|  |  |  |
| --- | --- | --- |
| **Category** | **State** | **SGE Letter Code** |
| **Pending:** | **pending** | **qw** |
| **Pending:** | **pending, user hold** | **qw** |
| **Pending:** | **pending, system hold** | **hqw** |
| **Pending:** | **pending, user and system hold** | **hqw** |
| **Pending:** | **pending, user hold, re-queue** | **hRwq** |
| **Pending:** | **pending, system hold, re-queue** | **hRwq** |
| **Pending:** | **pending, user and system hold, re-queue** | **hRwq** |
| **Pending:** | **pending, user hold** | **qw** |
| **Pending:** | **pending, user hold** | **qw** |
| **Running** | **running** | **r** |
| **Running** | **transferring** | **t** |
| **Running** | **running, re-submit** | **Rr** |
| **Running** | **transferring, re-submit** | **Rt** |
| **Suspended** | **job suspended** | **s, ts** |
| **Suspended** | **queue suspended** | **S, tS** |
| **Suspended** | **queue suspended by alarm** | **T, tT** |
| **Suspended** | **all suspended with re-submit** | **Rs, Rts, RS, RtS, RT, RtT** |
| **Error** | **all pending states with error** | **Eqw, Ehqw, EhRqw** |
| **Deleted** | **all running and suspended states with deletion** | **dr, dt, dRr, dRt, ds, dS, dT, dRs, dRS, dRT** |

qalter -p 255 jobid 提高优先级

Cluster Grid Engine HOWTO: Job management

Show the list of execution hosts and the list of queues

qconf -sel

qconf -sql

Show the configuration of a queue

qconf -sq all.q

Show the states of a queue

?a (alarm) - At least one of the load thresholds defined in the load\_thresholds list of the queue configuration is currently exceeded. This state prevents N1GE from scheduling further jobs to that queue. For more information, see the queue\_confman page. ?au state might be an indicator for a shutdown host

?A (Alarm) - At least one of the suspend thresholds of the queue is currently exceeded. This state causes jobs running in that queue to be successively suspended until no threshold is violated. For more information, see the queue\_conf man page.

?c (configuration ambiguous) - The queue instance configuration specified using sge\_conf is ambiguous. The state resolves when the configuration becomes unambiguous again. This state prevents you from scheduling further jobs to that queue instance. You can find detailed reasons why a queue instance entered this state in the sge\_qmaster messages file. You can also see the reasons using the qstat command with -explain. For queue instances in this state, the cluster queue's default settings are used for the ambiguous attribute.

?C (Calendar suspended) - The queue has been disabled or suspended automatically using the N1GE calendar facility. See the calendar\_conf man page for more information.

?d (disabled) - This setting is assigned to queues and released using the qmod command. Suspending a queue will suspend all jobs executing in that queue.

?D (Disabled) - The queue has been disabled or suspended automatically using the N1GE calendar facility. See the calendar\_conf man page for more information.

?E (Error) - This setting appears when the N1GE daemon (sge\_execd) on that host was unable to locate the sge\_shepherd executable on that host in order to start a job. Check that daemon's error log for information how to resolve the problem. Enable the queue afterwards using the qmod command with the -c option (e.g., qmod -c all.q). To get explanations of E state, use qstat -explain E -qs E

?o (orphaned) - The current cluster queue's configuration and host group configuration no longer needs this queue instance. The queue instance is kept because unfinished jobs are still associated with it. The orphaned state prevents you from scheduling further jobs to that queue instance. It disappears from qstat output when these jobs finish. To help resolve an orphaned queue instance associated with a job, use the qdel command. You can revive an orphaned queue instance by changing the cluster queue configuration so that the configuration covers that queue instance.

?s (suspended) - Assigned to queues and released using the qmod command. Suspending a queue suspends all jobs executing in that queue.

?S (Subordinate) - The queue has been suspend due to subordination to another queue. See queue\_conf for details. When suspending a queue, regardless of the cause, all jobs executing in that queue are suspended too.

?u (unknown) - The corresponding sge\_execd(8) cannot be contacted.

Increase job priority

As root (or Grid Engine administrator), you can increase/decrease the priority of a job:

# increase priority to bypass jobs earlier in queue

qalter -p 255 jobid

Suspend and resubmit stalled jobs

# as user:qstat | grep neteler | tr -s ' ' ' ' | cut -d' ' -f2 > /tmp/to\_suspend.sgecat /tmp/to\_suspend.sge# as root (?):su -for i in `cat /tmp/to\_suspend.sge` ; do qmod -sj $i ; doneqstat# remove crashed blade from list of execution hosts:qconf -de blade14# delete host from list:qconf -mhgrp "@allhosts"# apply new list:qconf -shgrp "@allhosts"# verify queue stats:qstat -f# resubmit jobs to other nodes (as job user!!):exitfor i in `cat /tmp/to\_suspend.sge` ; do qresub $i ; doneqstat

Delete all your jobs

Ops, you discovered that you launched tons of jobs with a bug in your processing script? You can easily delete all your jobs (your login name is stored in the system variable $USER):

qdel -u $USER

Reactive broken blade

todo

[root@head ~]# qconf -mhgrp "@allhosts"root@head modified "@allhosts" in host group list[root@head ~]# qconf -shgrp "@allhosts"group\_name @allhosts hostlist c00 c01 c02[root@head ~]# ssh c01 "/etc/init.d/sgeexecd stop ; /etc/init.d/sgeexecd start"

Taking a node temporarily out

Sometimes a blade struggles, so it needs to be blacklisted.

Delete the execution host "foo":

qconf -de foo

Add host foo back to the execution list.

qconf -ae foo

Qstat errors: states Eqw and E

State Eqw: If your Grid Engine job is hanging with an Eqw state, try running:

qstat -j This will give you more than enough information to work with and usually the root cause of your problem - usually path errors.

State E: Some job is hanging. Get info with

qstat -explain E -qs E 查看错误节点。

Note: State "E" does not go away automatically To globally clear all E states in your SGE cluster WHEN YOU ARE SURE that this is ok:

qmod -c '\*' See also here. 清除楚错误

Multi-queue management: Suspend and resume queues

Suspend a queue (add -f in case sge\_execd is not reachable):

qmod -s q\_name1

Suspend two queues (add -f in case sge\_execd is not reachable):

qmod -s q\_name1, q\_name2

Resumes (unsuspend) a queue:

qmod -us -f q\_name1

Disable/Enable a particular queue for some reason

... for example for maintenance... Disable a particular queue: qconf -sql # add -f in case sge\_execd is not reachable qmod -d q\_name

To enable back the queue:

qmod -e q\_name

Wildcards can be used to specify a range of queues:

qmod -e q\_name\*

Accounting

You can obtain usage statistics with "qacct":

qacct -o

Statistics are generated by all jobs that have completed and that are contained in the in the cluster accounting file $SGE\_ROOT/$SGE\_CELL/common/accounting. In this case, qacct reports on the following:

?WALLCLOCK – Wall clock time, which is the time between when the job starts and when it finishes in seconds

?UTIME – CPU time spent in user processes in seconds

?STIME – CPU time spent in system calls in seconds

?CPU – CPU time usage in seconds

?MEMORY – The integral memory usage in Gbytes CPU seconds

?IO – The amount of data transferred in input/output operations

?IOW – The input/output wait time in seconds

Now logout of the compute node and try qstat again – your hung MPICH2 job should be gone.

Note that the dr state will also appear on deleted jobs that were running on a dead/crashed compute node. These nodes will show a state of au in the qstat -f listing and you will not be able to ssh or rsh to them. As of April of 2013, users should be able to force the deletion of such jobs:

[user@cluster ~]$ qdel -f 140693 140757 140645 140649

job 140693 is already in deletion

job 140757 is already in deletion

job 140645 is already in deletion

job 140649 is already in deletion

............

[user@cluster ~]$

The -f flag tells the qdel command to **f**orce the deletion of the job, even if the host(s) on which is was running cannot be contacted to confirm removal. The command will take upwards of 30 seconds to complete, with the sequence of dots being displayed to indicate the passage of time.