



Sun Cluster Data Service for Sun Grid Engine Guide for Solaris OS



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

Part No: 820-3042-10
January 2009, Revision A

Copyright 2009 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, Java, and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc. or its subsidiaries 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 Sun 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 2009 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, Java et Solaris sont des marques de fabrique ou des marques déposées de Sun Microsystems, Inc., ou ses filiales, 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

Preface	5
Installing and Configuring Sun Cluster HA for Sun Grid Engine	9
Sun Cluster HA for Sun Grid Engine Overview	10
Overview of Installing and Configuring Sun Cluster HA for Sun Grid Engine	10
Planning the Sun Cluster HA for Sun Grid Engine Installation and Configuration	11
Configuration Restrictions	12
Configuration Requirements	12
Configuration Considerations	14
Configuration Planning Questions	15
Preparing the Nodes and Disks	16
▼ How to Prepare the Nodes and Disks	17
Installing and Configuring Sun Grid Engine	18
▼ How to Install and Configure Sun Grid Engine	19
▼ How to Enable Sun Grid Engine to Run in a Cluster	20
Verifying the Installation and Configuration of Sun Grid Engine	21
▼ How to Verify the Installation and Configuration of Sun Grid Engine	21
Installing the Sun Cluster HA for Sun Grid Engine Packages	22
▼ How to Install the Sun Cluster HA for Sun Grid Engine Packages	23
Configuring the HASStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine	24
▼ How to Register and Configure an HASStoragePlus Resource	25
Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine	25
▼ How to Configure Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine	26
Registering and Configuring Sun Cluster HA for Sun Grid Engine	27
Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources	27

- ▼ How to Create and Enable Sun Cluster HA for Sun Grid Engine Resources 29
 - Setting Sun Cluster HA for Sun Grid Engine Extension Properties 30
- Verifying the Sun Cluster HA for Sun Grid Engine Installation and Configuration 31
 - ▼ How to Verify the Sun Cluster HA for Sun Grid Engine Installation and Configuration ... 31
- Tuning the Sun Cluster HA for Sun Grid Engine Fault Monitors 31
- Debugging Sun Cluster HA for Sun Grid Engine 32
 - ▼ How to Activate Debugging for Sun Cluster HA for Sun Grid Engine 32

- A Files for Configuring and Removing Sun Cluster HA for Sun Grid Engine Resources 35**
 - Listing of `sge_config` 35
 - Listing of `sge_register` 36
 - Listing of `sge_remove` 38

- Index 39**

Preface

Sun Cluster Data Service for Sun Grid Engine Guide for Solaris OS explains how to install and configure Sun™ Cluster HA for Sun Grid Engine.

Note – This Sun Cluster release supports systems that use the SPARC and x86 families of processor architectures: UltraSPARC, SPARC64, AMD64, and Intel 64. In this document, x86 refers to the larger family of 64-bit x86 compatible products. Information in this document pertains to all platforms unless otherwise specified.

This document is intended for system administrators with extensive knowledge of Sun software and hardware. Do not use this document as a planning or presales guide. Before reading this document, you should have already determined your system requirements and purchased the appropriate equipment and software.

The instructions in this book assume knowledge of the Solaris™ Operating System (Solaris OS) and expertise with the volume-manager software that is used with Sun Cluster software.

Using UNIX Commands

This document contains information about commands that are specific to installing and configuring Sun Cluster data services. The document does *not* contain comprehensive information about basic UNIX® commands and procedures, such as shutting down the system, booting the system, and configuring devices. Information about basic UNIX commands and procedures is available from the following sources:

- Online documentation for the Solaris Operating System
- Solaris Operating System man pages
- Other software documentation that you received with your system

Typographic Conventions

The following table describes the typographic conventions that are used in this book.

TABLE P-1 Typographic Conventions

Typeface	Meaning	Example
AaBbCc123	The names of commands, files, and directories, and onscreen computer output	Edit your <code>.login</code> file. Use <code>ls -a</code> to list all files. <code>machine_name% you have mail.</code>
AaBbCc123	What you type, contrasted with onscreen computer output	<code>machine_name% su</code> Password:
<i>aabbcc123</i>	Placeholder: replace with a real name or value	The command to remove a file is <i>rm filename</i> .
<i>AaBbCc123</i>	Book titles, new terms, and terms to be emphasized	Read Chapter 6 in the <i>User's Guide</i> . <i>A cache</i> is a copy that is stored locally. Do <i>not</i> save the file. Note: Some emphasized items appear bold online.

Shell Prompts in Command Examples

The following table shows the default UNIX system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.

TABLE P-2 Shell Prompts

Shell	Prompt
C shell	<code>machine_name%</code>
C shell for superuser	<code>machine_name#</code>
Bourne shell and Korn shell	<code>\$</code>
Bourne shell and Korn shell for superuser	<code>#</code>

Related Documentation

Information about related Sun Cluster topics is available in the documentation that is listed in the following table. All Sun Cluster documentation is available at <http://docs.sun.com>.

Topic	Documentation
Data service administration	<i>Sun Cluster Data Services Planning and Administration Guide for Solaris OS</i> Individual data service guides
Concepts	<i>Sun Cluster Concepts Guide for Solaris OS</i>
Overview	<i>Sun Cluster Overview for Solaris OS</i>
Software installation	<i>Sun Cluster Software Installation Guide for Solaris OS</i>
System administration	<i>Sun Cluster System Administration Guide for Solaris OS</i>
Hardware administration	<i>Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS</i> Individual hardware administration guides
Data service development	<i>Sun Cluster Data Services Developer's Guide for Solaris OS</i>
Error messages	<i>Sun Cluster Error Messages Guide for Solaris OS</i>
Command and function reference	<i>Sun Cluster Reference Manual for Solaris OS</i>

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster at <http://docs.sun.com>.

Related Third-Party Web Site References

Third-party URLs that are referenced in this document provide additional related information.

Note – Sun is not responsible for the availability of third-party web sites mentioned in this document. Sun does not endorse and is not responsible or liable for any content, advertising, products, or other materials that are available on or through such sites or resources. Sun will not be responsible or liable for any actual or alleged damage or loss caused or alleged to be caused by or in connection with use of or reliance on any such content, goods, or services that are available on or through such sites or resources.

Documentation, Support, and Training

The Sun web site provides information about the following additional resources:

- [Documentation](http://www.sun.com/documentation/) (<http://www.sun.com/documentation/>)
- [Support](http://www.sun.com/support/) (<http://www.sun.com/support/>)
- [Training](http://www.sun.com/training/) (<http://www.sun.com/training/>)

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. To share your comments, go to <http://docs.sun.com> and click Feedback.

Getting Help

If you have problems installing or using Sun Cluster, contact your service provider and provide the following information:

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the Solaris Operating System (for example, Solaris 10)
- The release number of Sun Cluster (for example, Sun Cluster 3.2)

Use the following commands to gather information about each node on your system for your service provider.

Command	Function
<code>prtconf -v</code>	Displays the size of the system memory and reports information about peripheral devices
<code>psrinfo -v</code>	Displays information about processors
<code>showrev -p</code>	Reports which patches are installed
<code>prtdiag -v</code>	Displays system diagnostic information
<code>/usr/cluster/bin/clnode show-rev</code>	Displays Sun Cluster release and package version information

Also have available the contents of the `/var/adm/messages` file.

Installing and Configuring Sun Cluster HA for Sun Grid Engine

This chapter explains how to install and configure Sun Cluster HA for Sun Grid Engine.

Note – Sun Grid Engine was formerly known as Sun ONE Grid Engine. In this book, references to Sun Grid Engine also apply to Sun ONE Grid Engine unless this book explicitly states otherwise.

This chapter contains the following sections.

- “Sun Cluster HA for Sun Grid Engine Overview” on page 10
- “Overview of Installing and Configuring Sun Cluster HA for Sun Grid Engine” on page 10
- “Planning the Sun Cluster HA for Sun Grid Engine Installation and Configuration” on page 11
- “Preparing the Nodes and Disks” on page 16
- “Installing and Configuring Sun Grid Engine” on page 18
- “Verifying the Installation and Configuration of Sun Grid Engine” on page 21
- “Installing the Sun Cluster HA for Sun Grid Engine Packages” on page 22
- “Configuring the HAStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine” on page 24
- “Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine” on page 25
- “Registering and Configuring Sun Cluster HA for Sun Grid Engine” on page 27
- “Verifying the Sun Cluster HA for Sun Grid Engine Installation and Configuration” on page 31
- “Tuning the Sun Cluster HA for Sun Grid Engine Fault Monitors” on page 31
- “Debugging Sun Cluster HA for Sun Grid Engine” on page 32

Sun Cluster HA for Sun Grid Engine Overview

Sun Grid Engine is a distributed resource management program, which runs jobs in parallel on multiple machines. To minimize the loss of work that a failure of a machine might cause, nodes in the management tier must be protected against failure. However, protection of individual execution nodes in the grid against failure is not required. Failure of an individual execution node in a grid causes only a minor loss of work.

To eliminate single points of failure in the management tier of a Sun Grid Engine system, Sun Cluster HA for Sun Grid Engine provides fault monitoring and automatic fault recovery for the following Sun Grid Engine daemons:

- Queue master daemon
- Scheduling daemon

You must configure Sun Cluster HA for Sun Grid Engine as a failover service.

For conceptual information about failover data services and scalable data services, see [Sun Cluster Concepts Guide for Solaris OS](#).

Because the management tier relies on the Sun Grid Engine file system, the NFS server that exports this file system must also be protected against failure. To eliminate single points of failure in the NFS server, use the Sun Cluster HA for NFS data service. For more information about this data service, see [Sun Cluster Data Service for NFS Guide for Solaris OS](#).

Each component of Sun Grid Engine has a data service that protects the component when the component is configured in Sun Cluster. See the following table.

TABLE 1 Protection of Sun Grid Engine Components by Sun Cluster Data Services

Sun Grid EngineComponent	Data Service
Sun Grid Engine daemons:	Sun Cluster HA for Sun Grid Engine
▪ Queue master daemon (<code>sge_qmaster</code>)	The resource type is <code>SUNW.gds</code> .
▪ Scheduling daemon (<code>sge_schedd</code>)	
NFS server	Sun Cluster HA for NFS
	The resource type is <code>SUNW.nfs</code> .

Overview of Installing and Configuring Sun Cluster HA for Sun Grid Engine

The following table summarizes the tasks for installing and configuring Sun Cluster HA for Sun Grid Engine and provides cross-references to detailed instructions for performing these tasks. Perform the tasks in the order that they are listed in the table.

TABLE 2 Tasks for Installing and Configuring Sun Cluster HA for Sun Grid Engine

Task	Instructions
Plan the installation	“Sun Cluster HA for Sun Grid Engine Overview” on page 10 “Planning the Sun Cluster HA for Sun Grid Engine Installation and Configuration” on page 11
Prepare the nodes and disks	“Preparing the Nodes and Disks” on page 16
Install and configure Sun Grid Engine	“Installing and Configuring Sun Grid Engine” on page 18
Verify Sun Cluster HA for Sun Grid Engine installation and configuration	“Verifying the Installation and Configuration of Sun Grid Engine” on page 21
Install Sun Cluster HA for Sun Grid Engine Packages	“Installing the Sun Cluster HA for Sun Grid Engine Packages” on page 22
Configure the HASToragePlus resource type to work with Sun Cluster HA for Sun Grid Engine	“Configuring the HASToragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine” on page 24
Configure Sun Cluster HA for NFS for use with Sun Cluster HA for Sun Grid Engine	“Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine” on page 25
Register and Configure Sun Cluster HA for Sun Grid Engine	“Registering and Configuring Sun Cluster HA for Sun Grid Engine” on page 27
Verify Sun Cluster HA for Sun Grid Engine installation and configuration	“Verifying the Sun Cluster HA for Sun Grid Engine Installation and Configuration” on page 31
Tune Sun Cluster HA for Sun Grid Engine fault monitors	“Tuning the Sun Cluster HA for Sun Grid Engine Fault Monitors” on page 31
Debug Sun Cluster HA for Sun Grid Engine	“Debugging Sun Cluster HA for Sun Grid Engine” on page 32

Planning the Sun Cluster HA for Sun Grid Engine Installation and Configuration

This section contains the information that you need to plan your Sun Cluster HA for Sun Grid Engine installation and configuration.

Note – Before you begin, consult your Sun Grid Engine documentation for configuration restrictions and requirements that are not imposed by Sun Cluster software.

Configuration Restrictions

The configuration restrictions in the subsections that follow apply only to Sun Cluster HA for Sun Grid Engine.



Caution – Your data service configuration might not be supported if you do not observe these restrictions.

Sun Grid Engine Shadow Daemon

Do *not* use the Sun Grid Engine shadow daemon. The Sun Grid Engine shadow daemon provides an optional mechanism for recovery from failures. This mechanism interferes with the automatic fault recovery that Sun Cluster provides.

Sun Grid Engine Berkley DB spooling server

Do *not* choose the option to use a Berkley DB spooling server. Either choose the Classic spooling method or the local Berkley DB spooling method. Currently it is not possible to configure the Berkley DB spooling server in a highly available way within the Sun Cluster framework.

Start at Boot Option

Do *not* choose the start at boot option when installing Sun Grid Engine. To ensure that Sun Cluster HA for Sun Grid Engine can provide fault monitoring and automatic fault recovery, Sun Grid Engine must be started *only* by Sun Cluster.

Configuration Requirements

The configuration requirements in this section apply only to Sun Cluster HA for Sun Grid Engine.



Caution – If your data service configuration does not conform to these requirements, the data service configuration might not be supported.

Sun Grid Engine Software Version Requirements

Use Sun Grid Engine version 6.0 or 6.1. Make sure to apply the most recent available Patches to the Sun Grid Engine software.

Operating System for the Sun Grid Engine Management Tier

The Sun Grid Engine management tier must run on Sun Cluster nodes. Because Sun Cluster runs only on the Solaris Operating System, the Sun Grid Engine management tier must also run on the Solaris Operating System. However, Sun Grid Engine supports other operating systems. Therefore, this requirement applies only to the management tier, *not* to individual execution nodes in the grid.

Memory Requirements

Ensure that enough free memory is available on the cluster nodes where you plan to run the Sun Grid Engine master.

The amount of free memory that is required on each cluster node depends on the number of jobs that are running on the grid. For example:

- If 100 jobs are running, 10 Mbytes of free memory are required.
- If 10,000 jobs are running, 1 Gbyte of free memory is required.

Disk Space Requirements

Ensure that you have enough disk space in the Sun Grid Engine file system and on the local disk of each node.

The disk space requirements for each type of file or directory in the Sun Grid Engine file system are listed in the following table.

File Type or Directory Type	Required Disk Space
Binary files	15 Mbytes for each architecture
Spool directories	30–200 Mbytes
Installation tar file	40 Mbytes

On the local disk of each node, 10–20 Mbytes of disk space are required. If you are installing the Sun Grid Engine software on the local disk of a node, 15 Mbytes of disk space are additionally required for the binary files.

Sun Cluster HA for Sun Grid Engine Configuration Requirements

Configure Sun Cluster HA for Sun Grid Engine as a failover data service. You cannot configure Sun Cluster HA for Sun Grid Engine as a scalable data service. For more information, see:

- [“How to Enable Sun Grid Engine to Run in a Cluster” on page 20](#)
- [“Registering and Configuring Sun Cluster HA for Sun Grid Engine” on page 27](#)

Note – If you are using the Solaris 10 OS, install and configure this data service to run only in the global zone. At publication of this document, this data service is not supported in non-global zones. For updated information about supported configurations of this data service, contact your Sun service representative.

NFS Configuration for the Sun Grid Engine File System

The Sun Grid Engine file system must reside on a multihost disk. This disk must be available to the other nodes in the cluster that will be used for the Sun Grid Engine administrative services,

You must use NFS to export the Sun Grid Engine file system to the noncluster nodes. The NFS server that exports this file system must also be protected against failure. To protect the NFS server against failure, use the Sun Cluster HA for NFS data service. For more information about this data service, see *Sun Cluster Data Service for NFS Guide for Solaris OS*.

Sun Cluster HA for NFS Configuration Requirements

Configure the resources for the Sun Grid Engine management tier in the same resource group as the resource for NFS. For more information, see “[Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine](#)” on page 25.

Dependencies Between Sun Grid Engine Components

The dependencies between Sun Grid Engine components are shown in the following table.

TABLE 3 Dependencies Between Sun Grid Engine Components

Sun Grid Engine Component	Dependency
Sun Grid Engine queue master daemon (sge_qmaster)	SUNW.HAStoragePlus resource
Sun Grid Enginescheduling daemon (sge_schedd)	Sun Grid Engine queue master daemon (sge_qmaster) resource

These dependencies are set when you register and configure Sun Cluster HA for Sun Grid Engine. For more information, see “[Registering and Configuring Sun Cluster HA for Sun Grid Engine](#)” on page 27.

Configuration Considerations

The configuration considerations in the subsections that follow affect the installation and configuration of Sun Cluster HA for Sun Grid Engine.

Location of the Sun Grid Engine Binary Files

You can install Sun Grid Engine on one of the following locations:

- A highly available local file system
- The cluster file system

For the advantages and disadvantages of placing the Sun Grid Engine binary files on a highly available local file system and the cluster file system, see “[Configuration Guidelines for Sun Cluster Data Services](#)” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Tip – To enable the type of file system to be identified from the mount point, use a prefix that indicates the type of file system as follows:

- For mount points on a highly available local file system, use the `/local` prefix.
 - For mount points on the cluster file system, use the `/global` prefix.
-

File Systems for Spool Directories and Binary Files

The optimum distribution of spool directories and binary files among file systems depends on the grid configuration. See the following table.

Grid Configuration	File System Configuration
The execution tier contains fewer than 200 hosts.	Use a single shared NFS file system under the root of the Sun Grid Engine file system for the spool directories and binary files.
The execution tier contains about 200 hosts, or the applications are disk intensive.	Use a separate area on an NFS file system for the spool directories.
The execution tier contains more than 200 hosts, or NFS performance is likely to be a problem.	See the Sun Grid Engine documentation for alternate grid configurations.

Configuration Planning Questions

Use the questions in this section to plan the installation and configuration of Sun Cluster HA for Sun Grid Engine. Write the answers to these questions in the space that is provided on the data service worksheets in “[Configuration Worksheets](#)” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

- Which resource group will you use for the following resources:
 - Logical host name resource
 - HASStoragePlus resource
 - NFS resource

- Sun Grid Engine application resources

Use the answer to this question when you perform the following procedures:

- “How to Enable Sun Grid Engine to Run in a Cluster” on page 20
- “Configuring the HAStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine” on page 24
- “Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine” on page 25
- “Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources” on page 27

- What is the logical host name for the Sun Grid Engine resource? Clients access the data service through this logical host name.

Use the answer to this question when you perform the procedure “How to Enable Sun Grid Engine to Run in a Cluster” on page 20.

- Which resources will you use for the components of Sun Grid Engine?

You require one resource for each component in the following list:

- Queue master daemon
- Scheduling daemon

Use the answer to this question when you perform the procedure “Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources” on page 27.

- Where will the system configuration files reside?

See “Configuration Guidelines for Sun Cluster Data Services” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for the advantages and disadvantages of using the local file system instead of the cluster file system.

Preparing the Nodes and Disks

Preparing the nodes and disks modifies the configuration of the operating system to enable Sun Cluster HA for Sun Grid Engine to eliminate single points of failure in a Sun Grid Engine system.

Before you begin, ensure that the requirements in the following sections are met:

- “Memory Requirements” on page 13
- “Disk Space Requirements” on page 13

▼ How to Prepare the Nodes and Disks

1 Become superuser on all the cluster nodes where you are installing Sun Grid Engine.

2 Create an administrative user account for Sun Grid Engine on all those cluster nodes.

Either select an existing user account other than root for the grid administration, or create an account specifically for grid administration.

Tip – For consistency with the Sun Grid Engine documentation, name the account `sgeadmin`.

3 Create a directory for the root of Sun Grid Engine file system.

```
# mkdir sge-root-dir
```

Note – The `sge-root-dir` must reside in the cluster filesystem. Refer to “[Configuring the HAStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine](#)” on page 24 for more details.

4 Change the owner of the root of the Sun Grid Engine file system to the administrative user whose account you created in [Step 2](#).

```
# chown sge-admin sge-root-dir
```

5 Set the mode of the root of Sun Grid Engine file system to `drwxr-xr-x`.

```
# chmod 755 sge-root-dir
```

6 Specify the port number and protocol for the `sge_qmaster` and `sge_execd` services.

Choose an unused port number below 1024. The `sge_qmaster` and `sge_execd` services are to be provided through Transmission Control Protocol (TCP).

To specify the port number and protocol, add the following line to the `/etc/services` file.

```
sge_qmaster    port-no/tcp
sge_execd     port-no/tcp
```

7 For each type of host in the grid, create a plain text file that contains the names of all hosts of that type in the grid.

The `install_qmaster` script uses these files when you install Sun Grid Engine. Create a separate file for each type of host in the grid:

- Execution hosts
- Administrative hosts
- Submit hosts

Example 1 Preparing the Nodes and Disks for the Installation of Sun Grid Engine

This example shows how to prepare the nodes and disks for a Sun Grid Engine installation that is to be configured as follows:

- The root of Sun Grid Engine file system is the `/global/gridmaster` directory. This directory resides in the cluster file system.
- The account for grid administration is named `sgeadmin`.
- The `sge_qmaster` service is to be provided through port 536 and TCP.
- The `sge_execd` service is to be provided through port 537 and TCP.

The sequence of operations for preparing the nodes and disks for the installation of Sun Grid Engine is as follows:

1. To create the `/global/gridmaster` directory for the root of Sun Grid Engine file system, the following command is run:

```
# mkdir /global/gridmaster
```

2. To change the owner of the `/global/gridmaster` directory to the `sgeadmin` user, the following command is run:

```
# chown sgeadmin /global/gridmaster
```

3. To set the mode of the `/global/gridmaster` directory to `drwxr-xr-x`, the following command is run:

```
# chmod 755 /global/gridmaster
```

4. To specify that the `sge_qmaster` service is to be provided through port 536 and TCP, and that the `sge_execd` service is to be provided through port 537 and TCP, the following line is added to the `/etc/services` file:

```
sge_qmaster    536/tcp
sge_execd     537/tcp
```

Installing and Configuring Sun Grid Engine

The procedure that follows explains only the special requirements for installing Sun Grid Engine for use with Sun Cluster HA for Sun Grid Engine. For complete information about installing and configuring Sun Grid Engine, see your Sun Grid Engine documentation.

To enable Sun Grid Engine to run in a cluster, you must modify Sun Grid Engine to use a logical host name.

▼ How to Install and Configure Sun Grid Engine

Before you begin, ensure that you have the host names of all hosts in the grid. Create a separate list of host names for each type of host in the grid:

- Execution hosts
- Administrative hosts
- Submit hosts

1 Become superuser of the cluster node where you are installing Sun Grid Engine.

2 Install the Sun Grid Engine distribution files. You have to choose between the `tar.gz` format and the `pkgadd` format.

Follow the instructions outlined in “How to Load the Distribution Files On a Workstation” in the *N1 Grid Engine 6 Installation Guide*.

Note – If you choose the `pkgadd` format, you need to make sure to install Patches for the Sun Grid Engine software on exactly the same node the Sun Grid Engine packages are registered on.

3 Set the `SGE_ROOT` environment variable to the directory for the root of Sun Grid Engine file system that you created in “[Preparing the Nodes and Disks](#)” on page 16.

```
# SGE_ROOT=sge-root-dir
# export SGE_ROOT
```

4 Go to the directory for the root of Sun Grid Engine file system.

```
# cd sge-root-dir
```

5 Start the script that installs the Sun Grid Engine master host.

```
# ./install_qmaster
```

6 Follow the prompts on screen to provide or confirm the following information:

- The name of the Sun Grid Engine administrative user
- The value of the `SGE_ROOT` environment variable
- The TCP port number
- The name of the Sun Grid Engine cell to be configured
- The path to the spool directory
- The setup for the correct file permissions
- Details of your domain name service (DNS) domains

7 When you are asked whether you want to use classic spooling or Berkley DB, do not choose to use a Berkely DB spooling Server.

Either choose the `classic` spooling method, or choose `Berkley DB` with `local` spooling.

8 When you are prompted, specify the range of group IDs for Sun Grid Engine to use.

To ensure that you allocate enough group IDs, specify a range of approximately 100 group IDs, for example, 20000-20100.

9 Follow the prompts on screen to provide or confirm the following information:

- The path to the spooling directory for the execution daemon
- The email address of the user who should receive problem reports
- Confirm the configuration parameters

10 When you are asked if you want to install the script that starts Sun Grid Engine at boot time, reply no.

You are asked if you want to install the script that starts Sun Grid Engine at boot time.

```
We can install the startup script that will
start qmaster/scheduler at machine boot (y/n) [y] >> n
```

To ensure that Sun Cluster HA for Sun Grid Engine can provide fault monitoring and automatic fault recovery, Sun Grid Engine must be started *only* by Sun Cluster.

11 Follow the prompts on screen to provide or confirm the following information:

- Specify the list of execution, admin and submit hosts
- Do not use a shadow host
- Select a scheduler profile

▼ How to Enable Sun Grid Engine to Run in a Cluster

1 Become superuser of a node in the cluster that will host Sun Grid Engine.**2 Create a failover resource group to contain the Sun Cluster HA for Sun Grid Engine resources.**

Use the resource group that you identified when you answered the questions in [“Configuration Planning Questions” on page 15](#).

```
# clresourcegroup create -p Pathprefix=sge-root-dir sge-rg
```

```
-p Pathprefix= sge-root-dir
```

Specifies a directory on a cluster file system that Sun Cluster HA for NFS uses to maintain administrative and status information. This directory must be the directory that you created for the root of the Sun Grid Engine file system in [“Preparing the Nodes and Disks” on page 16](#).

```
sge-rg
```

Specifies that the resource group that you are creating is named *sge-rg*.

- 3 **Add a resource for the Sun Grid Engine logical host name to the failover resource group that you created in [Step 2](#).**

```
# clreslogicalhostname create \  
-g sge-rg \  
-h hostlist \  
sge-lh-rs
```

-g *sge-rg*
Specifies that the logical host name resource is to be added to the failover resource group that you created in [Step 2](#)

-h *hostlist*
Specifies a comma-separated list of host names that are to be made available by this logical host name resource

sge-lh-rs
Specifies that the resource that you are creating is named *sge-lh-rs*

Verifying the Installation and Configuration of Sun Grid Engine

Before you install the Sun Cluster HA for Sun Grid Engine packages, verify that the Sun Grid Engine software is correctly installed and configured to run in a cluster. This verification does *not* verify that the Sun Grid Engine application is highly available because the Sun Cluster HA for Sun Grid Engine data service is not yet installed.

Note – If any step in this procedure fails, see your Sun Grid Engine documentation for more information about how to verify the Sun Grid Engine installation.

▼ How to Verify the Installation and Configuration of Sun Grid Engine

You verify the installation and configuration of Sun Grid Engine by submitting a dummy job and checking that the required processes are running.

- 1 **Log in to the master host as the administrative user whose account you created in “[Preparing the Nodes and Disks](#)” on page 16.**
- 2 **Set the `SGE_ROOT` environment variable to the directory for the root of Sun Grid Engine file system that you created in “[Preparing the Nodes and Disks](#)” on page 16.**

```
$ SGE_ROOT=sge-root-dir  
$ export SGE_ROOT
```

3 Start the script that modifies your environment to enable Sun Grid Engine to run.

```
$ . $SGE_ROOT/default/common/settings.sh
```

4 Submit a dummy job to Sun Grid Engine.

```
$ qsub $SGE_ROOT/examples/jobs/sleeper.sh  
your job 1 (*Sleeper*) has been submitted
```

5 On the master host, confirm that these processes are running:

- sge_qmaster
- sge_schedd

```
# ps -ef | grep sge_  
root 429 1 0 Jul 27 3:37 /global/gridmaster/bin/solaris64/sge_qmaster  
root 429 1 0 Jul 27 3:37 /global/gridmaster/bin/solaris64/sge_schedd
```

6 View the global configuration of the grid.

- If you are using the command line, type the following command:

```
$ qconf -sconf
```
- If you are using the QMON graphical user interface (GUI), select Cluster Configuration.

7 On at minimum one execution host, confirm that these processes are running:

- sge_execd

```
# ps -ef | grep sge_  
root 451 1 0 Jul 27 3:37 /global/gridmaster/bin/solaris64/sge_execd
```

Installing the Sun Cluster HA for Sun Grid Engine Packages

If you did not install the Sun Cluster HA for Sun Grid Engine packages during your initial Sun Cluster installation, perform this procedure to install the packages. To install the packages, use the Sun Java™ Enterprise System Installation Wizard.

Note – You need to install the Sun Cluster HA for Sun Grid Engine packages in the global cluster and not in the zone cluster.

▼ How to Install the Sun Cluster HA for Sun Grid Engine Packages

Perform this procedure on each cluster node where you are installing the Sun Cluster HA for Sun Grid Engine packages.

You can run the Sun Java Enterprise System Installation Wizard with a command-line interface (CLI) or with a graphical user interface (GUI). The content and sequence of instructions in the CLI and the GUI are similar.

Before You Begin Ensure that you have the Sun Java Availability Suite DVD-ROM.

If you intend to run the Sun Java Enterprise System Installation Wizard with a GUI, ensure that your DISPLAY environment variable is set.

1 On the cluster node where you are installing the data service packages, become superuser.

2 Load the Sun Java Availability Suite DVD-ROM into the DVD-ROM drive.

If the Volume Management daemon `vold(1M)` is running and configured to manage DVD-ROM devices, the daemon automatically mounts the DVD-ROM on the `/cdrom` directory.

3 Change to the Sun Java Enterprise System Installation Wizard directory of the DVD-ROM.

- **If you are installing the data service packages on the SPARC® platform, type the following command:**

```
# cd /cdrom/cdrom0/Solaris_sparc
```

- **If you are installing the data service packages on the x86 platform, type the following command:**

```
# cd /cdrom/cdrom0/Solaris_x86
```

4 Start the Sun Java Enterprise System Installation Wizard.

```
# ./installer
```

5 When you are prompted, accept the license agreement.

If any Sun Java Enterprise System components are installed, you are prompted to select whether to upgrade the components or install new software.

6 From the list of Sun Cluster agents under Availability Services, select the data service for Sun Grid Engine.

- 7 If you require support for languages other than English, select the option to install multilingual packages.**
English language support is always installed.
- 8 When prompted whether to configure the data service now or later, choose Configure Later.**
Choose Configure Later to perform the configuration after the installation.
- 9 Follow the instructions on the screen to install the data service packages on the node.**
The Sun Java Enterprise System Installation Wizard displays the status of the installation. When the installation is complete, the wizard displays an installation summary and the installation logs.
- 10 (GUI only) If you do not want to register the product and receive product updates, deselect the Product Registration option.**
The Product Registration option is not available with the CLI. If you are running the Sun Java Enterprise System Installation Wizard with the CLI, omit this step.
- 11 Exit the Sun Java Enterprise System Installation Wizard.**
- 12 Unload the Sun Java Availability Suite DVD-ROM from the DVD-ROM drive.**
 - a. To ensure that the DVD-ROM is not being used, change to a directory that does *not* reside on the DVD-ROM.**
 - b. Eject the DVD-ROM.**

```
# eject cdrom
```

Next Steps Refer to the *Sun Cluster Data Service for NFS Guide for Solaris OS* on how to also install the Sun Cluster HA for NFS packages.

Configuring the HASStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine

For maximum availability of the Sun Grid Engine application, resources that Sun Cluster HA for Sun Grid Engine requires must be available before the Sun Grid Engine management tier is started. An example of such a resource is the Sun Grid Engine file system. To ensure that these resources are available, configure the HASStoragePlus resource type to work with Sun Cluster HA for Sun Grid Engine.

For information about the relationship between resource groups and disk device groups, see “Relationship Between Resource Groups and Device Groups” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Configuring the `HASStoragePlus` resource type to work with Sun Cluster HA for Sun Grid Engine involves the following operations:

- Synchronizing the startups between resource groups and disk device groups as explained in “Synchronizing the Startups Between Resource Groups and Device Groups” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
- Registering and configuring an `HASStoragePlus` resource

▼ How to Register and Configure an `HASStoragePlus` Resource

1 Become superuser on a node in the cluster that will host Sun Grid Engine.

2 Register the `SUNW.HASStoragePlus` resource type.

```
# clresourcetype register SUNW.HASStoragePlus
```

3 Add an `HASStoragePlus` resource for the Sun Grid Engine file system to the resource group that you created in “How to Enable Sun Grid Engine to Run in a Cluster” on page 20.

```
# clresource create \  
-g sge-rg \  
-t SUNW.HASStoragePlus \  
-p FilesystemMountPoints=sge-root \  
sge-hasp-rs
```

```
-g sge-rg
```

Specifies that the resource is to be added to the resource group that you created in “How to Enable Sun Grid Engine to Run in a Cluster” on page 20

```
-p FilesystemMountPoints=sge-root
```

Specifies that the mount point for this file system is the root of the Sun Grid Engine file system

```
sge-hasp-rs
```

Specifies that the resource that you are creating is named `sge-hasp-rs`

Configuring Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine

You must use NFS to export the Sun Grid Engine file system to the noncluster nodes. The NFS server that exports this file system must also be protected against failure. To protect the NFS server against failure, use the Sun Cluster HA for NFS data service.

The procedure that follows explains only the special requirements for using Sun Cluster HA for NFS with Sun Cluster HA for Sun Grid Engine. For complete information about installing and configuring Sun Cluster HA for NFS, see *Sun Cluster Data Service for NFS Guide for Solaris OS*.

▼ How to Configure Sun Cluster HA for NFS for Use With Sun Cluster HA for Sun Grid Engine

Note – Commands in this procedure assume that you have set the `$SGE_ROOT` environment variable to specify the root of the Sun Grid Engine file system.

1 Register the `SUNW.nfs` resource type.

```
# clresourcetype register SUNW.nfs
```

2 From any cluster node, create a directory for NFS configuration files.

Create the directory under root of the Sun Grid Engine file system. Name the directory `SUNW.nfs`.

```
# mkdir -p $SGE_ROOT/SUNW.nfs
```

3 In the directory that you created in [Step 2](#), create a file that contains the `share` command for the root of the Sun Grid Engine file system.

Name the file the `dfstab.sge-nfs-rs`, where `sge-nfs-rs` is the name of the NFS resource that you will create in [Step 4](#).

```
# echo "share -F nfs -o rw sge-root" \  
> $SGE_ROOT/SUNW.nfs/dfstab.sge-nfs-rs
```

4 Add a `SUNW.nfs` resource to the failover resource group that you created in [“How to Enable Sun Grid Engine to Run in a Cluster”](#) on page 20.

```
# clresource create \  
-g sge-rg \  
-t SUNW.nfs \  
-p Resource_dependencies=sge-hasp-rs \  
sge-nfs-rs
```

Example 2 Creating a `dfstab` File for the Root of the Sun Grid Engine File System

This example shows the command for creating a `dfstab` file for the root of the Sun Grid Engine file system.

- The root of the Sun Grid Engine file system is `/global/gridmaster`.
- The name of the NFS resource for which this file is created is `sge-nfs-rs`.

```
# echo "share -F nfs -o rw /global/gridmaster" \  
> /global/gridmaster/SUNW.nfs/dfstab.sge-nfs-rs
```

Registering and Configuring Sun Cluster HA for Sun Grid Engine

Before you perform this procedure, ensure that the Sun Cluster HA for Sun Grid Engine data service packages are installed.

Use the configuration and registration files in the `/opt/SUNWscsge/util` directory to register the Sun Cluster HA for Sun Grid Engine resources. The files define the dependencies that are required between Sun Grid Engine components. For information about these dependencies, see [“Dependencies Between Sun Grid Engine Components” on page 14](#). For a listing of these files, see [Files for Configuring and Removing Sun Cluster HA for Sun Grid Engine Resources](#).

Registering and configuring Sun Cluster HA for Sun Grid Engine involves the tasks that are explained in the following sections:

1. [“Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources” on page 27](#)
2. [“How to Create and Enable Sun Cluster HA for Sun Grid Engine Resources” on page 29](#)

Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources

Sun Cluster HA for Sun Grid Engine provides scripts that automate the process of configuring and removing Sun Cluster HA for Sun Grid Engine resources. These scripts obtain configuration parameters from the `sge_config` file in the `/opt/SUNWscsge/util/` directory. To specify configuration parameters for Sun Cluster HA for Sun Grid Engine resources, edit the `sge_config` file.

Each configuration parameter in the `sge_config` file is defined as a keyword-value pair. The `sge_config` file already contains the required keywords and equals signs. For more information, see [“Listing of `sge_config`” on page 35](#). When you edit the `sge_config` file, add the required value to each keyword. Use the values that you identified in [“Configuration Planning Questions” on page 15](#).

The keyword-value pairs in the `sge_config` file are as follows:

```
QMASTERRS=sge-qmaster-rs  
SCHEDDRS=sge-schedd-rs  
MASTERRG=sge-rg  
MASTERLH=sge-lh-rs
```

```

MASTERPORT=portno
MASTERHASP=sge-hasp-rs
SGE_ROOT=sge-root-dir
SGE_CELL=cell-name
SGE_VER=6.0
    
```

The meaning and permitted values of the keywords in the `sge_conf ig` file are as follows:

`QMASTERRS=sge-qmaster-rs`

Specifies the name that you are assigning to the resource for the Sun Grid Engine queue master daemon `sge_qmaster`. This must be defined.

`SCHEDDRS=sge-schedd-rs`

Specifies the name that you are assigning to the resource for the Sun Grid Engine scheduling daemon `sge_schedd`. This must be defined.

`MASTERRG=sge-rg`

Specifies the name of the resource group that contains the Sun Cluster HA for Sun Grid Engine resources. This name must be the name that you assigned when you created the resource group as explained in [“How to Enable Sun Grid Engine to Run in a Cluster”](#) on page 20. This must be defined.

`MASTERLH=sge-lh-rs`

Specifies the name of the logical host name resource for Sun Grid Engine. This name must be the name that you assigned when you created the resource in [“How to Enable Sun Grid Engine to Run in a Cluster”](#) on page 20. This must be defined.

`MASTERPORT=portno`

Specifies the port number that is configured for `sge_qmaster`, the default is set to 536. It must be an integer and must be defined.

`MASTERHASP=sge-hasp-rs`

Specifies the name of the `SUNW.HASStoragePlus` resource for Sun Grid Engine. This name must be the name that you assigned when you created the resource in [“Configuring the HASStoragePlus Resource Type to Work With Sun Cluster HA for Sun Grid Engine”](#) on page 24. If this resource is used it must be defined.

`SGE_ROOT=sge-root-dir`

Specifies the root directory of the Sun Grid Engine file system. This directory must be the directory that you created for root of the Sun Grid Engine file system in [“Preparing the Nodes and Disks”](#) on page 16. This must be defined.

`SGE_CELL=cell-name`

Specifies the cell that Sun Grid Engine references. This must be defined.

`SGE_VER=6.0`

Specifies the version of the installed Sun Grid Engine configuration. This keyword must be defined and set the value to "6.0".

Note – You must set the `SGE_VER` keyword to “6.0”, even if you are using Sun Grid Engine version 6.1.

EXAMPLE 3 Sample `sge_config` File

This example shows an `sge_config` file in which configuration parameters are set as follows:

- The name of the resource for the Sun Grid Engine queue master daemon `sge_qmaster` is `sge_qmaster-rs`.
- The name of the resource for the Sun Grid Engine scheduling daemon `sge_schedd` is `sge_schedd-rs`.
- The name of the resource group that contains the Sun Cluster HA for Sun Grid Engine resources is `sge-rg`.
- The name of the logical host name resource for Sun Grid Engine is `sge-lh-rs`.
- The port number for `sge_qmaster` is set to 536.
- The name of the `SUNW.HASStoragePlus` resource for Sun Grid Engine is `sge-hasp-rs`.
- The root directory of the Sun Grid Engine file system is `/global/gridmaster`.
- Sun Grid Engine references the default cell.
- The version for Sun Grid Engine is set to `6.0`.

```
QMASTERRS=sge_qmaster-rs
SCHEDDRS=sge_schedd-rs
MASTERRG=sge-rg
MASTERLH=sge-lh-rs
MASTERPORT=536
MASTERHASP=sge-hasp-rs
SGE_ROOT=/global/gridmaster
SGE_CELL=default
SGE_VER=6.0
```

▼ How to Create and Enable Sun Cluster HA for Sun Grid Engine Resources

Before you begin, ensure that you have edited the `sge_config` file or a copy of it to specify configuration parameters for Sun Cluster HA for Sun Grid Engine resources. For more information, see [“Specifying Configuration Parameters for Sun Cluster HA for Sun Grid Engine Resources”](#) on page 27.

1 Register the `SUNW.gds` resource type.

```
# clresourcetype register SUNW.gds
```

- 2 Go to the directory that contains the script for creating the Sun Grid Engine resources.

```
# cd /opt/SUNWscsge/util/
```

- 3 Run the script that creates the Sun Grid Engine resources.

```
# ./sge_register -f /mypath/sge_config
```

- 4 Bring online the failover resource group that you created in [“How to Enable Sun Grid Engine to Run in a Cluster” on page 20](#).

This resource group contains the following resources:

- Logical host name resource
- HASStoragePlus resource
- NFS resource
- Sun Grid Engine application resources

```
# clresourcegroup online -M sge-rg
```

sge-rg Specifies the resource group that you created in [“How to Enable Sun Grid Engine to Run in a Cluster” on page 20](#) is to be brought online



Caution – Make sure that the Sun Grid Engine daemons (`sge_qmaster` and `sge_schedd`) are not running before bringing the failover resource group online. They may be running because the `install_qmaster` installation script started them or they are still running after performing the verification described in [“How to Verify the Sun Cluster HA for Sun Grid Engine Installation and Configuration” on page 31](#).

Setting Sun Cluster HA for Sun Grid Engine Extension Properties

Extension properties for Sun Cluster HA for Sun Grid Engine resources are set when you run the script that creates these resources. You need to set these properties only if you require values other than the values that are set by the script. For information about Sun Cluster HA for Sun Grid Engine extension properties, see the `SUNW.gds(5)` man page. You can update some extension properties dynamically. You can update other properties, however, only when you create or disable a resource. The Tunable entry indicates when you can update a property.

To update an extension property of a resource, run the `clresource(1CL)` command with the following option to modify the resource:

```
-p property=value
```

`-p property` Identifies the extension property that you are setting

value Specifies the value to which you are setting the extension property

You can also use the procedures in [Chapter 2, “Administering Data Service Resources,” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*](#) to configure resources after the resources are created.

Verifying the Sun Cluster HA for Sun Grid Engine Installation and Configuration

After you install, register, and configure Sun Cluster HA for Sun Grid Engine, verify the Sun Cluster HA for Sun Grid Engine installation and configuration. Verifying the Sun Cluster HA for Sun Grid Engine installation and configuration determines if the Sun Cluster HA for Sun Grid Engine data service makes the Sun Grid Engine application highly available.

▼ How to Verify the Sun Cluster HA for Sun Grid Engine Installation and Configuration

- 1 Become superuser a node that will host Sun Grid Engine.
- 2 Verify that all Sun Grid Engine resources are online.

```
# cluster status -t rg,rs
```
- 3 If a Sun Grid Engine resource is not online, enable the resource.

```
# clresource enable sge-rs
```
- 4 Switch the Sun Grid Engine resource group to another cluster node.

```
# clresourcegroup switch -n node sge-rg
```

Tuning the Sun Cluster HA for Sun Grid Engine Fault Monitors

The Sun Cluster HA for Sun Grid Engine fault monitors verify that the following daemons are running correctly:

- Queue master daemon `sge_qmaster`
- Scheduling daemon `sge_schedd`

Each Sun Cluster HA for Sun Grid Engine fault monitor is contained in the resource that represents Sun Grid Engine component. You create these resources when you register and

configure Sun Cluster HA for Sun Grid Engine. For more information, see [“Registering and Configuring Sun Cluster HA for Sun Grid Engine” on page 27](#).

System properties and extension properties of these resources control the behavior of the fault monitor. The default values of these properties determine the preset behavior of the fault monitor. The preset behavior should be suitable for most Sun Cluster installations. Therefore, you should tune the Sun Cluster HA for Sun Grid Engine fault monitor *only* if you need to modify this preset behavior.

Tuning the Sun Cluster HA for Sun Grid Engine fault monitors involves the following tasks:

- Setting the interval between fault monitor probes
- Setting the timeout for fault monitor probes
- Defining the criteria for persistent faults
- Specifying the failover behavior of a resource

For more information, see [“Tuning Fault Monitors for Sun Cluster Data Services” in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*](#).

Debugging Sun Cluster HA for Sun Grid Engine

The config file in the `/opt/SUNWscsge/etc` directory enables you to activate debugging for Sun Grid Engine resources. This file enables you to activate debugging for all Sun Grid Engine resources or for a specific Sun Grid Engine resource on a particular node. If you require debugging for Sun Cluster HA for Sun Grid Engine to be enabled throughout the cluster, repeat this procedure on all nodes.

▼ How to Activate Debugging for Sun Cluster HA for Sun Grid Engine

1 Determine whether debugging for Sun Cluster HA for Sun Grid Engine is active.

If debugging is inactive, `daemon.notice` is set in the file `/etc/syslog.conf`.

```
# grep daemon /etc/syslog.conf
*.err;kern.debug;daemon.notice;mail.crit      /var/adm/messages
*.alert;kern.err;daemon.err                  operator
#
```

2 If debugging is inactive, edit the `/etc/syslog.conf` file to change `daemon.notice` to `daemon.debug`.

3 Confirm that debugging for Sun Cluster HA for Sun Grid Engine is active.

If debugging is active, `daemon.debug` is set in the file `/etc/syslog.conf`.

```
# grep daemon /etc/syslog.conf
*.err;kern.debug;daemon.debug;mail.crit      /var/adm/messages
*.alert;kern.err;daemon.err                  operator
#
```

4 Restart the syslogd daemon.

- If your operating system is Solaris 9, perform:

```
# pkill -1 syslogd
```

- If your operating system is Solaris 10, perform:

```
# svcadm restart system-log
```

5 Edit the `/opt/SUNWscsge/etc/config` file to change `DEBUG=` to `DEBUG=ALL` or `DEBUG=sgc-rs`.

```
# cat /opt/SUNWscsge/etc/config
#
# Copyright 2006 Sun Microsystems, Inc. All rights reserved.
# Use is subject to license terms.
#
# ident "@(#)config 1.1 06/02/18 SMI"
#
# Usage:
#     DEBUG=<RESOURCE_NAME> or ALL
#
DEBUG=ALL
#
```

Note – To deactivate debugging, reverse the preceding steps.



Files for Configuring and Removing Sun Cluster HA for Sun Grid Engine Resources

The /opt/SUNWscsge/util directory contains files that automate the process of configuring and removing Sun Cluster HA for Sun Grid Engine resources. Listings of these files are provided in the following sections:

- [“Listing of sge_config” on page 35](#)
- [“Listing of sge_register” on page 36](#)
- [“Listing of sge_remove” on page 38](#)

Listing of sge_config

```
#
# Copyright 2006 Sun Microsystems, Inc. All rights reserved.
# Use is subject to license terms.
#
# ident      "@(#)sge_config    1.3    06/07/14 SMI"
#
# This file will be sourced in by sge_register and the parameters
# listed below will be used.
#
# These parameters can be customized in (key=value) form and are always needed:
#
# QMASTERRS - name of the resource for the sge_qmaster daemon.
# SCHEDDRS  - name of the resource for the sge_schedd daemon.
# MASTERRG  - name of the resource group containing the resources
#             for sge_qmaster and sge_schedd.
# MASTERPORT - tcp port number that sge_qmaster will listen to
#             (normally 536).
# MASTERLH  - name of the LogicalHostname SC resource within MASTERRG.
# MASTERHASP - name of the HAStoragePlus SC resource within MASTERRG.
# SGE_ROOT  - SGE_ROOT of this Sun GridEngine Installation.
# SGE_CELL  - SGE_CELL of this Sun GridEngine Installation.
# SGE_VER   - Version of this Sun GridEngine Installation.
```

```
#           Currently this can only be set to 6.0 .
#

QMASTERRS=
SCHEDDRS=
MASTERRG=
MASTERPORT=536
MASTERLH=
MASTERHASP=
SGE_ROOT=
SGE_CELL=
SGE_VER=6.0
```

Listing of sge_register

```
#!/bin/ksh
#
# Copyright 2006 Sun Microsystems, Inc. All rights reserved.
# Use is subject to license terms.
#
# ident    "@(#)sge_register.ksh    1.6    06/07/14 SMI"
#

MYCONFIG=""
MYNAME='basename $0'
GDSTYPE=SUNW.gds

typeset opt

while getopts 'f:' opt
do
    case "${opt}" in
        f)    MYCONFIG=${OPTARG};;
        *)    echo "ERROR: ${MYNAME} Option ${OPTARG} unknown - early End."
              echo "Only the option -f <filename> is valid."
              exit 1;;
    esac
done

if [ -n "${MYCONFIG}" ] && [ -f "${MYCONFIG}" ]
then
    echo "sourcing ${MYCONFIG}"
    . ${MYCONFIG}
else
    PKGCONF='dirname $0'/sge_config
    echo "sourcing ${PKGCONF}"
    . ${PKGCONF}
fi
```

```

fi

if [ "${SGE_VER}" = "6.0" ]; then
    # Register resource for sge_qmaster
    /usr/cluster/bin/scrgadm -a -j ${QMASTERRS} -g ${MASTERRG} -t ${GDSTYPE} \
-x Start_command="/opt/SUNWscsge/bin/sge_qmaster6/start_sge_qmaster \
-R ${QMASTERRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-x Stop_command="/opt/SUNWscsge/bin/sge_qmaster6/stop_sge_qmaster \
-R ${QMASTERRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-x Probe_command="/opt/SUNWscsge/bin/sge_qmaster6/probe_sge_qmaster \
-R ${QMASTERRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-y Port_list=${MASTERPORT}/tcp -y Network_resources_used=${MASTERLH} \
-x Stop_signal=9 \
-x probe_timeout=90 -y Thorough_probe_interval=120 \
-y retry_count=2 -y retry_interval=900 \
-y Resource_dependencies=${MASTERHASP}
    St=$?
    if [ "${St}" -ne 0 ]; then
        echo "Error: Registration of resource ${QMASTERRS} failed."
        echo "Please correct the wrong parameters."
        exit 1
    else
        echo "Registration of resource ${QMASTERRS} succeeded"
    fi

    # Register resource for sge_schedd
    /usr/cluster/bin/scrgadm -a -j ${SCHEDDRS} -g ${MASTERRG} -t ${GDSTYPE} \
-x Start_command="/opt/SUNWscsge/bin/sge_schedd6/start_sge_schedd \
-R ${SCHEDDRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-x Stop_command="/opt/SUNWscsge/bin/sge_schedd6/stop_sge_schedd \
-R ${SCHEDDRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-x Probe_command="/opt/SUNWscsge/bin/sge_schedd6/probe_sge_schedd \
-R ${SCHEDDRS} -G ${MASTERRG} -S ${SGE_ROOT} -C ${SGE_CELL}" \
-y Port_list=${MASTERPORT}/tcp -y Network_resources_used=${MASTERLH} \
-x Stop_signal=9 \
-x probe_timeout=90 -y Thorough_probe_interval=120 \
-y retry_count=2 -y retry_interval=900 \
-y Resource_dependencies=${QMASTERRS}
    St=$?
    if [ "${St}" -ne 0 ]; then
        echo "Error: Registration of resource ${SCHEDDRS} failed."
        echo "Please correct the wrong parameters and try again."
        echo "Removing resource ${QMASTERRS}."
        /usr/cluster/bin/scrgadm -r -j ${QMASTERRS}
        exit 1
    else
        echo "Registration of resource ${SCHEDDRS} succeeded"
    fi
fi

```

```
else
    echo "Fatal: Please set variable SGE_VER properly in 'dirname $0'/sge_config!"
    exit 1
fi
```

Listing of sge_remove

```
#!/bin/ksh
#
# Copyright 2006 Sun Microsystems, Inc. All rights reserved.
# Use is subject to license terms.
#
# ident    "@(#)sge_remove.ksh    1.3    06/04/24 SMI"
#

MYCONFIG=""
MYNAME='basename $0'

typeset opt

while getopts 'f:' opt
do
    case "${opt}" in
        f)    MYCONFIG=${OPTARG};;
        *)    echo "ERROR: ${MYNAME} Option ${OPTARG} unknown - early End."
              echo "Only the option -f <filename> is valid."
              exit 1;;
    esac
done

if [ -n "${MYCONFIG}" ] && [ -f "${MYCONFIG}" ]
then
    echo "sourcing ${MYCONFIG}"
    . ${MYCONFIG}
else
    PKGCONF='dirname $0'/sge_config
    echo "sourcing ${PKGCONF}"
    . ${PKGCONF}
fi

/usr/cluster/bin/scswitch -n -j ${SCHEDDRS}
/usr/cluster/bin/scswitch -n -j ${QMASTERRS}

/usr/cluster/bin/scrgadm -r -j ${SCHEDDRS}
/usr/cluster/bin/scrgadm -r -j ${QMASTERRS}
```

Index

A

administrative user account, 17

B

Berkley DB spooling server, Sun Grid Engine
application, 12

binary files

- determining location, 15
- disk space requirements, 13
- distribution among file systems, 15

boot-time options

- restrictions, 12
- setting, 20

C

`clnode` command, 8

commands, node information, 8

`config` file, 32

configuration parameters, Sun Grid
Engineresources, 27-29

configuring

- file systems, 24-25
- HASStoragePlus resource, 24-25
- Sun Cluster HA for NFS, 25-27
- Sun Cluster HA for Sun Grid Engine
 - overview, 10-11
 - performing, 27-31
 - planning, 11-16

configuring (*Continued*)

Sun Grid Engine application, 18-21

D

daemons, shadow, 12

debugging, Sun Cluster HA for Sun Grid Engine, 32-33

dependencies, Sun Grid Engine components, 14

device groups, 24

`dfstab` file, 26

directories

`/opt/SUNWscsge/util`, 27

spool

- disk space requirements, 13
- distribution among file systems, 15

disk space requirements, 13

disks, preparing for Sun Grid Engine

installation, 16-18

E

environment variables

Sun Grid Engine root directory, 19, 21

`/etc/services` file, 17

extension properties

- effect on fault monitors, 32
- Sun Cluster HA for Sun Grid Engine, 30-31

F

fault monitors, tuning, 31-32

file systems

configuring, 24-25

distribution of files on, 15

performance, 15

requirements, 14

Sun Grid Engine

configuration requirements, 14

creating root directory, 17

environment variables, 19, 21

installing software on, 19

files

config, 32

determining location, 15

dfstab, 26

/etc/services, 17

/opt/SUNWscsge/etc/config, 32

services, 17

settings.sh, 22

sge_config

listing, 35-36

syntax, 27-29

sge_register

listing, 36-38

running, 30

sge_remove, 38

system configuration, 16

G

global zone, 14

grid configuration, effect on file systems, 15

group IDs, 20

H

HASStoragePlus resource

configuring, 24-25

creating, 25

enabling, 30

help, 8

hosts, file system requirements, 15

I

installing

Sun Cluster HA for Sun Grid Engine, 22-24

overview, 10-11

verifying installation, 31

Sun Grid Engine application, 18-21

J

jobs, memory requirements, 13

L

location, Sun Grid Engine files, 15

logical host names

resource group for

creating, 20

planning, 15

resources

creating, 21

enabling, 30

planning, 16

M

management tier, operating system for, 13

memory requirements, 13

messages file, 8

N

network addresses, *See* logical host names

NFS file system

configuration requirements, 14

distribution of files on, 15

performance, 15

resources

configuring, 25-27

enabling, 30

nodes, preparing for Sun Grid Engine

installation, 16-18

non-global zones, 14

O

operating systems, requirements on, 13
 /opt/SUNWscsge/util directory, 27
 /opt/SUNWscsge/etc/config file, 32

P

packages, 22-24
 performance, NFS file system, 15
 port numbers
 sge_execd service, 17
 sge_qmaster service, 17
 protocols
 sge_execd service, 17
 sge_qmaster service, 17
 prtconf -v command, 8
 prtdiag -v command, 8
 psrinfo -v command, 8

R

registering
 SUNW.gds resource type, 29
 SUNW.HAStoragePlus resource type, 25
 SUNW.nfs resource type, 26
 resource groups
 creating, 20
 enabling, 30
 planning, 15
 resource types
 fault monitors, 31-32
 SUNW.gds
 extension properties, 30-31
 registering, 29
 SUNW.HAStoragePlus
 dependencies on, 14
 registering, 25
 SUNW.nfs
 registering, 26

resources

dependencies between, 14
 HAStoragePlus
 creating, 25
 enabling, 30
 logical host name
 creating, 21
 enabling, 30
 planning, 16
 NFS file system
 configuring, 25-27
 enabling, 30
 Sun Grid Engine application
 debugging, 32-33
 enabling, 30
 extension properties, 30-31
 planning, 16
 Sun Grid Engine application
 configuration parameters, 27-29
 restrictions
 Sun Cluster HA for Sun Grid Engine
 configuration, 12
 zones, 14
 root directory
 creating, 17
 environment variables, 19, 21
 installing software on, 19

S

services file, 17
 settings.sh file, 22
 sge_config file
 listing, 35-36
 syntax, 27-29
 sge_execd service, 17
 sge_qmaster service, 17
 sge_register file
 listing, 36-38
 running, 30
 sge_remove file, 38
 SGE_ROOT environment variable, 19, 21
 sgeadmin user account, 17
 shadow daemon, Sun Grid Engine application, 12

- share command, 26
 - show- rev subcommand, 8
 - showrev -p command, 8
 - software packages, 22-24
 - spool directories
 - disk space requirements, 13
 - distribution among file systems, 15
 - spooling server, Berkley DB, 12
 - starting
 - Sun Grid Engine application, 12, 20
 - Sun Cluster HA for NFS
 - configuration requirements, 14
 - configuring, 25-27
 - Sun Cluster HA for Sun Grid Engine
 - overview, 10
 - configuration
 - performing, 27-31
 - planning, 11-16
 - with HASToragePlus resource, 24-25
 - debugging, 32-33
 - fault monitors, 31-32
 - installing, 22-24
 - overview, 10-11
 - verifying installation, 31
 - service configuration, 13-14
 - software packages, installing, 22-24
 - Sun Grid Engine, Berkley DB spooling server, 12
 - Sun Grid Engine application
 - administrative user account, 17
 - configuring, 18-21
 - dependencies between components, 14
 - disk space requirements, 13
 - disks, 16-18
 - fault monitor, 31-32
 - file system, 14
 - creating root directory, 17
 - environment variables, 19, 21
 - installing software on, 19
 - group IDs, 20
 - installing, 18-21
 - location of binary files, 15
 - memory requirements, 13
 - nodes, 16-18
 - Sun Grid Engine application (*Continued*)
 - resource group for
 - creating, 20
 - enabling, 30
 - planning, 15
 - resources
 - configuration parameters, 27-29
 - enabling, 30
 - extension properties, 30-31
 - planning, 16
 - shadow daemon, 12
 - software versions, 12
 - start at boot option, 12, 20
 - verifying installation and configuration, 21-22
 - SUNW.gds resource type
 - extension properties, 30-31
 - registering, 29
 - SUNW.HASToragePlus resource type
 - dependencies on, 14
 - registering, 25
 - SUNW.nfs resource type, registering, 26
 - system configuration files, location, 16
 - system properties, effect on fault monitors, 32
- ## T
- TCP (Transmission Control Protocol), 17
 - technical support, 8
 - Transmission Control Protocol (TCP), 17
 - tuning, fault monitors, 31-32
- ## U
- user accounts, 17
- ## V
- /var/adm/messages file, 8
 - verifying, Sun Grid Engine application, 21-22
 - version requirements, Sun Grid Engine application, 12

Z

zones, 14

