Univa Corporation

Grid Engine Documentation

Grid Engine Release Notes

Author:
Univa Engineering

Version:
8.3.1

August 28, 2015

Contents

1 License

TERM SOFTWARE LICENSE AND SUPPORT AGREEMENT

This agreement is between the individual or entity agreeing to this agreement and Univa Corporation, a Delaware corporation (Univa) with its registered office at 2300 N Barrington Road, Suite 400, Hoffman Estates, IL 60195.

1. SCOPE: This agreement governs the licensing of the Univa Software and Support provided to Customer.

   - Univa Software is defined as the Univa software described in the order, all updates and enhancements provided under Support, its software documentation, and license keys (Univa Software), which are licensed under this agreement. This Univa Software is only licensed and is not sold to Company.
   - Third-Party Software/Open Source Software licensing terms are addressed on the bottom of this agreement.

2. LICENSE. Subject to the other terms of this agreement, Univa grants Customer, under an order, a non-exclusive, non-transferable, renewable term license up to the license capacity purchased to:

   (a) Operate the Univa Software in Customer's business operations and

   (b) Make a reasonable number of copies of the Univa Software for archival and backup purposes.

   Customer's contractors and majority owned affiliates are allowed to use and access the Univa Software under the terms of this agreement. Customer is responsible for their compliance under the terms of this agreement.

   The initial term of this license is for a period of one year from date hereof to be automatically renewed at each anniversary unless a written notification of termination has been received 60 days prior to each anniversary.

3. RESTRICTIONS. Univa reserves all rights not expressly granted. Customer is prohibited from:

   (a) assigning, sublicensing, or renting the Univa Software or using it as any type of software service provider or outsourcing environment or

   (b) causing or permitting the reverse engineering (except to the extent expressly permitted by applicable law despite this limitation), decompiling, disassembly, modification, translation, attempting to discover the source code of the Univa Software or to create derivative works from the Univa Software.

4. PROPRIETARY RIGHTS AND CONFIDENTIALITY.

   (a) Proprietary Rights. The Univa Software, workflow processes, designs, know-how and other technologies provided by Univa as part of the Univa Software are the proprietary property of Univa and its licensors, and all rights, title and interest in and to such items, including all associated intellectual property rights, remain only with Univa.

The Univa Software is protected by applicable copyright, trade secret, and other intellectual property laws. Customer may not remove any product identification, copyright, trademark or other notice from the Univa Software.

(b) Confidentiality. Recipient may not disclose Confidential Information of Discloser to any third party or use the Confidential Information in violation of this agreement.

(c) Confidential Information means all proprietary or confidential information that is disclosed to the recipient (Recipient) by the discloser (Discloser), and includes, among other things:

- any and all information relating to Univa Software or Support provided by a Discloser, its financial information, software code, flow charts, techniques, specifications, development and marketing plans, strategies, and forecasts
- as to Univa the Univa Software and the terms of this agreement (including without limitation, pricing information).

(ii) Confidential Information excludes information that:

- was rightfully in Recipient's possession without any obligation of confidentiality before receipt from the Discloser
- is or becomes a matter of public knowledge through no fault of Recipient
- is rightfully received by Recipient from a third party without violation of a duty of confidentiality
- is independently developed by or for Recipient without use or access to the Confidential Information or
- is licensed under an open source license.

Customer acknowledges that any misuse or threatened misuse of the Univa Software may cause immediate irreparable harm to Univa for which there is no adequate remedy at law. Univa may seek immediate injunctive relief in such event.

5. PAYMENT. Customer will pay all fees due under an order within 30 days of the invoice date, plus applicable sales, use and other similar taxes.

6. WARRANTY DISCLAIMER. UNIVA DISCLAIMS ALL EXPRESS AND IMPLIED WARRANTIES, INCLUDING WITHOUT LIMITATION THE IMPLIED WARRANTY OF TITLE, MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE UNIVA SOFTWARE MAY NOT BE ERROR FREE, AND USE MAY BE INTERRUPTED.

7. TERMINATION. Either party may terminate this agreement upon a material breach of the other party after a 30 day notice/cure period, if the breach is not cured during such time period. Upon termination of this agreement or expiration of an order, Customer must discontinue using the Univa Software, de-install it and destroy or return the Univa Software and all copies, within 5 days. Upon Univa's request, Customer will provide written certification of such compliance.

8. SUPPORT INCLUDED. Univa's technical support and maintenance services (Support) is included with the fees paid under an order. Univa may change its Support terms, but Support will not materially degrade during any paid term. More details on Support are located at www.univa.com/support

9. LIMITATION OF LIABILITY AND DISCLAIMER OF DAMAGES. There may be situations in which, as a result of material breach or other liability, Customer is entitled to make a claim for damages against Univa. In each situation (regardless of the form of the legal action (e.g. contract or tort claims)), Univa is not responsible beyond:

   (a) the amount of any direct damages up to the amount paid by Customer to Univa in the prior 12 months under this agreement and

   (b) damages for bodily injury (including death), and physical damage to tangible property, to the extent caused by the gross negligence or willful misconduct of Univa employees while at Customer's facility.

   Other than for breach of the Confidentiality section by a party, the infringement indemnity, violation of Univa's intellectual property rights by Customer, or for breach of Section 2 by Customer, in no circumstances is either party responsible for any (even if it knows of the possibility of such damage or loss):

   (a) loss of (including any loss of use), or damage to: data, information or hardware

   (b) loss of profits, business, or goodwill or

   (c) other special, consequential, or indirect damages

10. INTELLECTUAL PROPERTY INDEMNITY. If a third-party claims that Customer's use of the Univa Software under the terms of this agreement infringes that party's patent, copyright or other proprietary right, Univa will defend Customer against that claim at Univa's expense and pay all costs, damages, and attorney's fees, that a court finally awards or that are included in a settlement approved by Univa, provided that Customer:

    (a) promptly notifies Univa in writing of the claim and

    (b) allows Univa to control, and cooperates with Univa in, the defense and any related settlement.

    If such a claim is made, Univa could continue to enable Customer to use the Univa Software or to modify it. If Univa determines that these alternatives are not reasonably available, Univa may terminate the license to the Univa Software and refund any unused fees.

    Univa's obligations above do not apply if the infringement claim is based on the use of the Univa Software in combination with products not supplied or approved by Univa in writing or in the Univa Software, or Customer's failure to use any updates within a reasonable time after such updates are made available.

    This section contains Customer's exclusive remedies and Univa sole liability for infringement claims.

11. GOVERNING LAW AND EXCLUSIVE FORUM. This agreement is governed by the laws of the State of Illinois, without regard to conflict of law principles. Any dispute arising out of or related to this agreement may only be brought in the state of Illinois. Customer consents to the personal jurisdiction of such courts and waives any claim that it is an inconvenient forum. The prevailing party in litigation is entitled to recover its attorney's fees and costs from the other party.

12. MISCELLANEOUS.

(a) Inspection. Univa, or its representative, may audit Customer's usage of the Univa Software at any Customer facility. Customer will cooperate with such audit. Customer agrees to pay within 30 days of written notification any fees applicable to Customer's use of the Univa Software in excess of the license.

(b) Entire Agreement. This agreement, and all orders, constitute the entire agreement between the parties, and supersedes all prior or contemporaneous negotiations, representations or agreements, whether oral or written, related to this subject matter.

(c) Modification Only in Writing. No modification or waiver of any term of this agreement is effective unless signed by both parties.

(d) Non-Assignment. Neither party may assign or transfer this agreement to a third party, except that the agreement and all orders may be assigned upon notice as part of a merger, or sale of all or substantially all of the business or assets, of a party.

(e) Export Compliance. Customer must comply with all applicable export control laws of the United States, foreign jurisdictions and other applicable laws and regulations.

(f) US Government Restricted Rights. The Univa Software is provided with RESTRICTED RIGHTS. Use, duplication, or disclosure by the U.S. government or any agency thereof is subject to restrictions as set forth in subparagraph (c)(I)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 or subparagraphs (c)(1) and (2) of the Commercial Computer Software Restricted Rights at 48 C.F.R. 52.227-19, as applicable.

(g) Independent Contractors. The parties are independent contractors with respect to each other.

(h) Enforceability. If any term of this agreement is invalid or unenforceable, the other terms remain in effect.

(i) No PO Terms. Univa rejects additional or conflicting terms of a Customer's form-purchasing document.

(j) No CISG. The United Nations Convention on Contracts for the International Sale of Goods does not apply.

(k) Survival. All terms that by their nature survive termination or expiration of this agreement, will survive.

Additional software specific licensing terms:

Grid Engine incorporates certain third-party software listed at the URL below. These licenses are accepted by use of the software and may represent license grants with restrictions in which Univa is bound to provide. We are hereby notifying you of these licenses.

Unicloud Kits

- Third Party Software is defined as certain third-party software which is provided along with the Univa Software, and such software is licensed under the license terms located at: http://www.univa.com/resources/licenses/
- Open Source Software is defined as certain opens source software which is provided along with the Univa Software, and such software is licensed under the license terms located at: http://www.univa.com/resources/licenses/

Grid Engine

- Third Party Software is defined as certain third-party software which is provided along with the Univa Software, and such software is licensed under the license terms located at: http://www.univa.com/resources/licenses/
- Open Source Software is defined as certain opens source software which is provided along with the Univa Software, and such software is licensed under the license terms located at: http://www.univa.com/resources/licenses/

Rev: August 2014

2 Supported Operating Systems, Versions and Architectures

Univa Grid Engine supports various platforms, hardware architectures and versions of operating systems. Find the full list in following table:

Operating System	Version	Architecture
SLES	10,11	x86, x86-64
RHEL	5 or higher, 6 or higher, 7	x86, x86-64
CentOS	5 or higher, 6 or higher, 7	x86, x86-64
Oracle Linux	5 or higher, 6 or higher, 7	x86, x86-64
Ubuntu	10.04LTS - 14.04LTS	x86, x86-64
Oracle Solaris	10, 11	x86_64, SPARC 64bit
HP-UX	11.0 or higher	64bit
IBM AIX	6.1 or later	64bit
Apple OS X	10.8 (Mountain Lion) or higher	x86, x86-64
Microsoft Windows	XP Professional (SP3)	32 bit
Microsoft Windows	Server 2003 / 2003 R2	32 bit
Microsoft Windows	Vista Enterprise / Ultimate	32 and 64bit
Microsoft Windows	Server 2008 / 2008 R2	32 and 64bit
Microsoft Windows	7 Professional / Enterprise / Ultimate	32 and 64bit
Microsoft Windows	Server 2012 / 2012 R2	32 and 64bit
Microsoft Windows	8 / 8.1 Pro / Enterprise	32 and 64bit

Table 1: Supported Operating Systems, Versions and Architectures

PLEASE NOTE: Hosts running the Microsoft Windows operations system cannot be used as master or shadow hosts.

PLEASE NOTE: Univa Grid Engine qmaster is fully supported on Linux and Solaris. We provide binaries in Univa Grid Engine for running the qmaster on other operating systems but they are not supported and delivered as a courtesy. If you require qmaster support on other architectures please contact us at support@univa.com.

PLEASE NOTE:: if you require Univa Grid Engine support for older versions of the above operating systems please contact our sales or support team.

3 Fixes and Enhancements

3.1 Summary

3.1.1 Manual Preemption

Univa Grid Engine 8.3 provides the first step of job preemption.

The preemption of jobs in order to enable other jobs to start running can be triggered manually. Later versions of Univa Grid Engine will provide the means to configure and do automatic job preemption.

Find more information concerning preemption in the man page sge_preemption(5) or in the chapter 'Advanced Concepts'->'Manual, Semi-Automatic and Automatic Preemption' in the Users Guide.

3.1.2 Different Resource Requests for Master and Slave Tasks of Parallel Jobs

Univa Grid Engine 8.3 allows to specify different resource requests for the master task and slave tasks of parallel jobs. Resources for the master task can be specified with the **-masterl** switch of submit commands. A parameter named **masterl** is also available in job classes and job submission verifiers to predefine job templates for parallel jobs or to adjust jobs during the time of job submission.

Find more information concerning **masterl** in the man pages qsub.1, sge_job_class.5 or jsv.1.

PLEASE NOTE: The **-masterl** switch affects many different components of Univa Grid Engine for now only the common use cases for **-masterl** have been implemented. Some usage of **-masterl** may conflict with other Univa Grid Engine features.

These are some of the problems that are to be expected when using the **-masterl** switch:

- the **-masterl** switch conflicts with the usage of soft resource requests of the same resource, i.e. if a request like "-masterl resource=val1 -soft -l resource=val2" is specified, the soft request should apply to all slave tasks, but is ignored.
- when enforcing limits by the QMaster is enabled by setting the "qmaster_params ENABLE_ENFORCE_MASTER_LIMIT=true" configuration value, limits specified by the **-masterl** switch are ignored.
- the resources requested using the **-masterl** switch are not taken into account for urgency calculation.
- backfilling of jobs before jobs using the **-masterl** switch might fail.
- requesting different "m_mem_free" limits with **-masterl** und **-l** does not work.

3.1.3 Possibility to Change Resource Selection During Runtime of Jobs

It is now possible to change the amount of granted resources during runtime of a job. Granted resource can either be increased or decreased. To do this **qalter -l** can be used in combination with running jobs. Users have also to specify the **-when** switch to express that the change should happen immediately and not when the job is restarted.

The attempt to increase the granted resource set of a job will be rejected with an error message if not enough resources are available during the time qalter is executed.

This new functionality is especially helpfull to increase h_rt for running jobs or to decrease the amount comsumables (like license requests) when those resources are only required for a certain amount of time but not the full runtime of the job.

Please note that many resource requests such as h_vmem or h_cpu are set by the kernel of the operating system (via the setrlimit() function) so that they cannot be modified during runtime.

The modification of resources that were granted based on "-soft" requested consumables is currently not supported.

Find more information concerning resource adjustements in the -l and -when section of qsub.1.

3.1.4 Possibility to use Consumable Resources for Soft Requests

It is now possible to request a consumable of type INT, MEMORY, RSMAP or DOUBLE in the soft request part (e.g.: qsub -soft -l cpu=4). If there are free resources the job will get it. If there are no resources available the request will be ignored and the job will be started. It is also possible to specify hard and soft requests (e.g.: qsub -soft -l cpu=4 -hard -l cpu=2). The job will only be started if the hard request is available, but if the soft request is also available the soft request will be granted.

3.1.5 Possibility to Specify a Range of Consumable Resource for Soft Requests

It is now also possible to request a range value for consumables in the soft request part (e.g.: qsub -soft -l cpu=4-6). The consumable must have type INT, MEMORY, RSMAP or DOUBLE. The resulting granted request value is reported by qstat -j call. Find more information and examples in the complex (5) man page. The soft consumable range is not supported for the NUMA binding of m_mem_free consumable.

3.1.6 Find Used/Available Resources

Output commands of Univa Grid Engine 8.3 have been enhanced so that they show resources in various states. **qconf -sp pe_name** will now show the amount of used resources. Previous versions showed only the amount of configured resources.

3.1.7 Find free-but-still-bound Resources for Preempted Jobs

Due to the preemption enhancement resources can be in a state called free-but-still-bound. Resources in such state are in process to be released from jobs that got preempted so that they will be available for preemptor jobs of higher priority.

qstat -F, **qhost -F**, **qquota** will show resources in that state additionally to the maximum capacity or available/used resources.

3.1.8 Users Can increase POSIX Priority of Own Jobs

In the past users where only allowed to decrease the POSIX priority of own jobs. Univa Grid Engine 8.3 allows also to reincrease the priority up to the value 0 when an administrator allowed this for a cluster by defining the parameter **ALLOW_INCREASE_POSIX_PRIORITY=1** in the **qmaster_params** of the cluster configuration.

3.1.9 Configureable Timeout for Client Side Suspended qrsh

When qrsh is run from a terminal and placed into the background the qrsh is suspended by the terminal session and will remain suspended indefinitely until it is placed back into the foreground by the user. Qrsh provides a command line switch **-bgio** that controls how qrsh behaves when backgrounded, however many users are not aware of this switch or they forget to properly use the switch and backgrounded qrsh tasks can remain backgrounded forever. A new parameter **SGE_DEFAULT_MAX_IJS_CLIENT_WAIT_TIME controls** how long the qrsh will wait before exiting. Please read sge_conf.5 in the man pages for more details.

3.1.10 Lost job Detection

The qmaster parameters “lost_job_timeout” and “enable_lost_job_reschedule” can be used to enable lost job detection at qmaster. If qmaster worker threads will monitor the jobs reported by the execution daemons. If a task of a job that was started on an execution node is not reported for longer than the defined timeout the job is logged in the qmaster messages file. If “enable_lost_job_reschedule” is enabled the jobs for which the timeout was detected are set to error and will be rescheduled.

3.2 Full List of Fixes and Enhancements

Univa Grid Engine 8.3.0 alpha (also fixed for a 8.1 and/or 8.2 patch release)

```
GE-5020  SGE_HGR_ environment variable is not shown in case of host aliasing
GE-5031  reduce the level of the execd message "master task of pe job*finished
         - still*slave tasks are running" to INFO
GE-5127  drmaa client failed receiving gdi request response for mid=65535
         (got syncron message receive timeout error)
GE-5151  extensive logging in qmaster messages file
GE-5154  qdel may crash and cause communication error loggings at qmaster
GE-5158  massive qdel request stresses qmaster daemon
GE-5159  event client (e.g. scheduler) may get triggered events delayed if event
         interval is changed
GE-5256  qhost, qstat and qquota should disconnect from qmaster before printing
         out the data
```

Univa Grid Engine 8.3.0 alpha (also fixed in a 8.2 patch release)

```
GE-5088  qmaster is crashing when used in combination with LO
GE-3177  max_unheard parameter not working as expected
GE-3435  potential deadlock in commlib
GE-4954  Add configurable timeout for client-side suspended qrsh jobs
GE-5059  update script adding wrong default parameter for cgroups_params
GE-5072  stree-edit is not part of the distribution
GE-5078  RSMAP attribute in complex_values definition masks following attributes
GE-5086  if execd gets modified execd load report time the change is not
         immediately effective
GE-5108  execd installation fails with error message "./inst_sge: test: ] missing"
GE-5112  uninstallation fails with error message
         "./inst_sge: LO_ENABLE_QCONF_OPTIONS=1: is not an identifier"
GE-5120  qmaster is crashing due to lothread issue, when a array job is deleted
GE-5121  scheduler assigns already used resource map value to job
GE-5122  RSMAP not correctly set if job requests more than one value
GE-5235  race condition at connection set-up may overwrite MAX_IJS_CLIENT_WAIT_TIME
GE-5285  qconf -de nodexxx should not work if nodexxx is reference any where.
         It causes the qmaster to crash
GE-5038  qstat -j shows job present, but job not available and no shepherd present
         either
GE-5280  execd reports memory errors at startup
GE-5360  the notify switch does not work in combination with qrsh
GE-2638  advance reservations should support project based access lists
GE-3610  check for GDI-version mismatch at commlib level
GE-4207  qrsh -inherit to a cluster of different version dumps core
GE-4782  the use of binding switch breaks the functionality of -w v/p
GE-4833  gridengine ignores complex request and puts tasks into wrong queue instance
GE-4892  shepherd pid is not moved out of cgroup when shepherd_cmd is set
GE-4959  on native Windows, if the execd was started manually, it stops when the
```

```
          console is closed
  GE-4964 on native Windows, the job environment does not contain SGE_ and -V/-v
          variables
  GE-4973 finished jobs are not stored at all, even if the global config param
          finished_jobs is greater than zero
  GE-4995 Fix scripts/nonreentrant.sh script that fails to find non-thread-safe
          functions used in the source code
  GE-5018 cgroup setting "killing=true" causes shepherd to terminate incorrectly
  GE-5032 jsv jc parameter is not reset in server JSV (bourne shell, TCL) if it
          was set during previous job verification
  GE-5036 native Windows clients crash if the sgepasswd file is corrupted
  GE-5041 "sharelog" record timestamp in "reporting" file not in milliseconds
  GE-5042 man page UGE_Starter_Service.exe(8) contains wrong version
  GE-5046 aix platform needs libxml2.a to be available in LIBPATH
  GE-5047 sge_qmaster segmentation fault
  GE-5051 util/setfilperm.sh does not set ownership of install_execd.bat
  GE-5055 sge_qmaster daemon accepts requests from clients using older GDI version
  GE-5058 make the auto installer create certificates even if WIN_DOMAIN_ACCESS is false
  GE-5065 garbled error output of "save_sge_config.sh"
  GE-5066 GUI installer refers to UGE 8.2.0beta1
  GE-5068 upgrade procedure does not check for existence of "bc" command
  GE-5071 libdrmaa is missing in sol-sparc packages
  GE-5075 define a single point to set the Grid Engine version and GDI version
  GE-5080 invalid "gdi_request_limits" accepted by cluster config change although
          error message is printed
  GE-5091 automatic session cleanup does not work in root user systems
  GE-5092 cwd entry in accounting might break the accounting file format when ":"
          are used in dir or filenames.
  GE-5093 accounting does not filter "\n" in submission command line
  GE-5094 negative performance impact on qmaster due to logging into message file:
          "session <session_id>: processed all available events till unique ID <event_id>"
  GE-5097 new PE parameter daemon_forks_slave / master_forks_slave needs to be
          compatible with cgroups main memory limitation
  GE-5114 host is not set to error state if sgepasswd file cannot be read or is broken
  GE-5115 sge_execd and sge_shepherd depend on libgcc on sol-amd64
  GE-5116 sge_execd on hp11-ia64 does not start (/usr/lib/hpux64/dld.so: Unable to
          find library libxml2.so.11)
  GE-5117 jobs are not started on hp11-ia64 (failed 137 : invalid execution state)
  GE-5132 create dl script for native Windows
  GE-5139 on native Windows, execd crashes if a load sensor reports too much
          load at a time
  GE-5144 aimk -uge-version broken
  GE-5146 port qping to native Windows (win-x86)
  GE-5150 misleading error message for classic spooling qmaster installation
  GE-5215 on native Windows, the PATH environment variable contains UNIX style parts
  GE-5238 on native Windows, it is not possible to specify more than one load sensor
  GE-5243 upgrade script fails to upgrade accounting file to 8.2.x format
```

Univa Grid Engine 8.3.0 beta1

```
GE-921   Jobs blocked by -a <time> shall also be considered by resource reservation
GE-5157  colon in resource request breaks accounting file format
GE-5260  describe how to setup sgepasswd if $SGE_ROOT is mounted with nosetuid
GE-5265  describe how to write and configure native Windows load sensors
GE-5269  using job info as filter for jobs can cause segmentation fault of DRMAA2 app
GE-5270  update all information about UGE on Windows in doc and man pages
GE-5274  possible delayed configuration update for execution hosts
GE-5275  race condition can brake reschedule_unknown functionality
GE-5309  DRMAA2 C API lacks of a function for reaping single jobs from internal lists
GE-3256  qsub -p accepts anything as priority value
GE-3601  add a way that consumables can also be requested as soft-request
GE-4115  enable qalter to modify resource requests of running jobs
GE-4309  Allow for the active -increasing- of a resource request
GE-4936  If licenses is not in use it should be possible to release it before job ends
GE-5195  Visualisation of preempted jobs
GE-5213  users should be able to increase Posix priority up to 0 even when lowered before
GE-5217  qstat -help does not contain -explain m
GE-5227  JB_priority field has wrong value for negative Posix priorities in XML Output
GE-5228  users should be allowed to manually preempt their own jobs with qmod -p
GE-5247  Formatting in sge_conf man page partially incorrect
GE-5272  need command to show used slots of a PE
GE-5273  demo doc package contains user and admin manual
GE-5307  DRMAA2 enums do not contain UNSET values which causes issues for some compilers
GE-5318  DRMAA2 extensible data structures needs to contain additionial pointers
         (OGF issue 160)
GE-5328  qalter prints error and success message when consumables in hard resource lists
         should be changed for pending jobs
GE-5334  qstat -j should show granted resources for running jobs
GE-5343  qstat -xml with and without -j shows different values for JB_priority
GE-5353  qacct gets broken when accounting line is to long
GE-5375  qselect option -si is broken
GE-5376  sgeexecd stop fails on Solaris with running jobs
GE-5166  enhancements to tag resources that get available through preemption
         (preemptable-resources)
GE-5175  Client side of manual preemption
GE-5219  ability to change Posix priority and job share against JC access specifiers
GE-5288  inconsistent/wrong OS support matrices in UGE documentation set
GE-5322  Improve performance of DRMAA2 list implementation
GE-5323  DRMAA2 C job template methods needs to be renamed to jtemplate
```

Univa Grid Engine 8.3.0 beta2

```
GE-3295  Too many users in a group causes GE commands to fail
GE-3778  reservation could not be enabled in certain cases.
GE-5245  Reservation is made way too far in the future.
GE-5324  dbwriter creates additional entry in sge_job for array jobs
GE-5383  limit enforcement in the execd is broken
GE-5396  Due to changes of limit calculation qalter -when now is broken
GE-5405  submitting or modifying parallel jobs with '-mbind nlocal' but
```

```
          without '-masterl m_mem_free' is broken
  GE-5406  scheduling of parallel jobs that request different architectures
          for master and slave tasks is broken
  GE-5410  Reservation in UGE scheduler fails when available resources are reported by a loadsens
  GE-5413  qstat crashes, when /etc/groups entry exceeds a certain length
  GE-5419  manual preemption is not triggered although requested
  GE-5420  UGE scheduler does not reserve resources for the preemptor that
          are freed due to preemption
  GE-5424  Now-reservation causes qstat -j to print incorrect end time of a job
  GE-5425  default preemption action missing when not specified with qmod -p
  GE-5426  multi line preemption_msg is not indented
  GE-5431  Exit status 4294967295 in acct records in reporting file
  GE-5432  UGE jobs might stuck in pending state when it requested reservation
          and when licenses are managed with LO.
  GE-5435  qalter -clears l_hard does not support complex attribute shortcuts
  GE-5447  soft range algorithm might add granted resources to pe slave jobs
          running on hosts where the consumable is not available
  GE-5459  qmaster crashes if -l or -masterl requests of a running job are
          altered
  GE-5461  job limit enforcement uses wrong values if -masterl is specified
  GE-5463  usage value wallclock time is wrongly multiplied by the number
          of slots a parallel job reserves on the execution host
  GE-5466  qconf -Mattr fails although the request is valid
  GE-5467  execd crashes from time to time when executing a new job
```

Univa Grid Engine 8.3.0rc1

```
  GE-5354  shepherd creates panic file because it does not get the current working directory
  GE-5394  qmod -p should print out warnings if preemption is not configured properly
  GE-5397  execd crashes on modification of running job
  GE-5460  Implement a policy to limit permission for altering resource requests.
  GE-5468  'qlogin -N name' using ssh causes shepherd to continously log POLLHUP messages
  GE-5474  if consumables are specified in '-masterl' request, load_thresholds are not obeyed
  GE-5477  qquota prints incorrect usage (slots=11844674407370)
  GE-5485  upgrade procedure fails upgrading exechost objects
  GE-5490  parallel jobs requesting a consumable resource are not scheduled when capacity / reque
  GE-5494  add a means to indent debug trace output depending on the call tree level
  GE-5499  client job spools unexpectedly removed while execd_params KEEP_ACTIVE=true is set
  GE-5500  sgeexecd startup script may hang on shutdown (stop call)
  GE-5512  sgemaster -shadowd start/stop might not work because of hostname matching problem
```

Univa Grid Engine 8.3.0rc2

```
  GE-5453  add section to users guide that explains how to use preemption
  GE-5456  multiple intermediate accounting records are written between 00:00 and 00:10 per job
  GE-5506  on native Windows, qacct crashes immediately
  GE-5507  on native Windows, qquota dumps an assertion if a job is running a an RQS is defined
  GE-5514  setting gdi_request_limits to NONE is ignored by master when some limits where previou
  GE-5516  SSOS functionality broken
```

Univa Grid Engine 8.3.1beta1

GE-3989 add execution host and start time to qstat -j output of jobs
GE-4218 scheduler profiling can not be switched off, it just is kind of disabled, logging does not stop
GE-4488 interactive job with ssh setup (not the builtin method) does not handle SIGHUP signal
GE-4745 Extra ^H characters in trace file causes problems with email from epilog scripts
GE-4757 Specifying -t 2-2 with -sync leads to a hang in qsub
GE-4861 qsh and old interactive job support change owner of /dev/null on execution host
GE-4862 job doing a self reschedule using qmod -rj <own_job_id> are not rescheduled
GE-5060 print out the IP address for not resolvable hosts
GE-5107 resource quotas are running out of sync, limits are overcommited
GE-5131 all array job tasks get same tickets in functional, sharetree (and override) policy
GE-5294 enhanced observation of processes and their used group id in the PDC module
GE-5414 Reconnection of execd to qmaster takes for ever.
GE-5438 qrsh does not show the LD_LIBRARY_PATH removal message
GE-5484 GE-5465 deny the combination of m_mem_free soft request and -mbind
GE-5517 UGE 8.3.0 cannot submit jobs to LO prior to 1.1
GE-5518 decreasing of m_mem_free in combination with mbind is broken
GE-5520 Support for Linux 4.x Kernels
GE-5522 execd updates wallclock usage of running jobs far too often
GE-5523 jobs with h_rt=INFINITY get killed immediately after start because they exceeded h_rt
GE-5525 building fails on Windows (win-x86) if a directory defined in aimk.site.bat is not located on the current drive
GE-5526 array tasks are not sorted by priority in qstat output
GE-5527 bdb spooling error: error deleting record with key CONFIG:<HOSTNAME>
GE-5532 Queue instances lost after qmaster restart
GE-5534 qalter h_rt can not be altered in combination with exclusive flag
GE-5538 qstat shows pending job priority only after 2 scheduling runs
GE-5544 DRMAA2 C lib crashed when processing jobs in hold state
GE-5549 maxpss and maxrss values falling down to current pss and rss value
GE-5551 do basic integration of docker into UGE using load sensor and starter_method

Univa Grid Engine 8.3.1

GE-5591 qsh jobs hang at start on OpenIndiana
GE-5572 qmaster crashes when qalter -when now is used

4 Upgrade Notes

4.1 Upgrade Requirements

This is a summary of the Upgrade Matrix that describes how you can carry out the transition from Sun or Oracle Grid Engine 6.2uX, Univa Grid Engine 8.0.X, 8.1.X, 8.2.X to Univa Grid Engine 8.3 when you are currently using classic, BDB local spooling or PostgreSQL spooling. If the current version of Grid Engine you are using is missing in the overview, then please look at the full Upgrade Matrix located in the section Updating Univa Grid Engine in the Installation Guide.

Version	Upgrade Method
Univa Grid Engine 8.2.X	Backup/Restore
Univa Grid Engine 8.1.X	Backup/Restore
Univa Grid Engine 8.0.X	Backup/Restore
Oracle Grid Engine 6.2u6-6.2u8	Backup/Restore
Sun Grid Engine 6.2u5	Backup/Restore
Sun Grid Engine 6.2u1-6.2u4	Upgrade to SGE 6.2u5 and then Backup/Restore
Sun Grid Engine 6.2 FCS	Upgrade to SGE 6.2u5 and then Backup/Restore

Table 2: Upgrading from SGE, OGE, UGE 8.1.X, UGE 8.2.X to Univa Grid Engine 8.3.X

4.2 Upgrade Procedure

When you upgrade from a previous version of Univa Grid Engine to Univa Grid Engine 8.3.0 then the upgrade procedure will automatically adjust the existing complex configuration of the cluster. An additional attribute will be added to each complex. The attribute is named *aapre* (available after preemption) and it specifies that a resource bound by a job will be automatically reported as available within the Univa Grid Engine system when the corresponding job gets preempted.

As default after an upgrade the *aapre* attribute will be set to *YES* for all consumable resources that exist in the cluster. If you do not want such resources to be reported as available when a job is preempted then you can change this value to *NO* after the upgrade process.

Please note that the *aapre* value for the slots complex has always to be set to *YES* whereas all non-consumables can only have the value *NO*.

Find more information concerning complexes in the man page complex(5) or in the Univa Grid Engine administration guide.

5 Compatibility Notes

5.1 Reporting of Job Wallclock Time

For every job, wallclock time is reported as online usage and in accounting and reporting. Wallclock time is the time a job is *running*, time periods during which a job was suspended are excluded from wallclock time reporting.

This obsoletes the following attributes of the **execd_params** in the cluster configuration (see man page sge_conf.5):

- *ACCT_RESERVED_USAGE*
- *SHARETREE_RESERVED_USAGE*
- *ENABLE_REAL_CP*
- *ENABLE_REAL_MEM*

The **usage_weight_list** configured in the scheduler configuration (see man page sched_conf.5) has been extended by the attribute **wallclock**. The setting of **SHARETREE_RESERVED_USAGE** is replaced by adding **wallclock=1** to the **usage_weight_list**.

PLEASE NOTE: In Univa Grid Engine 8.3.0 and Univa Grid Engine 8.3.1, the calculation of the **wallclock** time usage is broken for parallel jobs running in a parallel environment that use the **accounting_summary** feature!

5.2 Reporting of Resource Capacities and Resource Quota

With the addition of manual job preemption an additional state of resources has been added: When a job gets preempted the resources which will get preempted will first enter a state “free but still bound”. They will stay in the state “free but still bound” until the job is actually preempted (sge_execd acknowledged receipt of the corresponding suspend signal).

The output of commands printing resource capacities show such “free but still bound” resources:

5.2.1 qstat -F / qhost -F

Example: A sequential job requests 100M of a memory resource having a capacity of 1G and consumes 1 of 10 job slots.

qstat -F shows:

```
        gc:mem_pre=924.000M
        qc:slots=9
3000000003 0.55476 sleep      joga         r     04/02/2015 11:07:53     1
```

When the job gets preempted, the preemptable resources will show up as free but still bound for a short time:

```
        gc:mem_pre=924.000M+100.000M
        qc:slots=9+1

3000000003 0.55476 sleep      joga         P     04/02/2015 11:07:53     1
```

Once the preemption action has been finished, the resources will be reported as free:

```
        gc:mem_pre=1.000G
        qc:slots=10

3000000003 0.55476 sleep      joga         P     04/02/2015 11:07:53     1
```

5.2.2 qquota

Example: We have a resource quota set on slots:

```
{
   name         slots
   description  NONE
   enabled      TRUE
   limit        to slots=5
}
```

When a sequential job is running it consumes 1 slot, qquota output shows:

```
resource quota rule limit                filter
--------------------------------------------------------------------------------
slots/1            slots=1/5            -
```

When the job gets preempted, qquota shows that one slot less is booked from the resource quota (but not yet fully free):

```
resource quota rule limit                filter
--------------------------------------------------------------------------------
slots/1            slots=1-1/5          -
```

5.2.3 qconf -sp

qconf -sp shows two additional lines, “used slots” and “bound slots”.

Example: A parallel environment “test.pe” provides 5 slots, qconf -sp test.pe shows:

```
pe_name             test.pe
slots               5
used_slots          0
bound_slots         0
...
```

A 4 times parallel job changes the output to:

```
...
used_slots          4
bound_slots         0
...
```

When the job gets preempted the 4 slots consumed by the job will show up as bound_slots for a short time period:

```
...
used_slots          4
bound_slots         4
...
```

Once the preemption action has finished no slots of the PE will be reported as in use:

```
...
used_slots          0
bound_slots         0
...
```

6 Known Issues and Limitations

Following list shows the known limitations that are part of Univa Grid Engine 8.3.1:

GE-5597 with accounting_summary=true, "wallclock" usage of PE jobs is wrong

This affects not only the accounting of the "wallclock" usage, but also the calculation of the tickets if a share tree is used that is based on the "wallclock" usage.

Following list shows the known limitations of the "-masterl" switch in Univa Grid Engine 8.3.1:

GE-5380 If a resource is requested both in "-soft -l resource=value1" and "-masterl resource=value2", the soft request is ignored.

GE-5381 qmaster doesn't enforce -masterl limits

If enforcing limits by the QMaster is enabled using the "qmaster_params ENABLE_ENFORCE_MASTER_LIMIT=true" configuration value, the QMaster still does not enforce limits defined by the "-masterl" switch.

GE-5382 backfilling doesn't work properly for jobs that use the -masterl switch

GE-5384 Resource Quota Sets (RQS) do not take the limits requested by the "-masterl" switch into account

GE-5385 advance reservation does not take -masterl requests into account

GE-5348 The "-masterl" switch cannot be used in combination with the Univa License Orchestrator

GE-5386 qmon does not support the -masterl switch