



Sun HPC ClusterTools™ 6 Software Release Notes

Sun Microsystems, Inc.
www.sun.com

Part No. 819-4129-11
September 2006, Revision A

Submit comments about this document at: <http://www.sun.com/hwdocs/feedback>

Copyright 2006 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 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 of the U.S. patents listed at <http://www.sun.com/patents> and one or more additional patents or pending patent applications in the U.S. and in other countries.

This document and the product to which it pertains are distributed under licenses restricting their use, copying, distribution, and decompilation. No part of the product or of this document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any.

Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

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 in other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, docs.sun.com, Solaris, Sun HPC ClusterTools, Sun Performance Library, and UltraSPARC are trademarks or registered trademarks of Sun Microsystems, Inc. in the U.S. and in other countries.

AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices.

All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and in 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.

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

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 2006 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, Etats-Unis. Tous droits réservés.

Sun Microsystems, Inc. a les droits de propriété intellectuelle relatant à la technologie incorporée dans le produit qui est décrit dans ce document. En particulier, et sans la limitation, ces droits de propriété intellectuelle peuvent inclure un ou plus des brevets américains énumérés à <http://www.sun.com/patents> et un ou les brevets plus supplémentaires ou les applications de brevet en attente dans les Etats-Unis et dans les autres pays.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a.

Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées des systèmes Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, docs.sun.com, Solaris, Sun HPC ClusterTools, Sun Performance Library, et UltraSPARC sont des marques de fabrique ou des marques déposées de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays.

AMD, Opteron, le logo AMD, et le logo AMD Opteron sont des marques de fabrique ou des marques déposées de Advanced Micro Devices.

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.

LA DOCUMENTATION EST FOURNIE "EN L'ÉTAT" 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 CONTREFAÇON.



Contents

Major New Features 1

Product Migration 2

Prism 2

Sun S3L 2

SCSL 2

Cluster Console Manager 2

RSM 3

Related Software 3

Solaris 10 OS Recommended Patches 4

Differences Between SGE and N1GE Resource Managers 4

Outstanding CRs 6

x64-Related Issues 6

SPARC-Based Platform Issues 6

Issues Related To Both SPARC- and x64-Based Platforms 7

Sun HPC ClusterTools 6 Software Release Notes

This document describes late-breaking news about the Sun HPC ClusterTools™ 6 software. The information is organized into the following sections:

- Major New Features
- Product Migration
- Related Software
- Outstanding Bugs
- Performance Issues

Major New Features

The major new features of the Sun HPC ClusterTools 6 software include:

- Installs on the Solaris 10 Operating System
- Support for the UltraSPARC™ IV+ processor and x64 32- and 64-bit applications on the AMD Opteron™ architecture
- Applications can be built with Sun Studio versions 8, 9, 10, and 11
- Example programs that can be used with DTrace, the dynamic tracing facility in the Solaris Operating System (Solaris OS) software. These example programs are located at `/opt/SUNWhpc/examples/mpi/dtrace`.

Product Migration

Several components of the ClusterTools 5 product are not available in ClusterTools 6. This section contains a list of the components that are not available, as well as suggestions for replacing the functionality.

Prism

The Prism debugger has been removed from ClusterTools 6 software. The TotalView debugger from Etnus (<http://www.etnus.com>) supports the debugging of Sun MPI programs on SPARC-based platforms running the Solaris OS. For Sun MPI programs on x64- and SPARC-based platforms, dbx is supported.

The DDT debugger from Allinea (<http://www.allinea.com>) supports the debugging of Sun MPI programs on AMD x64-based platforms running the Solaris OS. (See “[Sun MPI Message Queues Not Accurate or Visible from Allinea DDT \(CR 6396838\)](#)” on page 6 for more information on a known issue with DDT.)

Sun S3L

The Sun Scalable Scientific Subroutine Library (Sun S3L) is not supported in ClusterTools 6 software. Many of the S3L functions can be replaced with functions provided in the public domain libraries PETSc and ScaLAPACK.

SCSL

Sun ClusterTools 6 software is not released under the Sun Community Source License (SCSL).

Cluster Console Manager

The Cluster Console Manager (CCM) tools, `cconsole`, `ctelnet`, and `crlogin`, are no longer shipped with ClusterTools 6 software.

The CCM package, `SUNWwcon`, is bundled with the Sun Java Enterprise System (JES) and can be downloaded with JES at the following URL:

RSM

RSM (Remote Shared Memory) functionality has been removed from ClusterTools 6 software.

Related Software

The Sun HPC ClusterTools 6 software works with the following versions of related software:

- Solaris 10 03/05 or any subsequent Solaris 10 OS release that supports Sun HPC ClusterTools 6 software.

Note – If you have the Solaris 10 3/05 OS release installed, download and install the recommended patches for your platform type, as shown in [TABLE 1](#). These patches are available from SunSolve. Alternatively, you can upgrade to another release of the Solaris 10 OS, such as Solaris 10 1/06.

- Sun Studio 8, 9, 10 and 11 C, C++, and Fortran compilers
- Distributed resource management frameworks operating under integration with Sun CRE:
 - Sun N1GE (Sun N1 Grid Engine) Version 6.
 - Load Sharing Facility (LSF) HPC Version 6.2.
 - OpenPBS Portable Batch System (PBS) 2.3.16 and Altair PBS Professional 7.1 (formerly Veridian).
- Java Runtime Environment (JRE) 1.5 (or compatible release) for using the Sun HPC ClusterTools installation tool's graphic interface.
- TotalView 7.1 debugger supports debugging Sun MPI applications on SPARC-based systems running the Solaris OS. Compatibility limitations may exist. Please see www.etnus.com for compilers that TotalView supports.

Solaris 10 OS Recommended Patches

If you have the Solaris 10 3/05 release installed, you must install the following Solaris patches in order to run Sun HPC ClusterTools. The patch revisions shown in this table reflect the minimum version that supports Sun HPC ClusterTools.

TABLE 1 Solaris 10 3/05 Patches By Platform

Platform	Patches to Install
i386/x64	113000-07
	118344-06
	118844-28
	118885-01
	118891-01
	121127-01
	121208-02
SPARC	118822-27
	118884-01
	118890-01
	119689-06

Differences Between SGE and N1GE Resource Managers

Sun Grid Engine 5.2 (SGE) was supported on Sun HPC ClusterTools 5 software. Sun HPC ClusterTools software supports Sun N1 Grid Engine (N1GE) 6 as its SGE (Sun Grid Engine) resource manager. Previous versions of Sun HPC ClusterTools software supported Sun Grid Engine software. This section outlines the differences between the two resource managers.

Sun N1 Grid Engine (N1GE6) introduces new features and attributes that are different from those in SGE 5.2. These differences fall into four major categories:

- Parallel Environment (PE) attributes
- Queue attributes
- Modifying queues

For more information about N1GE, refer to the *Sun N1 Grid Engine 6 Administration Guide* (817-5677).

Parallel Environment (PE) Attributes

N1GE6 introduces a new parallel environment attribute:

urgency_slots min

The queue_list attribute in SGE 5.2 has been removed. The values for queue_list now appear as the values for the slots attribute in PE.

Queue Attributes

The following table illustrates the differences between certain queue attributes in N1GE6 and in SGE 5.2.

TABLE 2 Changes in Queue Attributes

Attribute	SGE 5.2 Value	N1GE6 Value
slots	Number of processes (for example, 8)	Values from SGE 5.2 queue_list attribute (for example, 1, [node1=4], [node2=4])
qtype	BATCH INTERACTIVE PARALLEL	BATCH INTERACTIVE (PARALLEL has been removed)
pe_list		cre make (pe_list must be specified for the cre PE that uses this queue)

In addition, N1GE provides enhanced suspend and resume scripts. Use these scripts to prevent MPI processes from continuing to run when SGE has issued a suspension.

The suspend_method script has the following attributes:

```
<sge-root>/mpi/SunHPCT5/suspend_sunmpi_ci.sh $job_pid $job_id
```

The resume_method script has the following attributes:

```
<sge-root>/mpi/SunHPCT6/suspend_sunmpi_ci.sh $job_pid $job_id
```

where <sge-root> is the path to the location where SGE is installed.

Modifying Queues

The command to modify a queue has changed Between SGE 5.2 and N1GE6. For example, to modify a queue in N1GE6, you would issue a command similar to the following (substituting the name of the queue for *queue-name*):

```
% qconf -mattr queue qtype "BATCH INTERACTIVE" queue-name
```

Note that the PARALLEL value for qtype has been removed in N1GE6.

The equivalent command in SGE 5.2 would be as follows:

```
% qconf -mqattr qtype "BATCH INTERACTIVE PARALLEL" queue-name
```

Outstanding CRs

This section highlights some of the outstanding CRs (Change Requests) for the Sun HPC ClusterTools 6 software components. A CR may be a defect, or it may be an RFE (Request For Enhancement).

Note – The heading of each CR description includes the CR's Bugster number in parentheses.

x64-Related Issues

Sun MPI Message Queues Not Accurate or Visible from Allinea DDT (CR 6396838)

This issue only affects the access of MPI message queues from the Allinea DDT debugger on the x64 platform. Message queue information will not be available. Other debugging functionality is not affected.

Workaround: Not available.

SPARC-Based Platform Issues

Job Aborts under Non-Global Zone When It is Run As a Non-root User (CR 6320925)

If a user who is not root attempts to run an HPC job under a Solaris 10 non-global zone, the job aborts and returns an error message similar to the following:

```
Job cre.6 on nodename: aborted due to an unexpected error.
```

This error occurs only on nodes with the Solaris 10 3/05 version installed. This error has been fixed in subsequent versions of the Solaris 10 OS. To fix the problem, either download and install patch 119689-06 (or higher) to Solaris 10 3/05, or upgrade to a more recent version of the Solaris 10 OS (such as Solaris 10 1/06).

TotalView, LSF, and ClusterTools Do Not Work Together (CR 6395112)

TotalView, LSF, and Sun HPC ClusterTools cannot be used together when debugging parallel MPI programs. the debugging process in TotalView does not work correctly when LSF is used as the resource manager.

Workaround: Use a different resource manager (such as CRE) when debugging with TotalView.

Issues Related To Both SPARC- and x64-Based Platforms

`pbs spawn` Failed in `pbsrun: tm_init()` Failed (TM_BADINIT) when `-np` is Not Specified (CR 6370836)

`pbsrun` returns failure messages from the PBS task manager `tm_init` (which returns the message TM_BADINIT) during spawn jobs if the `-np` option of `mprun` is not specified.

Workaround: Specify values for `-np` and `-nr` in `mprun` when you want to use the `MPI_Comm_spawn()` API or the `MPI_Comm_spawn_multiple()` API with PBS. Make sure that the total of the two values you specify does not exceed the total number of processes allocated in the PBS environment.

Node Failures Cause Stale Job Entries (CR 4692994)

If a node crashes while an MPI program is running, CRE does not remove the job entry from its database, so `mpps` continues to show the job indefinitely, often in states such as coring or exiting.

Workaround: To delete these stale jobs from the database, `su` to `root` and issue this command:

```
# mpkill -C
```

mpps and mpkill Do Not Work Properly With the LSF Integration (CR 6389722)

The `mpps` and `mpkill` commands do not work properly when running under LSF.

Workaround: Instead of using `mpps` and `mpkill`, use the LSF equivalent commands `bjobs` and `bkill`.

CRE Poor Diagnosability When Not Enough Physical Memory to Run Processes (CR 4857731)

When the physical memory requirements of all the processes an MPI application exceed the amount of memory available on a node, the `mprun` command returns an error similar to the following:

```
Job cre.1 on nodename: received signal KILL.
```

Workaround: Run the application on a node with sufficient physical memory.

mprun Shows Key Mismatch: Authentication Error Under SGE/N1GE6 (CR 6383190)

When you use SGE/Sun N1 Grid Engine as the resource manager, `mprun` can return an error message similar to the following:

```
mprun: tmrte_auth_verify_user: Key mismatch: Authentication  
error. Contact system administrator.
```

This condition occurs when SGE/N1GE6 has been improperly installed. `mprun` returns the error, but the error itself occurs within the resource manager.

Workaround: Reconfigure Sun N1 Grid Engine 6 with a new unused gid range.

mprun Picks the Wrong PBS Plugin for OpenPBS When Both PBS Professional and OpenPBS Are Running (CR 6397692)

If you have both the OpenPBS and PBS Professional resource managers installed and running on the same system, issuing the `mprun -x pbs` command causes `mprun` to always select PBS Professional.

Workaround: Shut down the PBS Professional daemons if you want to use OpenPBS. To shut down the PBS Professional daemons, type the following command:

```
# /etc/init.d/pbs stop
```

RPC Errors Cause ClusterTools Commands to Hang/Error Out w/S10 KU Patch 118833-17/18 or 118855-15 (CR 6459510)

This CR affects Sun HPC ClusterTools 6 software running on Solaris 10 6/06 OS, or a previous version of the Solaris 10 OS with kernel updates 118833-17 or -18 or 118855-15. The issue causes Sun HPC ClusterTools commands such as `mpinfo` to hang or to produce RPC-related error messages such as the following:

```
% mpinfo -N
mpinfo: tmrte_auth_create_context: host: RPC call
tmrte_auth_conf_3 timed out after 30 secs. Try using mprun -t to
increase timeout factor.
```

Workaround: Disable TCP fusion by typing the following command at the system prompt:

```
echo "set ip:do_tcp_fusion = 0x0" >> /etc/system
```

Next, reboot the system so this command can take effect.

