

Moab HPC Suite 8.1.1 Release Notes

The release notes file contains the following sections:

- [New Features on page 2](#)
- [Differences on page 5](#)
- [Installation and Upgrade Information on page 10](#)
- [Known Issues on page 11](#)
- [Resolved Issues on page 13](#)

New Features

The following is a summary of key new features in Moab HPC Suite.

- [All Components on page 2](#)
- [Moab Workload Manager on page 2](#)
- [Moab Web Services on page 3](#)
- [TORQUE Resource Manager on page 3](#)

All Components

[8.1.1/5.1.1](#)

Support for more OSs

These additional OSs are now supported:

- CentOS 7.x
- RHEL 7.x
- Scientific Linux 7.x
- SUSE Linux Enterprise Server 12

i Support for Red Hat 7-based and SUSE 12-based systems requires the 8.1.1.2/5.1.1.2 maintenance release or later.

Moab Workload Manager

[8.1.1](#)

No new features.

[8.1.0](#)

Elastic Computing Feature - Ability to Request Dynamic Nodes From an External Service

A new Elastic Computing feature is available to allow the Moab scheduler to utilize systems that can temporarily provide additional nodes (for example, to create new virtual machines or borrow physical nodes from another system) to fulfill increased workload demand so that job backlog is completed in a reasonable time frame. When this feature is enabled and configured, Moab accesses the dynamic nodes, also known as bursting, to handle the increased workload. Accessed nodes are then released once the demand is filled.

i Elastic Computing is only available with a Moab HPC Suite - Enterprise Edition license that has Elastic Computing enabled. Please contact your account manager at Adaptive Computing for further details and requirements for this feature.

Moab's Scheduling Cycle No Longer Needs to Wait on the Cluster Query

A new "threadedqueries" resource manager flag is available. When this flag is set for an individual RM (for example, "RMCFG[torque] TYPE=PBS FLAGS=threadedqueries"), the queries that Moab performs to get information from the RM are done in a separate thread from the main Moab process. This allows Moab to remain responsive during the query, and ultimately reduces the time spent in a scheduling cycle. If multiple RMs are being used, the effect can be more significant because all RMs will be queried in parallel.

Moab Web Services

8.1.1

No new features.

8.1.0

Support for Multi-Line (textarea) Configuration Parameters in Plugins

MWS now supports multi-line (textarea) configuration parameters in plugins. See **Configuration Constraints** in the *Moab Web Services Reference Guide* for more information.

Trigger Object includes New Type Field

The Trigger object in MWS has a new field called type. The type of a trigger can be either generic or elastic.

TORQUE Resource Manager

5.1.1

See [All Components on page 2](#).

5.1.0

Ability to Provide Condensed qstat Output

A 'qstat -C' option, which specifies that TORQUE will provide only a condensed output (job name, resources used, queue, state, and job owner) for jobs that have not changed recently (as per the job_full_report_time parameter), has been added. Jobs that have recently changed will continue to send a full output.

Performance Enhancements to MOM Clean-up Time

Some minor performance enhancements were made to improve MOM clean-up time.

Differences

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

- [All Components on page 5](#)
- [Moab Workload Manager on page 5](#)
- [Moab Web Services on page 7](#)
- [TORQUE Resource Manager on page 7](#)

All Components

[8.1.1/5.1.1](#)

Daemon Restart

If your configuration uses systemd to start or stop daemons, also use systemd to restart daemons instead of using the direct restart options.

For example, if you use `systemctl start moab.service`, use `systemctl restart moab.service` instead of the `mschedctl -R` option.

i Support for systemd (Red Hat 7-based and SUSE 12-based systems) requires the 8.1.1.2/5.1.1.2 maintenance release or later.

Moab Workload Manager

[8.1.1](#)

Set Job Flags in Identity Manager

Enabled the `JOBFLAGS` parameter on accounts through `moab.cfg` and identity managers.

New PREEMPTIONALGORITHM Parameter

The `PREEMPTIONALGORITHM` is added to designate how Moab handles preemption scheduling policies. Valid values are `PREEMPTORCENTRIC` or `PREEMPTTEECENTRIC`. `PREEMPTTEECENTRIC` is the default.

- `PREEMPTORCENTRIC` specifies Moab uses the normal scheduling policy and obeys all configured policies (such as `JOBNODEMATCHPOLICY`, `NODEALLOCATIONPOLICY`, `NODEACCESSPOLICY`). Previously, Moab did not support those policies for preemption.

- PREEMPTCENTRIC specified Moab uses the custom scheduling policy that ignores many policies to ensure the fewest and least important (by priority) preemptees are disturbed by the preemptor.

i Preemption now works with JOBMATCHPOLICY EXACTNODE.

Change to ALWAYS-EVALUATE-ALL-JOBS Configuration Parameter

The configuration parameter ALWAYS-EVALUATE-ALL-JOBS was changed from a boolean to an enumerated value. The possible values are ALWAYS (formerly TRUE), FIRSTRSV (formerly FALSE), and FULLRSV (an intermediate setting).

i No change is required when upgrading from earlier versions. The TRUE value will map to ALWAYS and the FALSE value will map to FIRSTRSV.

RSVSEARCHALGO by partition

Enabled "PARCFG[] FLAGS=WideRsvSearchAlgo" to allow for per-partition specific scheduling rules. See the **RSVSEARCHALGO** parameter in the *Moab Workload Manager Administrator Guide*.

FSSCALINGFACTOR Pre-Partition Setting

Enabled "PARCFG[] FSSCALINGFACTOR" for partition-specific fairshare usage scaling.

msub MOAB_SUBMITDIR Environment Variable

MOAB_SUBMITDIR is populated with the submission directory for jobs submitted with "msub -E".

msub -P

Added '-P' option to msub command to match functionality with 'qsub -P'. This option can only be used by users in the ADMINCFG[1] security level.

Clarify Jobs Submitted with -n

Added log entry when job submitted with '-n' (node_exclusive). 'n' sets the node's access policy as SINGLEJOB.

New Argument for checknode -h

"ALL" is now a valid argument in 'checknode -h'.

Mandatory Queuetimes on Reservation Owners

Added new parameter "SRCFG[] OWNERPREEMPTQT=XX:XX" to configure mandatory queuetimes on reservation owners.

checkjob Change for Node Availability Information on Large Clusters

checkjob will no longer report node availability information on large (>1000 node) clusters unless a '-v' flag is included.

mrsvctl -m Disallows the Modification of a Standing Reservation

mrsvctl -m now disallows the modification of a standing reservation and returns an error to the user if this is attempted.

8.1.0

Set Default Accounts on a Per Partition Basis

Added the ability to define default accounts per partition. Also available in fairshare trees.

Node Collection in the Moab Database in MongoDB

The node collection in the Moab database in MongoDB has an index on the attributes field. This field can grow too large to index.

- For existing installations, the following commands on the MongoDB server will fix the problem:

```
$ mongo moab -u moab_user -p secret2  
> db.node.dropIndex({"attributes":1})
```

The username and password for your database are most likely different from the above example. Check with your database administrator.

- For new installations using this and future releases, the index is no longer created and does not need to be dropped.

IDCFG[] Defaults to TRUE

The default value for IDCFG[] CREATECRED has been changed to TRUE. Moab will now create all credentials that it finds in the identity manager.

Moab Web Services

8.1.1

No known differences.

8.1.0

Reservation Statistics Value Changes

MWS reservation statistics (CIPS, CAPS, TAPS, and TIPS) values have been changed from floating decimal points (double) to long integers. This supports reservations now passing the consumption rate as an attribute (instead of an element).

TORQUE Resource Manager

5.1.1

\$prologalarm is Always Honored

\$prologalarm was ignored on the prologue for a job. Also when the epilogue was run the \$prologalarm value was ignored if it was more than 300. Now the \$prologalarm value is always honored regardless of how large it is for both prologue and epilogue scripts. The default timeout is still 300 seconds.

Disable the Automatic Requeuing of Jobs

Added the ability to disable the automatic requeuing of jobs due to transient failures.

pbs_mom now sets environment variable for NVIDIA GPUs

A new mom config parameter, \$cuda_visible_devices, was added to specify whether pbs_mom sets the CUDA_VISIBLE_DEVICES environment variable when it starts a job. The default is TRUE.

Log Milliseconds

Added milliseconds in TORQUE's log files.

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.

5.1.0

CLIENTRETRY Configuration Option Support for qdel

The qdel parameter now includes the -b option and CLIENTRETRY configuration option support. This feature functions similar to -b option for qsub. Specifically:

- -b <num> command line argument support
- CLIENTRETRY configuration option support
- PBS_CLIENTRETRY environment variable support

Jobs Deleted When the Dependency Can No Longer Be Satisfied

When a job is deleted because its dependency can no longer be satisfied, that job will follow the keep_completed parameter set (if any) for pbs_server. Previously, jobs were purged immediately.

pbs_server -t No Longer Supports hot|warm|cold Options

The pbs_server -t option no longer supports the hot|warm|cold options. Other options are still supported.

Installation and Upgrade Information



When installing or upgrading, it is *strongly* recommended that administrators configure Moab with mauth authentication with a complex key value. See **Mauth Authentication** in the *Moab Workload Manager Administrator Guide* for more information.

- [Compatibility Requirements on page 10](#)
- [Installing Moab HPC Suite 8.1.x on page 10](#)
- [Upgrading to Moab HPC Suite 8.1.x on page 10](#)

Compatibility Requirements

This section provides information on compatibility between the different components of the suite.

Moab Workload Manager and TORQUE Resource Manager

Although the recommended configuration is Moab version 8.1.x and TORQUE version 5.1.x, Moab version 8.1.x supports TORQUE version 4.2.9, 4.2.10, 5.0.x, and 5.1.x.

TORQUE 5.1.x requires Moab 8.1.x or 8.0.x.

Installing Moab HPC Suite 8.1.x

Please see **Requirements** and also see **Preparing for Installation** and **Installing Suite RPM** for manual or RPM-based installation instructions, respectively, in the *Moab HPC Suite Installation and Configuration Guide*.

Upgrading to Moab HPC Suite 8.1.x

Please see **Preparing for Upgrade** and **Upgrading from 7.2 or Upgrading from 8.0** for manual or RPM-based installation instructions, respectively, in the *Moab HPC Suite Installation and Configuration Guide*.

Known Issues

The following are known issues in Moab HPC Suite. Following each issue description is an associated issue number in parentheses. Known issues are aggregated and grouped by the release version for which they first occurred or were reported.

8.1.1/5.1.1

- When altering a GRES with 'mjobctl -m' on a job submitted with "-l software=" (instead of with "-l gres="), the change incorrectly reverts after an iteration. As a workaround, use '-l gres=' instead of '-l software='. The 'software' syntax will be deprecated in favor of 'gres'. (MOAB-7631)
- Requesting multiple GRESes with "-l software=" honors only the first license request. Use "-l gres=" instead. The 'software' syntax will be deprecated in favor of 'gres'. (MOAB-7630)
- Job user priority lost after running 'mjobctl -m userprio'. (MOAB-8094)
- routing queue jobs do not exit the routing queue if the job is submitted with a hold (e.g. qsub -h). (TRQ-2788)

8.1.0/5.1.0

- Jobs submitted with invalid credentials are put in a held state, instead of rejected, until the administrator can respond. The checkjob command gives administrators further information regarding why the job is held. Blindly assuming that all held jobs should in fact be running RIGHT NOW is not only unsafe, but circumvents intentional Moab policies and workflow. An administrator should exercise care when resolving held jobs. (CVE-2014-5375, MOAB-7478, MOAB-7526)
- When installing or upgrading, it is *strongly* recommended that administrators configure Moab with mauth authentication with a complex key value. See **Mauth Authentication** in the *Moab Workload Manager Administrator Guide* for more information. (CVE-2014-5376, MOAB-7525, MOAB-7480)
- If the Moab **JOBNODEMATCHPOLICY** is set to **EXACTNODE**, and if `requirements.tasksPerNode` is used in a job submission to MWS, then Moab will double the resources requested. To avoid this problem, use `requirements.resourcesPerTask.processors.dedicated` instead. (MOAB-7424)
- BACKFILLPOLICY BESTFIT does not support multi-req jobs. Only FIRSTFIT supports multi-req jobs. (MOAB-6824)
- Some limitations exist in the way that pbsdsh can be used. Please note the

following situations are not currently supported:

- Running multiple instances of pbsdsh concurrently within a single job. (TRQ-2851)
- Using the -o and -s options concurrently; although requesting these options together is permitted, only the output from the first node is displayed rather than output from every node in the chain. (TRQ-2690)

Resolved Issues

The following is a list of some key bugs fixed in Moab HPC Suite. Following each issue description is an associated issue number in parentheses.

Resolved issues are aggregated and grouped by the release version in which they were resolved.

8.1.1/5.1.1

- **Moab failed to register GRES update via qalter.** Fix applied to update changes made to job gres via qalter. (MOAB-7559)
- **showhist.moab.pl was missing output file and working directory in the output.** Added ability to display the output file and working directory in the output of showhist.moab.pl. (MOAB-5609)
- **mnodectl -m features with regex only updates one node.** Enabled mnodectl -m <features> x:<node_regex> for node features. (MOAB-7843)
- **Moab is calling salloc/srun with wrong options.** Changed call for Slurm's srun command from "-n1 -N1" to "-n<num_tasks>". (MOAB-6770)
- **mjobctl 'command-line arg info not available' is logged.** Added "starttime" to the event logged when either the 'showstart' or 'mjobctl -q starttime' command is run. (MOAB-7230)
- **Group missing from showq -b.** Added GROUP to output of "showq -b" and "showq -b -v". (MOAB-6762)
- **Unable to disable a RM that is down.** Added ability to set state of down RM to disabled. (MOAB-7481)
- **Unable to clear failures for ID in mdiag -R output.** Added an ID option to mrmctl -f to clear the ID failure messages in mdiag -R. (MOAB-7259)
- **Moab was reading in old job information and using it.** Added log warning if job id from RM is already found in the completed job table. (MOAB-7196)
- **mysql error code was missing in the list of ignored SQL error codes.** This issue is fixed. (MOAB-7758)
- **runjob ignores policies.** Added the enforcing of partition limits to runjob (mjobctl -x). (MOAB-6132)

- **Job submission rejected for insufficient resources yet still created template jobs.** Cleaned up job workflows that could not be submitted. (MOAB-7919)
- **Moab peer to peer grid with LOCALWORKLOADEXPORT results in livelock.** Enabled "RMCFG[] TYPE=MOAB FLAGS=asynccdelete" for more responsive grids. (MOAB-7415)
- **GPU usage did not show up in workload trace.** This issue is fixed. (MOAB-7498)
- **Potential race condition when node idle purge time exceeded.** Enabled automatic node draining when NODEIDLEPURGETIME is reached. (MOAB-7663)
- **Moab not escaping properly when sending XML to MAM.** Enabled sanitizing of credential names. (MOAB-7419)
- **Erroneous message reported about classes when nodes are down.** This issue is fixed. (MOAB-7770)
- **Moab log is incorrect.** Fixed log message about nodes located for job. (MOAB-7788)
- **"could not get user info" message displays for job triggers with flag "user".** Set job trigger exec user to default user of job if not explicitly set. (MOAB-7907)
- **Checkjob reporting misleading error.** Suppressed non-essential error output when unable to connect to the CLIENTUIPORT. (MOAB-7539)
- **Jobs not taking all procs when "flags=allprocs" is requested on the job and "set queue batch resources_default.ncpus = 1" is set in TORQUE.** This issue is fixed. (MOAB-7748)
- **Moab was scheduling jobs before setting up the rsv event table.** This issue is fixed. (MOAB-7953)
- **MPBSWorkloadQuery failure caused problems with jobs.** Enabled RMCFG[] FLAGS=NoCondensedQueries to disable Torque's condensed qstat queries. (MOAB-7958)
- **Problems reported with credential REST queries.** Changed max_idle_jobs, max_jobs, max_nodes, max_processors, and max_processor_seconds from integer to string. (WS-2388)
- **Queue default resources were being modified.** Corrected an issue where the all memory values were converted to bytes after a restart. (TRQ-3139)
- **Some completed jobs were taking quite a while to clear out.** Hardened the removal of completed jobs from pbs_server. (TRQ-3044)

- **Ability to have the procct values displayed on the queues where it was set was lost.** This issue is fixed. (TRQ-3135)
- **max_user_queuable issues.** Fixed some bugs causing incorrect max_user_queuable values. (TRQ-2841)
- **PBS_NUM_NODES on multi-req job wrong when using specific hostnames.** This issue is fixed. (TRQ-1949)
- **Dependency jobs fail to start during HA fail over.** Dependencies now display in the format of depend=type:jobid instead of depend=type:jobid@server. (TRQ-2332)

```
# Was
depend=afterok:12.napali@napali

# Now
depend=afterok:12.napali
```

- **Bug reported that caused jobs to not start when ALPS incorrectly returns a permanent confirmation failure.** This issue is fixed. (TRQ-3023)
- **Problem reported that caused mom restarts to intermittently fail.** This issue is fixed. (TRQ-2307)
- **TORQUE will not compile with --enable-debug on configure.** This issue is fixed. (TRQ-2969)
- **pbs_server would count completed jobs against queue limits when pbs_server was restarted.** This issue is fixed. (TRQ-3087)
- **Compute node offline state doesn't persist across server restarts.** This issue is fixed. (TRQ-2790)
- **With kill_delay and \$exec_with_exec set, a job would be set to a completed state after running qrerun instead of getting set back to queued.** This issue is fixed. (TRQ-2993)
- **qmgr refuses numerical node ids.** This issue is fixed. (TRQ-2946)
- **X11 very slow with qsub -I -X.** The port forwarding buffer size has been increased to improve performance when enabling X11 forwarding from qsub (-X switch). (TRQ-2740)
- **Problem reported with interactive jobs not staying on the node from which they were submitted.** This issue is fixed. (TRQ-3122)
- **Jobs were getting stuck in a running state when an asynchronous run failed.** This issue is fixed. (TRQ-3114)
- **Array slot limits were not getting decremented when a job is preempted or rerun.** This issue is fixed. (TRQ-3110)

8.1.0/5.1.0

- **mrmctl -f {messages|stats} failed to work for AM.** Fixed bug so that messages and statistics can be cleared for AM as described in the documentation.
- **Several mrmctl options ignored the option argument or did not honor them properly.** These issues are fixed.
- **Role permission updates were incorrectly applied to user permissions.** This issue is fixed. (WS-2340)
- **multireq jobs take hours to start.** Fix bug where multi-req jobs were slow to start in certain cases. (MOAB-6824)
- **Unauthorized error does not tell you what permission you are lacking.** Improved error messages related to accessing resources without sufficient permissions. (WS-2301)
- **Submitting remote job scripts (that is, including the script as part of the JSON payload) is not currently supported.** Job scripts can now be included in the POST body when submitting jobs via MWS. The new field is called commandScript. (WS-2112)
- **MWS could send phony node name to Moab.** MWS is now more robust when parsing the nodesRequested field during a job PUT. (WS-2352)
- **Reservation trigger parser is broken.** The rest/reservations resource now correctly returns trigger IDs (if any) on reservations. (WS-2342)
- **HOST_NAME_SUFFIX was no longer adding suffix to job names.** This issue is fixed. (TRQ-2956)
- **pbs_mom filling up the logs in a HA environment.** Reduced verbosity in error logging in HA environments. (TRQ-2863)
- **Delays reported with Preemption.** Speed up process data collection to help diminish time to rerun a job, including correctly sum job memory and cpu time resources for processes created under setsid() calls within a job, and avoid an infinite loop in pbs_mom when running under a Linux Container (lxc). (TRQ-3026)
- **Jobs were getting stuck in a running state when an asynchronous run failed.** This issue is fixed. (TRQ-3114)
- **Array slot limits were not getting decremented when a job is preempted or rerun.** This issue is fixed. (TRQ-3110)