



Sun N1 Grid Engine 6.1 Release Notes



Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Part No: 820-0700
May 2007

Copyright 2007 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. All rights reserved.

Sun Microsystems, Inc. has intellectual property rights relating to technology embodied in the product that is described in this document. In particular, and without limitation, these intellectual property rights may include one or more U.S. patents or pending patent applications in the U.S. and in other countries.

U.S. Government Rights – Commercial software. Government users are subject to the Sun Microsystems, Inc. standard license agreement and applicable provisions of the FAR and its supplements.

This distribution may include materials developed by third parties.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, the Solaris logo, the Java Coffee Cup logo, docs.sun.com, N1 Java, and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and SunTM Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

Products covered by and information contained in this publication are controlled by U.S. Export Control laws and may be subject to the export or import laws in other countries. Nuclear, missile, chemical or biological weapons or nuclear maritime end uses or end users, whether direct or indirect, are strictly prohibited. Export or reexport to countries subject to U.S. embargo or to entities identified on U.S. export exclusion lists, including, but not limited to, the denied persons and specially designated nationals lists is strictly prohibited.

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 2007 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. Tous droits réservés.

Sun Microsystems, Inc. détient les droits de propriété intellectuelle relatifs à la technologie incorporée dans le produit qui est décrit dans ce document. En particulier, et ce sans limitation, ces droits de propriété intellectuelle peuvent inclure un ou plusieurs brevets américains ou des applications de brevet en attente aux Etats-Unis et dans d'autres pays.

Cette distribution peut comprendre des composants développés par des tierces personnes.

Certains composants de ce produit peuvent être dérivées du logiciel Berkeley BSD, licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays; elle est licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, le logo Solaris, le logo Java Coffee Cup, docs.sun.com, N1 Java et Solaris sont des marques de fabrique ou des marques déposées de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui, en outre, se conforment aux licences écrites de Sun.

Les produits qui font l'objet de cette publication et les informations qu'il contient sont régis par la législation américaine en matière de contrôle des exportations et peuvent être soumis au droit d'autres pays dans le domaine des exportations et importations. Les utilisations finales, ou utilisateurs finaux, pour des armes nucléaires, des missiles, des armes chimiques ou biologiques ou pour le nucléaire maritime, directement ou indirectement, sont strictement interdites. Les exportations ou réexportations vers des pays sous embargo des Etats-Unis, ou vers des entités figurant sur les listes d'exclusion d'exportation américaines, y compris, mais de manière non exclusive, la liste de personnes qui font objet d'un ordre de ne pas participer, d'une façon directe ou indirecte, aux exportations des produits ou des services qui sont régis par la législation américaine en matière de contrôle des exportations et la liste de ressortissants spécifiquement désignés, sont rigoureusement interdites.

LA DOCUMENTATION EST FOURNIE "EN L'ETAT" ET TOUTES AUTRES CONDITIONS, DECLARATIONS ET GARANTIES EXPRESSES OU TACITES SONT FORMELLEMENT EXCLUES, DANS LA MESURE AUTORISEE PAR LA LOI APPLICABLE, Y COMPRIS NOTAMMENT TOUTE GARANTIE IMPLICITE RELATIVE A LA QUALITE MARCHANDE, A L'APTITUDE A UNE UTILISATION PARTICULIERE OU A L'ABSENCE DE CONTREFACON.

Contents

- 1 Sun N1 Grid Engine 6.1 Software Release Notes5**
 - Accessing Documentation5
 - Free 30–Day Email Support5
 - Contents of This Software Package6
 - Installing the Sun N1 Grid Engine 6.1 Software7
 - Supported Operating Systems and Platforms7
 - Using N1 Grid Engine 6.1 with an Existing 6.0 Cluster8
 - New Features in Sun N1 Grid Engine 6.1 Software8
 - Flexible Resource Quotas8
 - Master Bottleneck Analysis Using Solaris 10 DTrace8
 - New Command Options8
 - Support for Additional Operating Systems9
 - Support for Additional Database Software9
 - Other Changes9
 - Changed Features in N1 Grid Engine 6.1 Software9
 - Changed Command Options9
 - Software Support Changes in Sun N1 Grid Engine 6.1 Software9
 - Known Limitations and Workarounds 10
 - Known Limitations of Sun N1 Grid Engine 6.1 Software 10
 - Known Limitations and Workarounds for the Microsoft Windows Platform 12

Sun N1 Grid Engine 6.1 Software Release Notes

The release notes include the following information:

- “Accessing Documentation” on page 5
- “Free 30–Day Email Support” on page 5
- “Contents of This Software Package” on page 6
- “Installing the Sun N1 Grid Engine 6.1 Software” on page 7
- “New Features in Sun N1 Grid Engine 6.1 Software” on page 8
- “Changed Features in N1 Grid Engine 6.1 Software” on page 9
- “Known Limitations and Workarounds” on page 10

Accessing Documentation

You can view or print the most recent Sun N1 Grid Engine 6.1 documentation from the Sun documentation site at <http://docs.sun.com/app/docs/coll/1017.4>. The documentation includes the following:

- *Sun N1 Grid Engine 6.1 Installation Guide*
- *Sun N1 Grid Engine 6.1 Administration Guide*
- *Sun N1 Grid Engine 6.1 User’s Guide*
- *Sun N1 Grid Engine 6.1 Release Notes*

Free 30–Day Email Support

N1 Grid Engine 6.1 is available for free download from the www.sun.com web site. To receive 30–days of free email support for your download, fill in and send the [free evaluation questionnaire](http://www.javelinfeedback.com/sun/index.jsp?pi=c2b00c871c1f86177ac800c779c76fab) (<http://www.javelinfeedback.com/sun/index.jsp?pi=c2b00c871c1f86177ac800c779c76fab>).

Contents of This Software Package

The Sun N1 Grid Engine 6.1 software distribution is made up of the following components:

- The grid engine software binary packages, including all daemons, client programs, and libraries. You must load and install one binary package for each operating system architecture you intend to use.
- The grid engine software common package, containing install scripts, and other architecture-independent utilities.
- The optional Accounting and Reporting Console (ARCo) software, which is made up of three separate packages:
 - The Sun Java Web Console package. You must select the package appropriate for the operating system architecture on which you plan to run the web console server.

Note – You can also download the Sun Java Web Console 2.2.6 software from the Sun web site at <http://www.sun.com/download/products.xml?id=461d58be>.

- The dbwriter package, written in Java and therefore available in only one version.
- The ARCo module package, usable across different supported architectures.

Note – In order to operate ARCo, you also must set up a PostgreSQL, MySQL, or Oracle database server. PostgreSQL, MySQL, and Oracle are not included in the Sun N1 Grid Engine 6.1 software distribution. For more information, see Chapter 8, “Installing the Accounting and Reporting Console,” in *Sun N1 Grid Engine 6.1 Installation Guide*.

The Sun N1 Grid Engine 6.1 software distribution kit contains the following top-level directory hierarchy:

- 3rd_party – Contains information about freeware, public domain, and public license software
- bin – Grid engine software executables
- catman – Online manual pages organized into admin and user commands
- ckpt – Sample checkpointing configurations
- dbwriter – DbWriter software used by the accounting and reporting console
- dtrace – DTrace based monitoring utilities for Solaris 10
- examples – Sample script files, configuration files, and application programs
- include – DRMAA header file
- lib – Required shared libraries and DRMAA Java™ binding jar file
- man – Online manual pages in nroff format

- `mpi` – A sample parallel environment interface for the MPI message-passing system
- `pvm` – A sample parallel environment interface for the PVM message-passing system
- `qmon` – Pixmaps, resource, and help files for QMON, the graphical user interface
- `reporting` – Accounting and reporting console software
- `util` – Some utility shell procedures used for installation tasks and some template grid engine system shutdown and boot scripts
- `utilbin` – Some utility programs that are mainly required during the installation

Installing the Sun N1 Grid Engine 6.1 Software

To install the Sun N1 Grid Engine 6.1 software, follow the instructions in *Sun N1 Grid Engine 6.1 Installation Guide*.

Supported Operating Systems and Platforms

The Sun N1 Grid Engine 6.1 software supports the following operating systems and platforms:

- Solaris 10, 9 and 8 Operating Systems (SPARC Platform Edition)
- Solaris 10 and 9 Operating Systems (x86 Platform Edition)
- Solaris 10 Operating System (x64 Platform Edition)
- Apple Mac OS X 10.4 (Tiger), PPC platform
- Apple Mac OS X 10.4 (Tiger), x86 platform
- Hewlett Packard HP-UX 11.00 or higher, 32 bit
- Hewlett Packard HP-UX 11.00 or higher, 64 bit (including HP-UX on IA64)
- IBM AIX 5.1, 5.3
- Linux x86, kernel 2.4, 2.6, glibc \geq 2.3.2
- Linux x64, kernel 2.4, 2.6, glibc \geq 2.3.2
- Linux IA64, kernel 2.4, 2.6, glibc \geq 2.3.2
- Silicon Graphics IRIX 6.5
- Microsoft Windows Server 2003, Windows XP Professional with Service Pack 1 or later, Windows 2000 Server with Service Pack 3 or later, or Windows 2000 Professional with Service Pack 3 or later

Using N1 Grid Engine 6.1 with an Existing 6.0 Cluster

You can install the N1 Grid Engine 6.1 software in an environment that has an existing N1 Grid Engine 6.0 cluster. To run the 6.1 software in parallel with an existing N1 Grid Engine environment, follow these rules:

- Use a different `$SGE_ROOT` directory and different TCP ports for the `qmaster` and execution daemons.
- Do *not* select to install a system-wide startup script during manual or automatic installation. Installing a system-wide startup script would overwrite your N1 Grid Engine 6.0 startup script for `qmaster` and execution daemons.
- If you decide to install two execution daemons on one host, be sure to use a different “gid_range” from the global/local cluster configuration.
- On Microsoft Windows systems, you can install the optional “N1 Grid Engine Helper Service” only for *one* Grid Engine instance. If you already had installed this service for N1 Grid Engine 6.0, you may not install it for N1 Grid Engine 6.1 and, thus, you cannot run jobs that require a GUI on the Windows desktop for N1 Grid Engine 6.1.
- Verify that variables point to the correct instance of N1 Grid Engine. Specifically, check your port settings, your `PATH` variable, and the `LD_LIBRARY_PATH` variable. For Solaris and Linux, `LD_LIBRARY_PATH` does not need to be set anymore.

New Features in Sun N1 Grid Engine 6.1 Software

The Sun N1 Grid Engine 6.1 software includes several new features and expanded functionality.

Flexible Resource Quotas

The *resource quotas* feature enables you to limit the maximum number of running jobs per user, user group, and projects on arbitrary resources like queues, hosts, memory, and software licenses. A firewall-like rule syntax allows an unprecedented configuration flexibility.

For information about resource quotas, see Chapter 6, “Managing Resource Quotas,” in *Sun N1 Grid Engine 6.1 Administration Guide*. For additional details, see the `qquota(1)`, `sge_resource_quota(5)`, and `qconf(1)` man pages.

Master Bottleneck Analysis Using Solaris 10 DTrace

If your master component runs on a Solaris 10 machine, you can use the DTrace-based master monitor diagnosis utility to monitor the master and look for any bottlenecks. For more information, see “Using DTrace for Performance Tuning” in *Sun N1 Grid Engine 6.1 Administration Guide* and the `$SGE_ROOT/dtrace/README_dtrace.txt` file.

New Command Options

You can now use the `-wd` option to specify the job working directory for any of the following commands: `qsub`, `qalter`, `qsh`, `qrsh` and `qmon`. For more information, see the man pages.

Support for Additional Operating Systems

The Sun N1 Grid Engine 6.1 release adds support for the following operating systems:

- Linux on Itanium (IA64)
- Apple Mac OS X on the x86 platform

Support for Additional Database Software

ARCo supports the following database servers: PostgreSQL 7.4 - 8.2, MySQL 5.0, and Oracle 9i, 10.0, 10.1, and 10.2.

Other Changes

- Resource matching for string and host complex attributes has been extended to support a flexible boolean expression grammar (logical AND, OR and NOT operators).
- The Grid Engine Accounting and Reporting Console (ARCo) now can write the reporting data to the MySQL database.
- You no longer need to set the environment variable LD_LIBRARY_PATH on Solaris and Linux when using N1 Grid Engine commands. This change improves command execution and helps to avoid conflicts with system installed shared libraries, such as SSL and Berkeley DB libraries.
- The complex variable `display_win_gui` now enables you to schedule jobs only to Windows hosts that are running the “N1 Grid Engine Helper Service.” The helper service allows background applications to display their graphical user interfaces on the visible desktop of the Windows host.
- Minor changes to QMON to improve usability.

Changed Features in N1 Grid Engine 6.1 Software

Changed Command Options

For performance reasons, the default behavior of the `qstat -u` option has changed. Before N1 Grid Engine 6.1, `qstat` without the `-u` option printed the jobs of all users. Beginning with N1 Grid Engine 6.1, `qstat` without the `-u` prints only the jobs of the user who executed `qstat`.

To enforce the old `qstat` behavior, administrators can add `-u *` to the cluster-wide `$SGE_ROOT/$SGE_CELL/common/sge_qstat` file. Users can enforce the previous behavior by adding `-u *` to the user private file searched at `$HOME/.sge_qstat`.

Software Support Changes in Sun N1 Grid Engine 6.1 Software

The Sun N1 Grid Engine 6.1 software no longer supports the following operating systems:

- Solaris 7 (SPARC Platform Edition)
- Solaris 8 (x86 Platform Edition)

- IBM AIX 4.3
- Apple MacOS X 10.2 (Jaguar) and 10.3 (Panther) on PowerPC (PPC) Platform

In addition, the Sun N1 Grid Engine 6.1 software does not support the Grid Engine Management Module (GEMM) for Sun Control Station.

Known Limitations and Workarounds

The following sections contain information about product irregularities discovered during testing, but too late to fix or document.

Known Limitations of Sun N1 Grid Engine 6.1 Software

This Sun N1 Grid Engine 6.1 software release has the following limitations:

- The stack size for `sgc_qmaster` should be set to 16 MBytes. `sgc_qmaster` might not run with the default values for stack size on the following architectures: IBM AIX and HP UX 11.
- You should set a high file descriptor limit in the kernel configuration on hosts that are designated to run the `sgc_qmaster` daemon. You might want to set a high file descriptor limit on the shadow master hosts as well. A large number of available file descriptors enables the communication system to keep connections open instead of having to constantly close and reopen them. If you have many execution hosts, a high file descriptor limit significantly improves performance. Set the file descriptor limit to a number that is higher than the number of intended execution hosts. You should also make room for concurrent client requests, in particular for jobs submitted with `qsub -sync` or when you are running DRMAA sessions that maintain a steady communication connection with the master daemon. Refer to your operating system documentation for information about how to set the file descriptor limit.
- The number of concurrent dynamic event clients is limited by the number of file descriptors. The default is 99. Dynamic event clients are jobs submitted with the `qsub -sync` command and a DRMAA session. You can limit the number of dynamic event clients with the `qmaster_params` global cluster configuration setting. Set this parameter to `MAX_DYN_EC=n`. See the `sgc_conf(5)` man page for more information.
- The ARCo module is available only for the Solaris Sparc, Solaris Sparc 64 bit, Solaris x86, Solaris x64, Linux x86, and Linux 64 bit kernels.
- Only a limited set of predefined queries is currently shipped with ARCo. Later releases will include more comprehensive sets of predefined queries.
- Jobs requesting the amount `INFINITY` for resources are not handled correctly with respect to resource reservation. `INFINITY` might be requested by default in case no explicit request for a certain resource has been made. Therefore it is important to request that all resources be explicitly taken into account for resource reservation.

- Resource reservation currently takes only pending jobs into account. Consequently, jobs that are in a hold state due to the submit options `-a time` and `-hold_jid joblist`, and are thus not pending, do not get reservations. Such jobs are treated as if the `-R n` submit option were specified for them.
- Berkeley DB requires that the database files reside on the local disk, if `qmaster` is not running on Solaris 10 and uses a NFSv4 mount (full NFSv4 compliant clients and servers from other vendors are also supported, but have not yet been tested.) If the `sge_qmaster` cannot be run on the file server intended to store the spooling data (for example, if you want to use the shadow master facility), a Berkeley DB RPC server can be used. The RPC server runs on the file server and connects with the Berkeley DB `sge_qmaster` instance. However, Berkeley DB's RPC server uses an insecure protocol for this communication and so it presents a security problem. Do *not* use the RPC server method if you are concerned about security at your site. Use `sge_qmaster` local disks for spooling instead and, for fail-over, use a high availability solution such as Sun Cluster, which maintains host local file access in the fail-over case.
- Busy QMON with large array task numbers. If large array task numbers are used, you should use “compact job array display” in the QMON Job Control dialog box customization. Otherwise the QMON GUI will cause high CPU load and show poor performance.
- The automatic installation option does not provide full diagnostic information in case of installation failures. If the installation process aborts, check for the presence and the contents of an installation log file in `qmaster-spool-dir/install_hostname_timestamp.log` or in `/tmp/install.pid`.
- On IBM AIX, HP/UX 11, and SGI IRIX 6.5 systems, two different binaries are provided for `sge_qmaster`, `spooldefaults`, and `spoolinit`. One of these binaries is for the Berkeley DB spooling method, the other binary is for the classic spooling method. The names of these binaries are `binary.spool_db` and `binary.spool_classic`.

To change to the desired spooling method, modify three symbolic links before you install the master host. Do the following:

```
# cd sge-root/bin/arch
# rm sge_qmaster
# ln -s sge_qmaster.spool_classic sge_qmaster

# cd sge-root/utilbin/arch
# rm spooldefaults spoolinit
# ln -s spooldefaults.spool_classic spooldefaults
# ln -s spoolinit.spool_classic spoolinit
```

- The default Mac OS X installation does not include the OpenMotif library that QMON needs. You can get the OpenMotif library for the PowerPC and x86 architectures from various web sites, such as <http://dryden.biol.yorku.ca/macosex/>.
- PDF export in ARCo requires a lot of memory. Huge reports can result in a `OutOfMemoryException` when they are exported into PDF.

Workaround – Increase the JVM heap size for the Sun Java Web Console. The following command sets the max. heap size to 512 MB.

```
# smreg add -p java.options="... -Xmx512M ..."
```

A restart of the Sun Java Web Console is necessary to make the change effective as in this command:

```
# smcwebserver restart
```

- For DBWriter (part of ARCo) the 64-Bit support of the Java virtual machine needs to be installed on Solaris Sparc 64-bit and Solaris x64, and Linux 64-bit kernels.
- When you use Java bindings with DRMAA, verify that the LD_LIBRARY_PATH is set correctly.

Note – If you are using a 32-bit Java Virtual Machine (JVM), you must set the LD_LIBRARY_PATH to the 32-bit shared DRMAA library (for example, \$SGE_ROOT/lib/sol-sparc), even when your application actually runs on a 64-bit operating system platform.

- The N1 Grid Engine 6.1 version of the drmaa.jar file is not compatible with the previous drmaa.jar file. The old drmaa.jar file has been renamed to drmaa-0.5.jar.
- For a fully-featured automatic installation (not using CSP), you must grant the root user permissions to remote login through rsh or ssh without asking for a password. This enables the installation script to start the installation on the remote hosts. If this is not configured correctly, you have to log into each execution host and manually execute the automatic installation using the following command:

```
inst_sge -x -auto <conf-file> -noremove
```

- If you use local execd spool directories, the automatic installation will have problems. Local execd spool directories are defined on local hard disks. Using these local execd spool directories may lead to a hanging automatic installation.

Known Limitations and Workarounds for the Microsoft Windows Platform

- The installation of Services For UNIX (SFU) 3.5 requires a good administrative understanding of the Windows platform and its integration into a UNIX environment. For an overview of SFU, see Appendix A, “Microsoft Services For UNIX,” in *Sun N1 Grid Engine 6.1 Installation Guide*. You can find additional technical information and documentation about SFU on the Microsoft web site at <http://www.microsoft.com/windows/sfu/default.asp>.

Username mapping, NFS mounts, and hostname resolving in SFU require special attention to successfully install the Grid Engine execution daemon, submit host functionality, and integration of Windows hosts into a N1 Grid Engine cluster.

- You cannot install a Windows execution host remotely with the auto installation procedure. You can use the auto installation procedure through the `inst_sge -noremote` command to install locally.
- You cannot submit a job from a Windows submit host as the Windows “local Administrator” to a Unix or Linux execution host. However, you can submit a job as local Administrator from Windows to Windows, and you can submit as user root from Unix or Linux to Windows, Unix, or Linux execution hosts.

