ-2360)
- Torque enabled unused `scheduler_iteration` parameter by default. (TRQ-2161)

## Key Documentation Changes

This topic lists some key documentation changes made for Torque Resource Manager 6.1.1.

Features new to this release have been documented but are not listed on this page. For information about those new features and links to their respective documentation, see [New Features](#).

In this topic:

- [Torque Resource Manager](#)
- [Online Help](#)

### Torque Resource Manager

This section lists documentation-related changes to the Torque Resource Manager Administrator Guide.

#### [6.1.1](#)

- Updated description of momctl -C.
- Added documentation for pbs\_mom -w option.
- Add notes regarding --enable-geometry-requests is not compatible with --enable-cgroups.

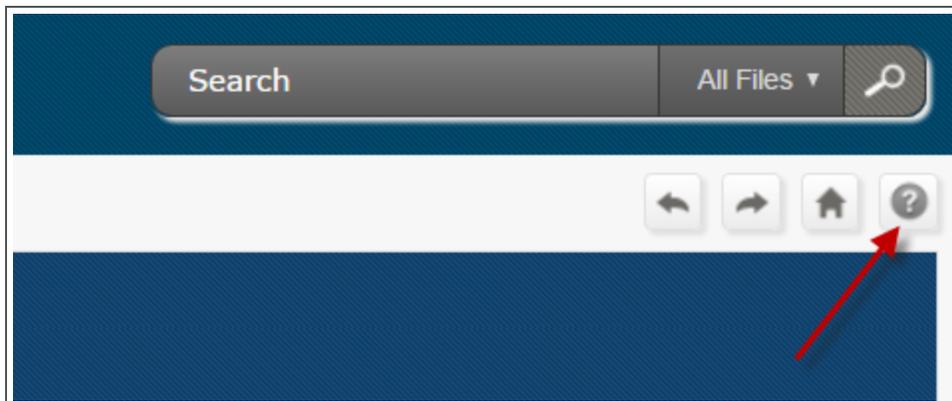
#### [6.1.0](#)

N/A

### Online Help

This section lists documentation changes made for the release.

To access the online help, click the ? icon in the toolbar.



9.1.1/6.1.1

Improved search filter sets for all online documentation.

9.1.0/6.1.0

Improvements have been made for searching the online documentation. In addition to the search filter sets added in 9.0.1, information is available on how to specify the search terms to limit the results displayed.