

TORQUE qsub

qsub is the TORQUE command used to submit jobs into a queue. There are several options available for qsub. The purpose of this handout is to explain the resource_list option and how it is used. Below you will see an excerpt from the qsub man page. Following is a detailed description of how to request resources using the -l option.

NAME

qsub - submit pbs job

SYNOPSIS

qsub [-a date_time] [-A account_string] [-b secs] [-c interval] [-C directive_prefix] [-d path] [-D path] [-e path] [-h] [-I] [-j join] [-k keep] [-l resource_list] [-m mail_options] [-M user_list] [-N name] [-o path] [-p priority] [-q destination] [-rc] [-S path_list] [-t num_jobs] [-T prologue/epilogue script_name] [-u user_list] [-v variable_list] [-V] [-W additional_attributes] [-X] [-z] [script]

DESCRIPTION

To create a job is to submit an executable script to a batch server. The batch server will be the default server unless the -q option is specified. See discussion of PBS_DEFAULT under Environment Variables below. Typically, the script is a shell script which will be executed by a command shell such as sh or csh.

Options on the qsub command allow the specification of attributes which affect the behavior of the job.

The qsub command will pass certain environment variables in the Variable_List attribute of the job. These variables will be available to the job. The value for the following variables will be taken from the environment of the qsub command: HOME, LANG, LOGNAME, PATH, MAIL, SHELL, and TZ. These values will be assigned to a new name which is the current name prefixed with the string "PBS_O_". For example, the job will have access to an environment variable named PBS_O_HOME which have the value of the variable HOME in the qsub command environment.

In addition to the above, the following environment variables will be available to the batch job.

PBS_O_HOST

the name of the host upon which the qsub command is running.

PBS_SERVER

the hostname of the pbs_server which qsub submits the job to.

PBS_O_QUEUE

the name of the original queue to which the job was submitted.

PBS_O_WORKDIR

the absolute path of the current working directory of the qsub command.

PBS_ARRAYID

each member of a job array is assigned a unique identifier (see -t)

PBS_ENVIRONMENT

set to **PBS_BATCH** to indicate the job is a batch job, or to **PBS_INTERACTIVE** to indicate the job is a PBS interactive job, see **-I** option.

PBS_JOBID

the job identifier assigned to the job by the batch system.

PBS_JOBNAME

the job name supplied by the user.

PBS_NODEFILE

the name of the file contain the list of nodes assigned to the job (for parallel and cluster systems).

PBS_QUEUE

the name of the queue from which the job is executed.

-l resource_list

Defines the resources that are required by the job and establishes a limit to the amount of resource that can be consumed. If not set for a generally available resource, such as CPU time, the limit is infinite. The *resource_list* argument is of the form:

resource_name[=[value]][,*resource_name*[=[value]],...]

Requesting Resources

The following can be found at <http://www.clusterresources.com/torquedocs21/2.1jobsubmission.shtml>

The table below itemizes the resources that can be requested using the **-l** option with **qsub**. The items under the “Resource” column equate to the *resource_name* portion of the syntax for the **-l** option.

For example:

```
qsub -l nodes=4:ppn=2,host=kmn,mem=200mb jobscript.sh
```

The **-l** option for this command requests Resources of nodes, host and mem.

Various resources can be requested at the time of job submission. A job can request a particular node, a particular node attribute, or even a number of nodes with particular attributes. Either native TORQUE resources, or external scheduler resource extensions may be specified. The native TORQUE resources are listed in the following table:

Resource	Format	Description
arch	string	Specifies the administrator defined system architecture required. This defaults to whatever the PBS_MACH string is set to in "local.mk".
cpuh	seconds, or [[HH:]MM:]SS	Maximum amount of CPU time used by all processes in the job.

Resource	Format	Description
file	size *	The amount of total disk requested for the job. (Ignored on Unicos.)
host	string	Name of the host on which the job should be run. This resource is provided for use by the site's scheduling policy. The allowable values and effect on job placement is site dependent.
mem	size *	Maximum amount of physical memory used by the job. (Ignored on Darwin, Digital Unix, Free BSD, HPUX 11, IRIX, NetBSD, and SunOS. Also ignored on Linux if number of nodes is not 1. Not implemented on AIX and HPUX 10.)
nice	integer	Number between -20 (highest priority) and 19 (lowest priority). Adjust the process execution priority.
nodes	{<node_count> <hostname>} [:ppn=<ppn>][:<property>[:<property>]...] [+ ...]	<p>Number and/or type of nodes to be reserved for exclusive use by the job. The value is one or more node_specs joined with the + (plus) character: <code>node_spec [+node_spec . . .]</code>. Each node_spec is a number of nodes required of the type declared in the node_spec and a name of one or more properties desired for the nodes. The number, the name, and each property in the node_spec are separated by a : (colon). If no number is specified, one (1) is assumed.</p> <p>The name of a node is its hostname. The properties of nodes are:</p> <ul style="list-style-type: none"> • ppn=# - specify the number of processors per node requested. Defaults to 1. • property - a string assigned by the system administrator specifying a node's features. Check with your administrator as to the node names and properties available to you. <p>See Example 1 (-l nodes) for examples.</p> <p>NOTE: By default, the node resource is mapped to a virtual node (that is, directly to a processor, not a full physical compute node). This behavior can be changed within Maui or Moab by setting the <code>JOBNODEMATCHPOLICY</code> parameter. (See Appendix F of the Moab Workload Manager Administrator's Guide for more information.)</p>
opsys	string	Specifies the administrator defined operating system as defined in the mom configuration file.
other	string	Allows a user to specify site specific information. This resource is provided for use by the site's scheduling policy. The allowable values and effect on job placement is site dependent.
pcput	seconds, or [[HH:]MM:]SS	Maximum amount of CPU time used by any single process in the job.
pmem	size *	Maximum amount of physical memory used by any single process of the job. (Ignored on Fujitsu. Not implemented on Digital Unix and HPUX.)

Resource	Format	Description
pvmem	size *	Maximum amount of virtual memory used by any single process in the job. (Ignored on Unicos.)
software	string	Allows a user to specify software required by the job. This is useful if certain software packages are only available on certain systems in the site. This resource is provided for use by the site's scheduling policy. The allowable values and effect on job placement is site dependent. (See Scheduler License Management in the Moab Workload Manager Administrator's Guide for more information.)
vmem	size *	Maximum amount of virtual memory used by all concurrent processes in the job. (Ignored on Unicos.)
walltime	seconds, or [[HH:]MM:]SS	Maximum amount of real time during which the job can be in the running state.

***size format:**

The size format specifies the maximum amount in terms of bytes or words. It is expressed in the form *integer[suffix]*. The suffix is a multiplier defined in the following table ('b' means bytes (the default) and 'w' means words). The size of a word is calculated on the execution server as its word size.

Suffix		Multiplier
b	w	1
kb	kw	1024
mb	mw	1,048,576
gb	gw	1,073,741,824
tb	tw	1,099,511,627,776

Example 1 (-l nodes)

Usage	Description
qsub -l > qsub -l nodes=12	request 12 nodes of any type
qsub -l > qsub -l nodes=2:server+14	request 2 "server" nodes and 14 other nodes (a total of 16) - this specifies two node_specs, "2:server" and "14"
qsub -l > qsub -l nodes=server:hippi+10:noserver+3:bigmem:hippi	request (a) 1 node that is a "server" and has a "hippi" interface, (b) 10 nodes that are not servers, and (c) 3 nodes that have a large amount of memory and have hippy

Usage	Description
<pre>qsub -l > qsub -l nodes=b2005+b1803+b1813</pre>	request 3 specific nodes by hostname
<pre>qsub -l > qsub -l nodes=4:ppn=2</pre>	request 2 processors on each of four nodes
<pre>qsub -l > qsub -l nodes=1:ppn=4</pre>	request 4 processors on one node
<pre>qsub -l > qsub -l nodes=2:blue:ppn=2+red:ppn=3+b1014</pre>	request 2 processors on each of two blue nodes, three processors on one red node, and the compute node "b1014"

Example 2

```
> qsub -l mem=200mb /home/user/script.sh
```

This job requests a node with 200 MB of available memory.

Example 3

```
> qsub -l nodes=node01,mem=200mb /home/user/script.sh
```

This job will wait until node01 is free with 200 MB of available memory.

Example 4

```
> qsub -l nodes=node01,mem=200mb -q regQueue /home/user/script.sh
```

The purpose of this example is to show that the `-l` option is a single option with one or more elements in the list. Other `qsub` options such as `-q` which designates from which queue a job will run can still be used to define the job.