Sun HPC ClusterTools™ 7.1 Software Migration Guide

Sun Microsystems, Inc.
www.sun.com

Part No. 819-7484-11
October 2007, Revision A

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

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

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

Sun, Sun Microsystems, the Sun logo, Java, Netra, Solaris, Sun Ray, docs.sun.com, Sun HPC ClusterTools, Sun Cluster, Sun Fire, UltraSPARC, Sun Performance Library and Sun 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 architecture developed by Sun Microsystems, Inc.

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

Products covered by and information contained in this service manual are controlled by U.S. Export Control laws and may be subject to the export or import laws in other countries. Nuclear, missile, chemical 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, California 95054, Etats-Unis. Tous droits réservés.

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

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, Java, Netra, Solaris, Sun Ray, docs.sun.com, Sun HPC ClusterTools, Sun Cluster, Sun Fire, UltraSPARC, Sun Performance Library et Sun 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.

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

Les produits qui font l'objet de ce manuel d'entretien et les informations qu'il contient sont regis par la legislation americaine en matiere de controle des exportations et peuvent etre soumis au droit d'autres pays dans le domaine des exportations et importations. Les utilisations finales, ou utilisateurs finaux, pour des armes nucleaires, des missiles, des armes biologiques et chimiques ou du nucleaire maritime, directement ou indirectement, sont strictement interdites. Les exportations ou reexportations vers des pays sous embargo des Etats-Unis, ou vers des entites figurant sur les listes d'exclusion d'exportation americaines, y compris, mais de maniere non exclusive, la liste de personnes qui font objet d'un ordre de ne pas participer, d'une facon directe ou indirecte, aux exportations des produits ou des services qui sont regi par la legislation americaine en matiere de controle des exportations et la liste de ressortissants specifiquement designes, 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.

Please
Recycle

Contents

Preface v

1. Differences Between Sun HPC ClusterTools Versions 1

Comparing `mprun` Options to `mpirun` Options 1

Passing MPI Options 4

Specifying Selected CPUs for Applications 5

Thread Safety 5

Sun HPC ClusterTools-Specific Commands 5

Updating Applications Compiled With Previous Versions of Sun HPC ClusterTools Software 6

Comparing Sun MPI Environment Variables to Open MPI MCA Parameters 7

More Information About Parameters and Compiling Applications 7

2. Migrating Sun HPC ClusterTools 6 Applications 9

Using the Wrapper Compilers 10

C++ Specific Issues 10

Fortran-Specific Issues 11

Using MPI_Accumulate on User-Defined Types 12

Using Non-Default Error Handlers 13

Error Codes 13

More Information About the Wrapper Compilers and About MCA Parameters 13

3. Migrating Third-Party Job Launchers 15

How Sun Grid Engine Works With Open MPI 15

Using Sun Grid Engine With ORTE 16

MCA Parameters 16

Submitting Jobs Under Sun Grid Engine Integration 17

More Information About Sun Grid Engine, the Open Source Grid Engine, and MCA Parameters 17

Preface

This document is designed to aid Sun HPC ClusterTools 6 users in the process of migrating their programs to work with Sun HPC ClusterTools 7.1 software.

This document is designed to aid Sun HPC ClusterTools 6 users in the process of migrating their programs to work with Sun HPC ClusterTools 7.1 software. This manual describes the differences between the two releases, and provides helpful information about migrating programs.

Before You Read This Document

The *Sun HPC ClusterTools™ 7 Software Release Notes* includes release note information for the other components in this suite. For information about writing MPI programs, refer to the Open MPI web site at:

`http://www.open-mpi.org`

For information about a specific distributed resource management package, refer to the documentation supplied with that package.

Using UNIX Commands

This document might not contain information about basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices. Refer to the following for this information:

- Software documentation that you received with your system

- Solaris™ Operating System documentation, which is at:
  `http://www.sun.com/documentation`

Shell Prompts

Shell	Prompt
C shell	*machine-name*`%`
C shell superuser	*machine-name*`#`
Bourne shell and Korn shell	`$`
Bourne shell and Korn shell superuser	`#`

Typographic Conventions

Typeface*	Meaning	Examples
`AaBbCc123`	The names of commands, files, and directories; on-screen computer output	Edit your `.login` file. Use `ls -a` to list all files. `% You have mail.`
`AaBbCc123`	What you type, when contrasted with on-screen computer output	`%` **`su`** `Password:`
AaBbCc123	Book titles, new words or terms, words to be emphasized. Replace command-line variables with real names or values.	Read Chapter 6 in the *User's Guide*. These are called *class* options. You *must* be superuser to do this. To delete a file, type `rm` *filename*.

* The settings on your browser might differ from these settings.

Related Documentation

This book focuses on Sun MPI and assumes familiarity with the *MPI Standard*. The following materials provide useful background about using Sun MPI and about the *MPI Standard*.

Application	Title	Part Number
Sun HPC ClusterTools Documentation	*Read Me First: Guide to Sun HPC ClusterTools Software Documentation*	819-7477-10
Sun HPC ClusterTools Software	*Sun HPC ClusterTools 7.1 Software Release Notes*	819-7478-10
	Sun HPC ClusterTools 7.1 Software Installation Guide	819-7479-10
	Sun HPC ClusterTools 7.1 Software User's Guide	819-7480=10

For more information about Open MPI, see the Open MPI Web site at:

`http://www.open-mpi.org`

For more information about Sun Grid Engine software, see the Sun Grid Engine web site at:

`http://www.sun.com/software/gridware`

Documentation, Support, and Training

Sun Function	URL
Documentation	http://www.sun.com/documentation/
Support	http://www.sun.com/support/
Training	http://www.sun.com/training/

Third-Party Web Sites

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 by or in connection with the use of or reliance on any such content, goods, or services that are available on or through such sites or resources.

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. You can submit your comments by going to:

http://www.sun.com/hwdocs/feedback

Please include the title and part number of your document with your feedback:

Sun HPC ClusterTools 7.1 Software Migration Guide, part number 819-7484-11

CHAPTER 1

Differences Between Sun HPC ClusterTools Versions

This chapter outlines some of the differences between Sun HPC ClusterTools 6 and Sun HPC ClusterTools 7.1 software.

Sun HPC ClusterTools 6 and previous versions used Sun's own implementation of the MPI-2 standard. However, Sun HPC ClusterTools 7.1 is based on Open MPI 1.2 (OMPI), an open-source implementation of MPI. There are some differences between Sun MPI and Open MPI; you might need to update applications that use Sun MPI functions to run with Open MPI and Sun HPC ClusterTools 7.1.

The *Sun HPC ClusterTools 7.1 Software User's Guide* contains information about the Open MPI `mpirun` command and its options.

For more information about Open MPI, see the Open MPI web site at:

`http://www.open-mpi.org`

Comparing `mprun` Options to `mpirun` Options

In Sun HPC ClusterTools 6 and previous versions, the command to run MPI was `mprun`. In Sun HPC ClusterTools 7.1, this command is called `mpirun`. The `mpirun`(1) man page contains information about the usage of `mpirun` and its options.

For a complete listing of all Sun HPC ClusterTools 7.1 `mpirun` options and a sample usage message, type the following command:

```
% mpirun --help
```

The following table shows a mapping between Sun HPC ClusterTools 6 options and some of the options in Sun HPC ClusterTools 7.1.

Note – Although it appears that Sun HPC ClusterTools 7.1 does not possess options equivalent to those in Sun HPC ClusterTools 6 software, Sun HPC ClusterTools 7.1 contains many additional options that do not have equivalents in Sun HPC ClusterTools 6. For more information about the options in Sun HPC ClusterTools 7.1, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

TABLE 1-1 Comparison of `mprun` Options and `mpirun` Options

Description	Sun HPC ClusterTools 6/Sun MPI (`mprun`)	Sun HPC ClusterTools 7.1/Open MPI (`mpirun`)
Displays this help/usage text	`-h`	`-h`
Displays tool version information	`-V`	
Specifies the cluster to use	`-c [cluster]`	
Specifies the partition to use	`-p [partition]`	
Specify the argv [0] explicitly	`-A [aout]`	
Specify uid to execute as	`-U [uid]`	
Specify gid to execute as	`-G [gid]`	
Specify the I/O fd set to multiplex	`-I [iofds]`	
Specify CRE I/O (use with -x)	`-Is`	
Specify an alternate working directory	`-C [path]`	`-wdir --wdir`
Specify a project name	`-P [project]`	
Chroot to working dir before execution	`-r [path]`	
Show job id after exec	`-J`	
Specify the number of processes/threads in job	`-np [PxT]`	`-np [P] (P != 0)`
Specify the number of processes/threads to reserve	`-nr [PxT]`	
Specify Resource Requirement String	`-R [rrs]`	

TABLE 1-1 Comparison of `mprun` Options and `mpirun` Options

Description	Sun HPC ClusterTools 6/Sun MPI (`mprun`)	Sun HPC ClusterTools 7.1/Open MPI (`mpirun`)
Allow wrapping of hosts	`-W`	
Settle for available hosts	`-S`	
Run this job on same resources as [job name]	`-j [job name]`	
Only rank 0 gets stdin	`-i`	
Rank-tag stdout	`-o`	
Separate stdout/stderr streams	`-D`	
No stdio connections	`-N`	
Batch stream handling	`-B`	
No stdin connection	`-n`	
No spawning on SMPs	`-Ns`	
Enable spawning on SMPs	`-Ys`	
Group procs [n] to an SMP	`-Z [n]`	
Group/tile procs [n] to an SMP	`-Zt [n]`	
Specify rankmap string	`-l "[host] procs?[,...] "`	`-host --host -H`
Specify rankmap file	`-m [file]`	`-hostfile [file], --hostfile [file], -machinefile [file], --machinefile [file]`
Use any partition independent nodes	`-u`	
Multiply daemon and mprun timeouts by factor n	`-t [n]`	
Dump JID to a file	`-d [filename]`	

TABLE 1-1 Comparison of `mprun` Options and `mpirun` Options

Description	Sun HPC ClusterTools 6/Sun MPI (`mprun`)	Sun HPC ClusterTools 7.1/Open MPI (`mpirun`)
Verbose. Gives extra information during job startup.	`-v`	`-d`
Run processes under control of resource manager RM	`-x [RM]`	`[not needed]`
Multiple executables	`-np 8 : -np 2 exe1 : -np 6 exe2`	`-np 2 exe1 : -np 6 exe2`

For a complete listing of `mpirun` options and instructions on how to use them, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

Passing MPI Options

Both Sun MPI and Open MPI use environment variables to pass options to MPI. However, Open MPI also allows you to pass options in two additional ways:

- Using the `-mca` flag with `mpirun`
- Defining a configuration file with the name `$HOME/.openmpi/mca-params.conf`.

The *Sun HPC ClusterTools 7.1 Software User's Guide* describes Modular Component Architecture (MCA) parameters in detail and lists the ways in which they can be used with `mpirun`. For more information about using MCA parameters with Sun Grid Engine software, see "MCA Parameters" on page 16.

For more information about the available options that can be used with `-mca` or in the `mca-params.conf` file, see the Open MPI runtime tuning Frequently Asked Questions (FAQ) at:

`http://www.open-mpi.org/faq/?category=tuning`

MPI Environment Variables

Many of the environment variables that existed in Sun HPC ClusterTools 6 are not supported in Sun HPC ClusterTools 7.1. However, the MCA parameters in Sun HPC ClusterTools 7.1 supply much of the equivalent functionality.

For more information about the MCA parameters and their usage, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*. For additional information from the command line, see the `mpirun` man page, or type the following command at the system prompt:

```
% ompi_info -param all all
```

Specifying Selected CPUs for Applications

Sun HPC ClusterTools 6 uses a centralized database in the `tm.rdb` process to store the names of selected CPUs on which to run applications. However, Sun HPC ClusterTools 7.1 does not use a centralized database. You specify the CPUs on which to run applications in two ways:

- Specify the CPUs on which to run on the command line
- Create a hostfile listing the CPUs

For more information, see Chapter 4 in the *Sun HPC ClusterTools 7.1 Software User's Guide*.

Thread Safety

Sun HPC ClusterTools 6 software included both non-thread-safe (`libmpi.so`) and thread-safe (`libmpi_mt.so`) MPI libraries. Sun HPC ClusterTools 7.1 includes only a non-thread-safe version of the library. Thread-safe libraries are not supplied in Sun HPC ClusterTools 7.1.

For more information, refer to the *Sun HPC ClusterTools 7.1 Software Release Notes*.

Sun HPC ClusterTools-Specific Commands

Some commands and functionality in Sun HPC ClusterTools 6/Sun MPI have no equivalents in Sun HPC ClusterTools 7.1/Open MPI.

The following tools in Sun HPC ClusterTools 6 have no equivalents in Open MPI:

- `mpprof`
- `mpps`
- `mpinfo`
- `mpkill`

These four tools made use of root daemons that existed in ClusterTools 6. Sun HPC ClusterTools 7.1 does not have root daemons; therefore, the software currently does not have equivalent tools.

In place of `mpprof`, you might use the Solaris DTrace dynamic tracing utility, the PERUSE interface in Open MPI, or some other tracing tool. For more information about using DTrace with Sun HPC ClusterTools 7.1 software, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

For more information on the PERUSE interface, refer to the following paper on the Open MPI Web site:

`http://www.open-mpi.org/papers/euro-pvmmpi-2006-peruse/`

In place of `mpkill`, you might need to kill the processes on each node manually, or kill the processes with a batch system, such as Sun Grid Engine. You can also use Ctrl-C to kill each process, if needed.

In place of `mpps` and `mpinfo`, use your batch system to see the parallel job.

Because there are no root daemons, the activation step in Sun HPC ClusterTools 7.1 is optional, as it only creates symbolic links to the installed files. For more information about the differences in installation and activation steps, refer to the *Sun HPC ClusterTools 7.1 Software Installation Guide*.

In addition, Sun HPC ClusterTools 6 has the ability to do dynamic allocation of processors based on resource constraints, such as CPU load or memory usage. ORTE (the Open MPI Run-Time Environment) has no equivalent functionality. It also cannot determine the number of CPUs on a node.

For more information, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

Updating Applications Compiled With Previous Versions of Sun HPC ClusterTools Software

You must recompile applications created using Sun HPC ClusterTools 6 software (or an earlier version) before you can use that application with Sun HPC ClusterTools 7.1.

For more information about updating applications, see Chapter 2, "Migrating Sun HPC ClusterTools 6 Applications" on page 9.

Note – In most cases, you should not experience issues when migrating your applications to Sun HPC ClusterTools 7.1 software. The instructions in Chapter 2 of this manual describe what to do if you do run into a problem.

Comparing Sun MPI Environment Variables to Open MPI MCA Parameters

Sun HPC ClusterTools 7.1/Open MPI employs Open MPI MCA parameters to serve a similar function to the environment variables in Sun HPC ClusterTools 6/Sun MPI. The *Sun HPC ClusterTools 7.1 User's Guide* contains information about the available MCA parameters for Solaris OS installations, as well as instructions on how to use them.

The following table shows some equivalent functions between the two ClusterTools versions:

TABLE 1-2 Sun MPI and Open MPI Equivalents

Sun MPI Environment Variable	Open MPI MCA Parameter
MPI_FULLCONNINIT	mpi_preconnect_all
MPI_PRINTENV	mpi_show_mca_params
MPI_CHECK_ARGS	mpi_param_check
MPI_SPIN	mpi_yield_when_idle
MPI_PROC_BIND	mpi_paffinity_alone

More Information About Parameters and Compiling Applications

For more information about MCA parameters, refer to the following:

- *Sun HPC ClusterTools 7.1 Software User's Guide* (819-7480)

- *Sun HPC ClusterTools 7.1 Software Release Notes* (819-7478)
- Open MPI FAQ

  `http://www.open-mpi.org/faq/?category=tuning`

For more information about compiling applications with Sun HPC ClusterTools 7.1 software, see Chapter 2, "Migrating Sun HPC ClusterTools 6 Applications" on page 9.

CHAPTER 2

Migrating Sun HPC ClusterTools 6 Applications

This chapter describes changes you might need to make in your application code to recompile and run programs developed with Sun HPC ClusterTools 6 software (or a previous version of Sun HPC ClusterTools) in Sun HPC ClusterTools 7.1.

This chapter contains the following topics:

- "Using the Wrapper Compilers" on page 10
- "C++ Specific Issues" on page 10
- "Fortran-Specific Issues" on page 11
- "Using MPI_Accumulate on User-Defined Types" on page 12
- "Using Non-Default Error Handlers" on page 13
- "Error Codes" on page 13

Note – In most circumstances, the differences between Sun HPC ClusterTools 6 and Sun HPC ClusterTools 7.1 are relatively minor. If you do run into difficulty using your compiled application with Sun HPC ClusterTools 7.1 software, then the information in this chapter can help you troubleshoot the problem.

It is suggested that you recompile your applications using the `mpicc` compiler if you want them to be compatible with Sun HPC ClusterTools 7.1 software. The `tmcc` compiler is no longer supported in Sun HPC ClusterTools 7.1, and there is no backward compatibility.

Using the Wrapper Compilers

Sun HPC ClusterTools 7.1 supplies wrapper compilers for you to use instead of directly calling the compilers when compiling applications for use with the Sun HPC ClusterTools 7.1 software. These wrapper compilers do not actually perform the compilation and linking steps themselves, but they add the appropriate compiler and linker flags and call the compiler and linker.

Note – Using the wrapper compilers is strongly suggested. If you decide not to use them, the Open MPI Web site at `http://www.open-mpi.org` contains instructions about how to compile without using them.

The following wrapper compilers are available:

TABLE 2-1 Wrapper Compilers

Language	Wrapper Compiler
C	`mpicc`
C++	`mpiCC`, `mpicxx`, or `mpic++` (Note: `mpiCC` is for use on case-sensitive file systems only)
Fortran 77	`mpif77`
Fortran 90	`mpif90`

For more information about the wrapper compilers, their use, and troubleshooting, see the Open MPI FAQ at:

`http://www.open-mpi.org/faq/?category=mpi-apps`

C++ Specific Issues

This section describes issues you should address if you have a C++ application using Sun MPI and a previous version of Sun HPC ClusterTools. You might need to make changes to your code in order for it to work properly with Open MPI and Sun HPC ClusterTools 7.1.

`mpi++.h` Header File

Sun MPI supports the `mpi++.h` header file for C++ programs; however, there is no `mpi++.h` file in Open MPI. If your application has a C++ test for Sun MPI, you must change the header file from `mpi++.h` to `mpi.h` for it to work with Open MPI.

Function and Macro Definitions

The MPI-2 standard requires an MPI implementation to supply variants of the functions/macros NULL_COPY_FN, NULL_DUP_FN, and NULL_DELETE_FN that can be "specified from either C, C++, or Fortran." For C++ in Sun MPI, the definitions for these functions and macros are slightly different than they are for C and Fortran.

If your applications contain calls to any of the Sun MPI functions/macros listed in TABLE 2-2, you must change the function names to their Open MPI equivalents before your applications will work with Sun HPC ClusterTools 7.1. The following table lists the Sun MPI functions and their Open MPI equivalents.

TABLE 2-2 Function/Macro Equivalents for C++ Programs

Sun MPI Function/Macro	Open MPI Equivalent
MPI::Comm::COMM_NULL_COPY_FN	MPI_COMM_NULL_COPY_FN
MPI::Comm::COMM_DUP_FN	MPI_COMM_DUP_FN
MPI::Comm::COMM_NULL_DELETE_FN	MPI_COMM_NULL_DELETE_FN
MPI::Datatype::TYPE_NULL_COPY_FN	MPI_TYPE_NULL_COPY_FN
MPI::Datatype::TYPE_DUP_FN	MPI_TYPE_DUP_FN
MPI::Datatype::TYPE_NULL_DELETE_FN	MPI_TYPE_NULL_DELETE_FN
MPI::Win::WIN_NULL_COPY_FN	MPI_WIN_NULL_COPY_FN
MPI::Win::WIN_DUP_FN	MPI_WIN_DUP_FN
MPI::Win::WIN_NULL_DELETE_FN	MPI_WIN_NULL_DELETE_FN

Fortran-Specific Issues

This section lists issues you should address if you have a Fortran application using Sun MPI and a previous version of Sun HPC ClusterTools. You might need to make changes to your code in order for it to work properly with Open MPI and Sun HPC ClusterTools 7.1.

f90 Interfaces

When compiling code for which you need the f90 interfaces, either use the `mpif90` wrapper or add the following to your link line:

```
-lmpi_f90
```

Either the `mpif90` wrapper or the `-lmpif90` flag enables the application to compile correctly.

f90 Module Location

The f90 module `mpi.mod` has moved from `/opt/SUNWhpc/include` to `/opt/SUNWhpc/lib`. This should not affect your program if you make use of the compiler wrapper `mpif95`.

Datatypes

Four Fortran datatypes existed in previous versions of Sun HPC ClusterTools, but do not exist in Open MPI or Sun HPC ClusterTools 7.1. They are not defined in the MPI-2 standard. These datatypes are as follows:

- MPI_LOGICAL1
- MPI_LOGICAL2
- MPI_LOGICAL4
- MPI_LOGICAL8

Using MPI_Accumulate on User-Defined Types

Open MPI currently does not support MPI_Accumulate on user-defined types. If you try to use MPI_Accumulate in a program with a user-defined type, the following error message is displayed:

```
MPI_Accumulate currently does not support reductions with any
user-defined types. This will be rectified in a future release.
```

Using Non-Default Error Handlers

In Sun MPI, a non-default error handler would persist past the call to MPI_Finalize. Therefore, if you called an MPI call after MPI_Finalize, the non-default error handler would be used. In Open MPI, the non-default error handler does not persist, and the default error handler is used. This causes any call used after MPI_Finalize to be aborted.

Error Codes

The following table lists error codes that are unique to Sun HPC ClusterTools 6. Only one of these error codes has an Open MPI equivalent in Sun HPC ClusterTools 7.1 software.

TABLE 2-3 Error Codes Not Defined in the MPI-2 Standard

Sun HPC ClusterTools 6 Error Code	Open MPI Equivalent	Error Class
MPI_ERR_TIMEDOUT	None	MPI_ERR_TIMEDOUT
MPI_ERR_RESOURCES	MPI_ERR_SYSRESOURCES	MPI_ERR_RESOURCES
MPI_ERR_TRANSPORT	None	MPI_ERR_TRANSPORT
MPI_ERR_HANDSHAKE	None	MPI_ERR_HANDSHAKE

More Information About the Wrapper Compilers and About MCA Parameters

For more information about the wrapper compilers, their use, and troubleshooting, see the Open MPI FAQ at:

`http://www.open-mpi.org/faq/?category=mpi-apps`

For more information about MCA parameters, refer to the following:

- *Sun HPC ClusterTools 7.1 Software User's Guide* (819-7480)
- *Sun HPC ClusterTools 7.1 Software Release Notes* (819-7478)

- Open MPI FAQ

  http://www.open-mpi.org/faq/?category=tuning

CHAPTER 3

Migrating Third-Party Job Launchers

Sun HPC ClusterTools 6 and previous versions provided support for use with third-party job launchers, including Sun Grid Engine. The current version of SGE is now known as Sun Grid Engine. Sun HPC ClusterTools 7.1 supports Sun Grid Engine, in conjunction with ORTE (Open Run-Time Environment).

ORTE is the runtime support system for Open MPI that allows users to execute their applications in a distributed clustering environment.

Note – Sun HPC ClusterTools 6 supported the LSF job-launching software. LSF is not supported in Sun HPC ClusterTools 7.1.

How Sun Grid Engine Works With Open MPI

HPC sites use batch systems to share resources fairly and accountably, and also to guarantee that a job can obtain the resources it needs to run at maximum efficiency. To properly monitor a job's resource consumption, the batch system must be the agent that launches the job.

Sun Grid Engine, like many other batch systems, cannot launch multiple process jobs (such as MPI applications) on its own. In Sun HPC ClusterTools 7.1, ORTE launches the multiple process jobs and sets up the environment required by Open MPI.

When Sun Grid Engine launches a parallel job in cooperation with OpenRTE, Sun Grid Engine "owns" the resulting launched processes. Sun Grid Engine monitors the resources for these processes, thereby creating a tightly integrated environment for resource accounting. OpenRTE allows users to execute their parallel applications.

Note – There is also an open source version of Grid Engine (GE) hosted on http://www.sunsource.net. Although the Sun HPC ClusterTools 7.1/Open MPI integration is developed with Sun Grid Engine, this integration should work for the open source Grid Engine as well.

Using Sun Grid Engine With ORTE

Open MPI requires that its executables are in your PATH on every node that will run them.

Note – `mpirun` was called `mprun` in previous versions of Sun HPC ClusterTools.

For example, if Open MPI was installed into `/opt/SUNWhpc/HPC7.1/bin`, you would add this location to your PATH as shown in the following example:

```
% setenv PATH /opt/SUNWhpc/HPC7.1/bin:${PATH}
% mpirun -np 4 a.out
```

The `setenv` command will prefix the PATH on both the local and remote hosts with `/opt/SUNWhpc/HPC7.1/bin`.

For more information about the `mpirun` command, its usage, and its options, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

MCA Parameters

Open MPI provides MCA (Modular Component Architecture) parameters for use with the `mpirun` command. These parameters direct `mpirun` to perform the specified functions under Sun Grid Engine. To specify an MCA parameter, use the `-mca` flag with the `mpirun` command.

Note – In most cases, you do not need to change the values set in the `gridengine` MCA parameters. If you run into difficulty and want to change the values for debugging purposes, the option is available.

For more information about MCA parameters, refer to the *Sun HPC ClusterTools 7.1 Software User's Guide*.

Submitting Jobs Under Sun Grid Engine Integration

To submit jobs under Sun Grid Engine integration in Sun HPC ClusterTools 7.1, you must first create a Sun Grid Engine (SGE) environment using `qsub`, `qsh`, and so on. Instructions about how to set up the parallel environment (PE) and queue in Sun Grid Engine are described in the *Sun HPC ClusterTools 7.1 Software User's Guide.*

There are two ways to submit jobs under Sun Grid Engine integration: interactive mode and batch mode. These procedures are also described in the *Sun HPC ClusterTools 7.1 Software User's Guide*.

In previous versions of Sun HPC ClusterTools software, integration with Sun Grid Engine required you to set an `mprun.conf` file and to use the `mprun -x sge` option. When using Sun HPC ClusterTools 7.1, neither of these steps is necessary.

More Information About Sun Grid Engine, the Open Source Grid Engine, and MCA Parameters

For more information about Sun Grid Engine, refer to the following:

- Sun Grid Engine product page:

  `http://www.sun.com/software/gridware`
- *Sun Grid Engine 6 Administration Guide*

  `http://docs.sun.com/app/docs/doc/817-5677`
- *Sun Grid Engine 6 Installation Guide*

  `http://docs.sun.com/app/docs/doc/817-6118`
- *Sun Grid Engine 6 User's Guide*

http://docs.sun.com/app/docs/doc/817-6117

- Grid Engine project on SunSource:

  http://gridengine.sunsource.net

For more information about MCA parameters, refer to the following:

- *Sun HPC ClusterTools 7.1 Software User's Guide* (819-7480)
- *Sun HPC ClusterTools 7.1 Software Release Notes* (819-7478)
- Open MPI FAQ

  http://www.open-mpi.org/faq/?category=tuning

Index

A

activation, 6
applications
  migrating, 9
  recompiling, 6, 9
  updating, 1

B

batch systems, 15

C

C++ programs, 10
compilation, 10
compilers
  mpicc, 10
  mpicc compiler, 9
  mpiCC, mpicxx, mpic++, 10
  mpif77, 10
  mpif90, 10
  wrapper compilers, 10
compiling
  using the wrapper compilers, 10
CPUs
  specifying for applications, 5

D

documentation
  Open MPI FAQ, v
  product notes, v

E

Environment variables, 4, 7
error handlers, 13

F

Fortran
  datatypes, 12
Fortran applications, 11
functions and macros, 11

G

Grid Engine
  open source version, 16

H

hostfile, 5

I

include file, 11

J

job launchers, 15

L

linking, 10

M

macro definitions
  for C++, 11
MCA parameters, 4, 7, 16
  compared to environment variables, 4
  gridengine parameters, 17
mca-params.conf file, 4

MPI options
passing options, 4
MPI programs
writing, v
MPI Standard, vii, 1, 11
mpi++.h file, 11
MPI_Accumulate, 12
MPI_Finalize, 13
MPI-2, 1
mpicc compiler, 9
mpif77, 10
mpif90, 10
mpirun, 1
comparison with mprun, 1
-d option, 4
-h option, 2
--help option, 2
-host option, 3
-hostfile option, 3
-mca option, 4, 16
multiple executables, 4
-np option, 2
options, 2
-wdir option, 2

O

ompi_info command, 5
Open MPI, vii, 1, 5
and Sun N1 Grid Engine, 15
functions, 11
ORTE, 6, 15

P

parallel environment (PE), 17
Passing MPI options, 4
PATH environment variable, 16

R

recompiling applications, 6
run-time tuning, 4

S

submitting jobs
under Sun N1 Grid Engine, 17
Sun HPC ClusterTools 6
commands, 5
comparison with mpirun options, 2
error codes, 13
Fortran datatypes, 12
migrating applications, 9
recompiling applications, 6
updating applications, 1
Sun N1 Grid Engine, vii, 15
and Open MPI, 15
gridengine MCA parameters, 17

T

Thread safety, 5
tmcc compiler, 9

U

updating applications, 1

W

wrapper compilers, 12
Writing MPI programs, v