

Oracle Parallel Server™  
Administrator's Guide  
for Silicon Graphics® Systems

Document Number 007-2763-003

## CONTRIBUTORS

Written by Susan Ellis and Carolyn Curtis

Illustrated by Dan Young and Carolyn Curtis

Edited by Christina Cary

Production by Linda Rae Sande

St Peter's Basilica image courtesy of ENEL SpA and InfoByte SpA. Disk Thrower image courtesy of Xavier Berenguer, Animatica.

© 1997, Silicon Graphics, Inc.— All Rights Reserved

The contents of this document may not be copied or duplicated in any form, in whole or in part, without the prior written permission of Silicon Graphics, Inc.

## RESTRICTED RIGHTS LEGEND

Use, duplication, or disclosure of the technical data contained in this document by the Government is subject to restrictions as set forth in subdivision (c) (1) (ii) of the Rights in Technical Data and Computer Software clause at DFARS 52.227-7013 and/or in similar or successor clauses in the FAR, or in the DOD or NASA FAR Supplement. Unpublished rights reserved under the Copyright Laws of the United States. Contractor/manufacturer is Silicon Graphics, Inc., 2011 N. Shoreline Blvd., Mountain View, CA 94043-1389.

Silicon Graphics, the Silicon Graphics logo, CHALLENGE, and Indy are registered trademarks and IRISconsole, IRIS InSight, IRIX, IRIX NetWorker, O2, Origin, Origin2000, Performance Co-Pilot, and XFS are trademarks of Silicon Graphics, Inc. Oracle is a registered trademark and Oracle Parallel Server is a trademark of Oracle Corporation. UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd.

## FCC WARNING

This equipment has been tested and found compliant with the limits for a Class A digital device, pursuant to Part 15 of the FCC rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference in which case the user will be required to correct the interference at his own expense.

**ATTENTION**

This product requires the use of external shielded cables in order to maintain compliance pursuant to Part 15 of the FCC Rules.

**VDE 0871/6.78**

This equipment has been tested to and is in compliance with the Level A limits per VDE 0871.

**EUROPEAN UNION STATEMENT**

This device complies with the European Directives listed on the "Declaration of Conformity" which is included with each product. The CE mark insignia displayed on the device is an indication of conformity to the aforementioned European requirements.

**INTERNATIONAL SPECIAL COMMITTEE ON RADIO INTERFERENCE (CISPR)**

This equipment has been tested to and is in compliance with the Class A limits per CISPR publication 22.

**CANADIAN DEPARTMENT OF COMMUNICATIONS STATEMENT**

This digital apparatus does not exceed the Class A limits for radio noise emissions from digital apparatus as set out in the Radio Interference Regulations of the Canadian Department of Communications.

**ATTENTION**

Cet appareil numérique n'émet pas de perturbations radioélectriques dépassant les normes applicables aux appareils numériques de Classe A prescrites dans le Règlement sur les interférences radioélectriques établi par le Ministère des Communications du Canada.

## VCCI CLASS A STATEMENT FOR JAPAN

この装置は、第一種情報装置（商工業地域において使用されるべき情報装置）で商工業地域での電波障害防止を目的とした情報処理装置等電波障害自主規制協議会(VCCI)基準に適合しております。

従って、住宅地域またはその隣接した地域で使用すると、ラジオ、テレビジョン受信機等に受信障害を与えることがあります。

取扱説明書に従って正しい取り扱いをして下さい。

---

# Contents

**List of Figures** vii

**About This Guide** ix

Audience ix

Structure of This Document ix

Conventions x

1. **Features of the Oracle Parallel Server** 1
  - OPS Hardware and Software Requirements 1
  - OPS Configuration 2
  - OPS Instances and Domains 4
  - OPS Architectural Features 7
  - OPS Components and How They Work Together 8
    - Role of IRISconsole 10
    - Role of the CHALLENGE RAID Storage System 11
    - Role of the XFS Filesystem 12
    - Role of DBA (Database Accelerator) 13
  - Optional Silicon Graphics Software and OPS Performance 13
    - Performance Co-Pilot (PCP) 13
    - IRIX NetWorker 14
2. **Configuring, Starting, and Halting the OPS Software** 15
  - Checking OPS Software Configuration 15
    - Checking the opsconf Configuration File 16
    - Checking the sidconf Configuration File 17
    - Checking the nconf Configuration File 17
  - Creating System Controller Password Files 18
  - Starting the Node Controller Software on the Indy or O2 Workstation 18
  - Starting the OPS Software Manually on the Servers 18

Starting the OPS Software Automatically on the Servers	19
Starting a Server for Single-Host Operation	20
Halting the OPS Software on the Servers	20
Halting the OPS Software Manually	20
Checking Log Files	21

---

## List of Figures

<b>Figure 1-1</b>	OPS Hardware Configuration	3
<b>Figure 1-2</b>	OPS Software Configuration	4
<b>Figure 1-3</b>	One Domain on an OPS Configuration	5
<b>Figure 1-4</b>	OPS Configuration With Two Domains	6
<b>Figure 1-5</b>	Example OPS Site With Two Vaults and One CHALLENGE RAID Storage System	9
<b>Figure 1-6</b>	SCSI-2 Bus and Internal Buses (Front View)	12



---

## About This Guide

The Oracle Parallel Server™ (OPS) system represents a significant development in RDBMS capability for Oracle® users. The ability to cluster two computer systems and coordinate their access to a single shared Oracle database provides for increased throughput and capacity with higher levels of application and database availability.

This guide was prepared in conjunction with Release 1.2 of OPS software for Silicon Graphics systems.

### Audience

This guide is written for the person who administers the OPS system. The OPS administrator is familiar with Oracle RDBMS in general and the specific Oracle database instances running on the CHALLENGE® or Origin™ servers. The OPS administrator is also familiar with the operation of the workstation and the CHALLENGE or Origin servers, as well as the optional Vault storage systems and CHALLENGE RAID storage system, if they are used in the OPS configuration. The OPS administrator has acquired familiarity with the Oracle Parallel Server system and uses IRISconsole™ to control the OPS systems.

### Structure of This Document

This guide contains the following chapters and appendix:

- Chapter 1, “Features of the Oracle Parallel Server,” explains capabilities of the Oracle Parallel Server and how system components work together.
- Chapter 2, “Configuring, Starting, and Halting the OPS Software,” explains how to configure and start the OPS system software and how to halt the OPS system.

## Conventions

These type conventions and symbols are used in this guide:

*Italics*                Filenames, IRIX™ commands, manual or book titles, variable command line arguments, and variables to be supplied by the user in examples, code, and syntax statements

Fixed-width type        Error messages, prompts, and onscreen text

**Bold fixed-width type**    User input, including keyboard keys (printing and nonprinting); literals supplied by the user in examples, code, and syntax statements

---

## Features of the Oracle Parallel Server

This chapter introduces the features and capabilities of the Oracle Parallel Server (OPS) system. It contains these major sections:

- “OPS Hardware and Software Requirements” on page 1
- “OPS Configuration” on page 2
- “OPS Instances and Domains” on page 4
- “OPS Architectural Features” on page 7
- “OPS Components and How They Work Together” on page 8
- “Optional Silicon Graphics Software and OPS Performance” on page 13

### OPS Hardware and Software Requirements

The Silicon Graphics® OPS system consists of the following hardware components:

- two same-model CHALLENGE servers (except CHALLENGE S) running IRIX 6.2 or two Origin2000™ servers running IRIX 6.4
- one IRISconsole: either an Indy® workstation running IRIX 6.2 or an O2™ workstation running IRIX 6.3, along with an IRISconsole serial port multiplexer, cables, and software
- one CHALLENGE RAID desktide storage system with two storage-control processors (SPs) configured as RAID level 5 (for the databases)
- two plexed CHALLENGE Vaults or one CHALLENGE RAID desktide storage system configured as RAID level 1 (for Oracle REDO logs)
- required hardware upgrades and cables

The required software for and OPS system consists of the follow components:

- IRIX 6.2 with patches on CHALLENGE systems or IRIX 6.4 with patches on Origin2000 systems
- Release 1.2 of OPS software for Silicon Graphics systems on the CHALLENGE or Origin systems
- IRISconsole software on the Indy or O2 workstation
- software for the component systems, such as the CHALLENGE RAID storage system
- Release 7.3.2 or later of the Oracle RDBMS (obtained from Oracle Corporation)
- the Parallel Server Option (obtained from Oracle Corporation)

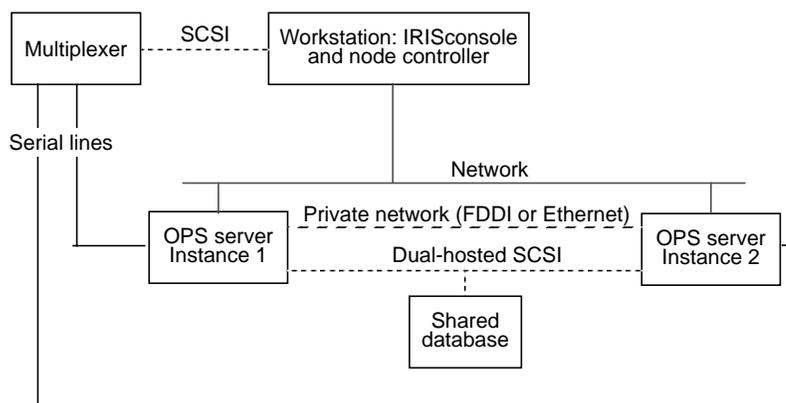
Optional software includes the following:

- Performance Co-Pilot™ (PCP)
- Database Accelerator (DBA)
- IRIX NetWorker™

## OPS Configuration

An OPS system is a collection of Oracle instances running on separate CHALLENGE or Origin servers, providing simultaneous access to the same physical database. The physical database is the same as that for an ordinary non-OPS (nonparallel) Oracle RDBMS, except that it has separate redo logs and rollback segments for each instance. The redo log file is a compressed record of changes that a transaction has made.

For Silicon Graphics systems, the OPS system is available in a dual-host configuration; each server can access the same shared disk storage. The servers can be two CHALLENGE DM, CHALLENGE L, CHALLENGE XL, or Origin2000 systems. A third system, an Indy or O2 workstation running the IRISconsole software, functions as the OPS Node Controller and as a single point of administration for the OPS system. Figure 1-1 diagrams Silicon Graphics' OPS hardware configuration.



**Figure 1-1** OPS Hardware Configuration

Besides the IRISconsole software, the IRISconsole workstation also runs the OPS node control software, *opsnc*. The OPS Node Controller command *opsnc* implements a fail-stop mechanism: in the event of a private network partition (a private network failure that results in the OPS instances being isolated from each other), only one OPS instance is permitted to continue providing service. The other instance is forced to crash and must be restarted by the system administrator.

Each OPS instance consists of the following software components:

- Oracle RDBMS processes, including PMON, SMON, DBWR, LGWR
- OPS Distributed Lock Manager (DLM) processes: *dlnmon* and *dlmd*

An OPS system allows multiple servers access to the same shared physical database. The Distributed Lock Manager has the functionality that enables sharing.

- OPS Connection Manager (CM) process: *opscm*

The OPS Connection Manager *opscm* implements a heartbeat protocol across both servers to detect server and private network failures, monitors local DLM processes to detect lock manager failure, and provides a sync service to coordinate recovery for server failure and reintegration.

Figure 1-2 diagrams OPS software configuration.

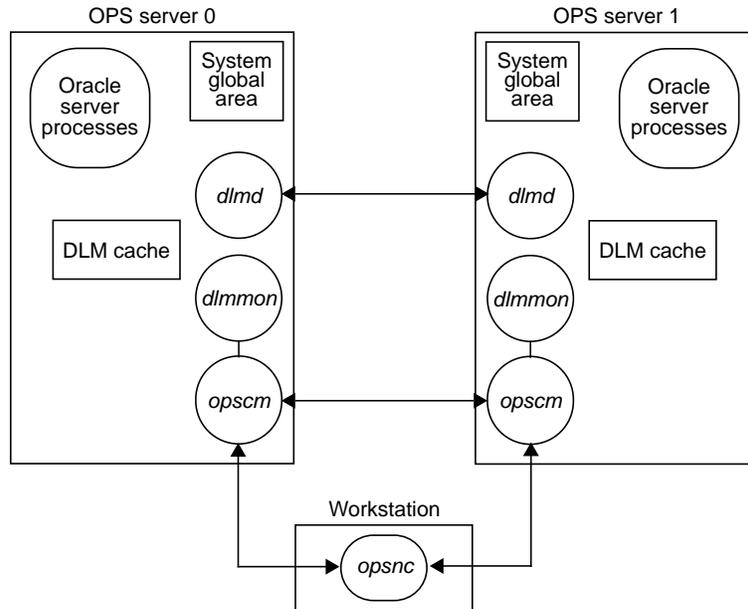
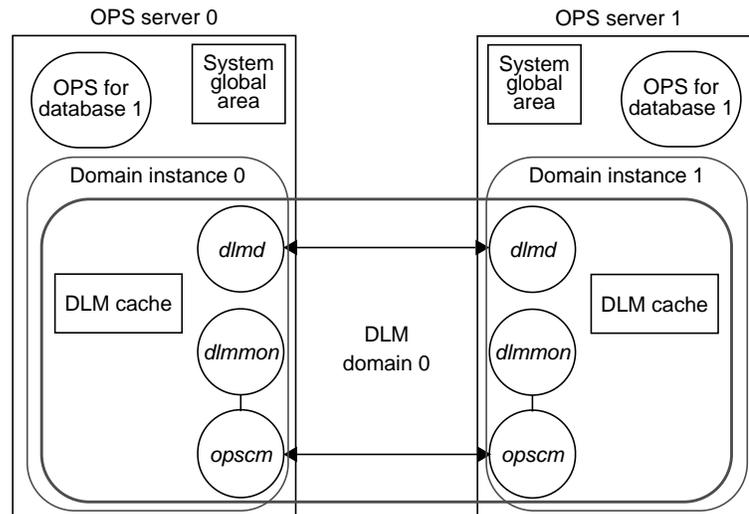


Figure 1-2 OPS Software Configuration

## OPS Instances and Domains

An Oracle instance consists of a *system global area* (SGA) and a set of server processes that access the physical database located on disk. The SGA is a section of shared memory accessed by each of the OPS server processes in an instance. In an OPS system, multiple Oracle instances constitute one *domain*. Each instance has its own SGA, server processes, redo log files, rollback segments, Distributed Lock Manager (DLM) processes, and Connection Manager (CM) process. All instances in the domain access the same physical database.

DLM domains are numbered starting with 0; DLM instances are numbered 0 and 1. Figure 1-3 diagrams an OPS configuration with a single domain.



**Figure 1-3** One Domain on an OPS Configuration

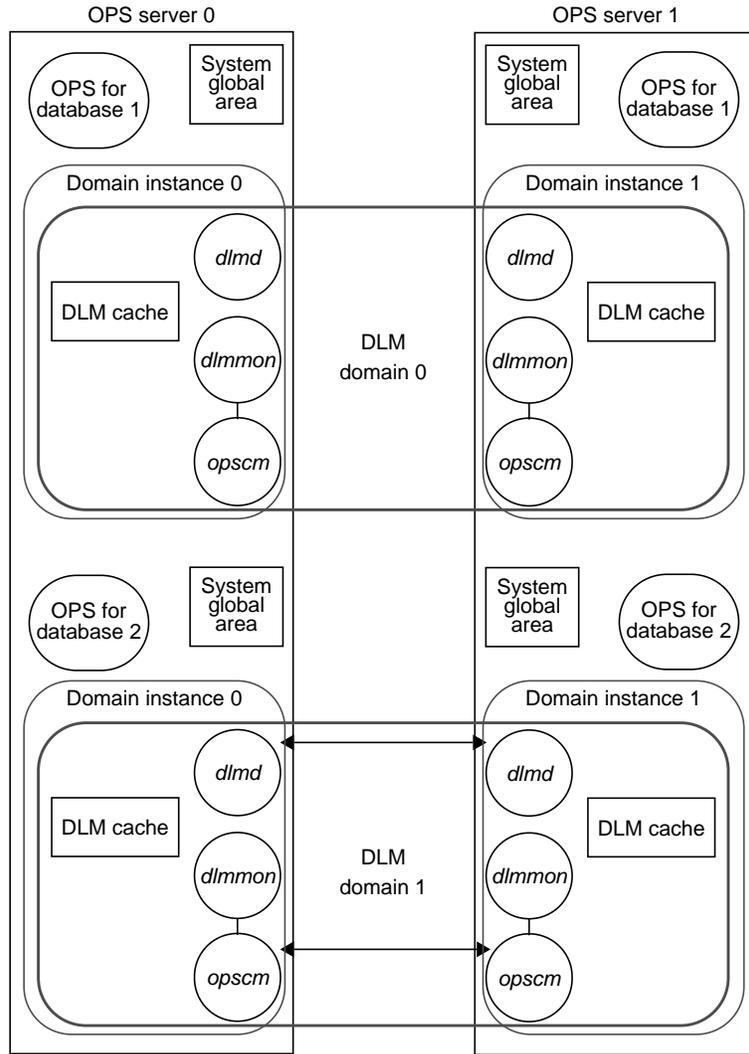
In the configuration shown in Figure 1-3, the instances in the DLM domain are

- 0,0 (DLM domain 0, domain instance 0)
- 0,1 (DLM domain 0, domain instance 1)

Figure 1-4 diagrams an OPS configuration with two domains. In this configuration, the instances in the DLM domain are

- 0,0 (DLM domain 0, domain instance 0)
- 0,1 (DLM domain 0, domain instance 1)
- 1,0 (DLM domain 1, domain instance 0)
- 1,1 (DLM domain 1, domain instance 1)

In any of these cases, all Oracle instances that use the same database must use the same domain.



**Figure 1-4** OPS Configuration With Two Domains

## OPS Architectural Features

The major architectural features of the OPS system are

- high availability

High availability is provided at multiple levels:

- If a server fails, the database is still accessible from the surviving server.
- If the CHALLENGE RAID storage system is used, RAID-5 provides tolerance to any single point of failure within the RAID.
- Each redo log file can be mirrored, so that an instance can survive failure of a log file.

- consolidation of database administration, using the workstation as a node controller

- high performance

An OPS system utilizes the full power of CHALLENGE and Origin system memory and its high-speed system bus performance. Operating system enhancements include changes to virtual memory for more efficient multiprocessing, raw I/Os, multi-process networking, and process scheduling. Besides these enhancements, IRIX already supports real-time scheduling, CPU affinity, and, for the 64-bit operating system, CPU partitioning (the ability to steer interrupts to specific CPUs), which are critical for DBMS performance.

- distributed locks

Row-level locking, the finest level of locking granularity, minimizes the amount of data contention between transactions and maximizes concurrency. Oracle Parallel Server extends this feature by allowing multiple transactions on different servers to lock and update different rows of any table in the database.

Row-level locking is independent of the parallel cache manager's use of distributed lock, which is used to keep the SGAs consistent with each other. Row-level locking is achieved by the Oracle RDBMS internal concurrency control architecture. For distributed locking, the parallel cache manager uses a special background process, the LCNO process, which requests locks from the Distributed Lock Manager (DLM). The DLM is not used for row-level locking; thus its use is minimized and performance is enhanced.

For more information on Oracle RDBMS and OPS operation, consult Oracle documentation.

## **OPS Components and How They Work Together**

The OPS system allows Oracle7 instances running on the two servers to access a common Oracle database. This design allows users on multiple systems seamless access to common data, so that more computing resources are available to all applications that access the same database.

The OPS system is designed to allow any servers in the cluster to be brought down, either voluntarily or involuntarily, without interrupting access to the database from the other servers. If an Oracle instance or a server fails, users from the failed server can migrate to the other running server and reconnect to the database.

Figure 1-5 diagrams an example OPS installation with storage systems.

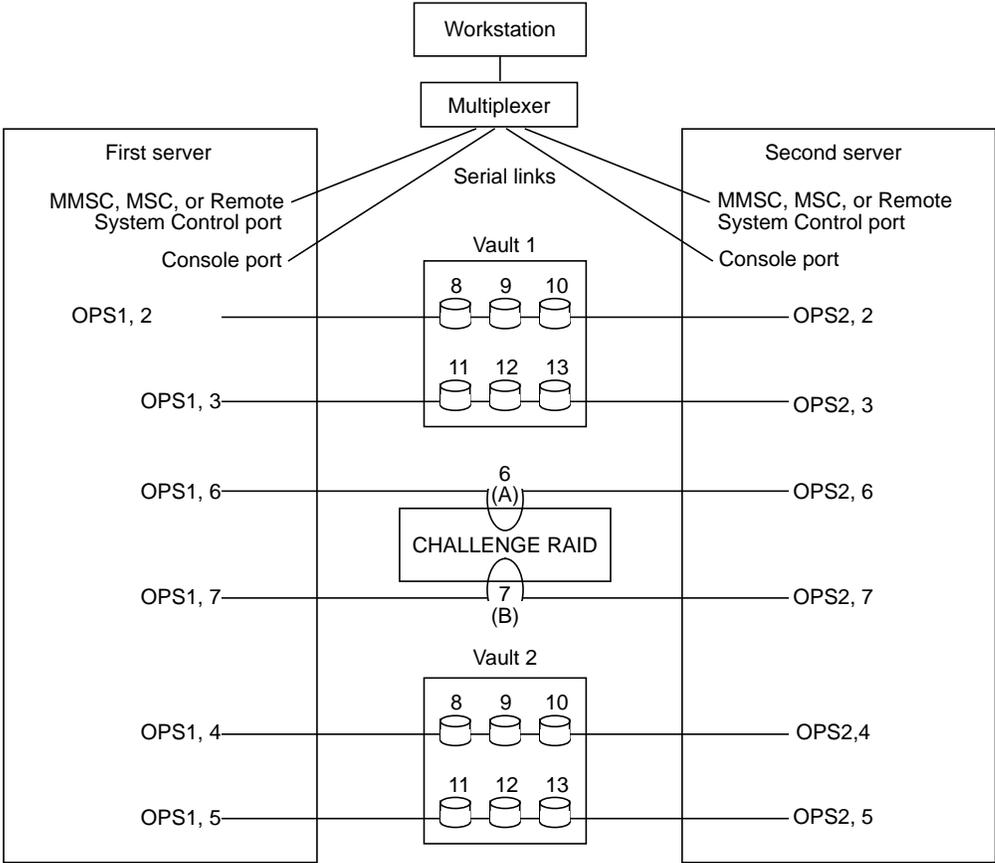


Figure 1-5 Example OPS Site With Two Vaults and One CHALLENGE RAID Storage System

The rest of this section describes specific components of the OPS system:

- IRISconsole
- CHALLENGE RAID storage system
- XFS™ filesystem
- Database Accelerator (DBA)

## Role of IRISconsole

For the OPS system, IRISconsole is made up of the following:

- IRISconsole software, including a graphical user interface, running under IRIX 6.2 (Indy) or IRIX 6.3 (O2)
- an IRISconsole multiplexer, including cabling connecting the workstation to the multiplexer
- a pair of serial cables included in the IRISconsole package, plus one additional pair, for connecting the two OPS servers to the multiplexer

The IRISconsole software monitors each OPS server through the server's Remote System Control and console ports using a serial connection to the serial port server (multiplexer). If a server fails, IRISconsole can automatically start procedures defined by the OPS system administrator in addition to the failover procedures provided for in the OPS software.

**Note:** For full OPS system and IRISconsole functionality, the Remote System Control ports on the Origin2000 MMSC or MSC system controller and the System Console ports on the CHALLENGE DM, L, or XL servers must be cabled to ports on the multiplexer.

The IRISconsole software enables the administrator to

- display, view, or take control of the console of an OPS server (or other attached system)
- view console activity logs and other system reports
- view real-time graphs of hardware operating statistics of an OPS server, such as voltage, operating temperature, and blower speeds; save the graphs as files and display them (for CHALLENGE systems only)
- set a threshold for operating statistics so that an alarm is activated when the threshold is reached and various activities can be triggered (for CHALLENGE systems only)

For complete information on the IRISconsole, see the documentation:

- *IRISconsole Administrator's Guide* (007-2872-*nnn*)
- *IRISconsole Multiplexer Installation Guide* (007-2839-*nnn*)

## **Role of the CHALLENGE RAID Storage System**

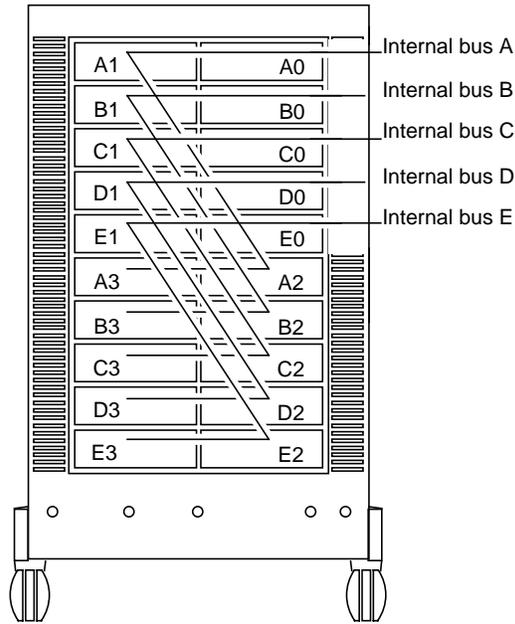
The CHALLENGE RAID (Redundant Array of Independent Disks) storage system provides a compact, high-capacity, high-availability source of disk storage for the OPS system in the form of multiple disk drive modules that you can replace when the storage system is powered on (hot-replaceable modules). Each CHALLENGE RAID storage system supports from five to twenty disk modules.

The CHALLENGE RAID storage system used for databases supports RAID level 5: a group of disk modules is bound together into a logical unit (LUN). A RAID-5 group maintains parity data that lets the disk group survive a disk module failure without losing data. In addition, in a CHALLENGE RAID storage system configured for an OPS system, the RAID-5 group can survive a single SCSI-2 internal bus failure, because each disk module in the group is bound on an independent SCSI-2 internal bus.

Through the storage-control processors (SPs), the SCSI-2 bus is split into five internal fast/narrow SCSI buses—A, B, C, D, and E—that connect the slots for the disk modules. For example, internal bus A connects the modules in slots A0, A1, A2, and A3, in that order. Figure 1-6 diagrams this configuration.

For an OPS system, the CHALLENGE RAID storage system must have two SPs. Each SP controls disk modules in a LUN. The second processor provides a second path to the disk modules as part of the failover strategy of the OPS system; see Figure 1-6. Each LUN is controlled by one of the SPs. The non-controlling SP takes over a LUN if its controlling SP fails.

In addition, both SPs are required for storage system caching to work: each processor temporarily stores modified data in its memory and writes the data to disk at the most expedient time.



**Figure 1-6** SCSI-2 Bus and Internal Buses (Front View)

For complete information on the CHALLENGE RAID storage system, see the *CHALLENGE RAID Owner's Guide (007-2532-*nnn*)*.

### Role of the XFS Filesystem

XFS is a journaled filesystem that allows for extremely fast recovery time of filesystem structures during reboot. Recovery of XFS filesystems is independent of filesystem size. For this reason, XFS is particularly useful for OPS operation.

On a traditional UNIX® filesystem, a full filesystem check takes an amount of time proportional to the size of the filesystem. On XFS, the recovery time is in seconds, because it is dependent upon system activity level, rather than filesystem size. Using XFS reduces the time required to bring a failed server back online.

For complete information on the XFS filesystem, see the guide *IRIX Admin: Disks and Filesystems (007-2825-*nnn*)*. This document is viewable in IRIS InSight™.

## Role of DBA (Database Accelerator)

The Database Accelerator (DBA) consists of kernel enhancements designed to boost performance specifically for Oracle. These kernel enhancements can help double the performance of write-intensive benchmarks, such as TPC-[AB], or building very large indexes for real-life applications. The kernel enhancements are as follows:

- *Postwait driver*, a kernel software driver, provides very fast multithreaded synchronization mechanism for Oracle processes. It replaces the standard SVR4 mechanism of semaphore, which is too slow for the high TPS rate.
- *Kernel list I/O*, an IRIX enhancement, allows the Oracle database writer to flush modified buffers to disks efficiently: a single Oracle database writer can flush at least 2000 buffers per second to disk drives. With only one system call, the database writer can initiate multiple writes to all disk drives in the system. Without this functionality, Oracle database writer performance would have to use shadow processes, thus incurring the overhead of process synchronization; another limitation would be the single-threaded nature of making one system call per disk write.

## Optional Silicon Graphics Software and OPS Performance

This section briefly explains how the following optional Silicon Graphics software products can enhance OPS performance:

- PCP (Performance Co-Pilot)
- IRIX NetWorker

### Performance Co-Pilot (PCP)

Performance Co-Pilot (PCP) provides a suite of tools for performance monitoring and performance management services across the spectrum of performance domains—hardware platforms, the operating systems, the DBMS, and the applications.

PCP runs in a client/server configuration: PCP agents (clients) monitor domains and send information to the PCP server, which graphically displays the information on the workstation. PCP can be used to monitor Oracle and system activity on both servers in the OPS system.

## **IRIX NetWorker**

IRIX NetWorker reliably protects files against loss across an entire network of systems. NetWorker saves valuable administrator time by speeding and simplifying daily backup operations. As NetWorker backs up data, it creates a database of the saved data, making it easy to locate a file for recovery. Furthermore, as the network and number of files expand, NetWorker has the capacity and performance to handle the load.

IRIX NetWorker includes extended support for autochangers (jukeboxes and tape libraries), and archiving and retrieval capability. Its ability to back up raw files makes it particularly suitable for use with an OPS system, because all Oracle files are XLV raw devices.

For complete information on NetWorker, see the documentation:

- *IRIX NetWorker Administrator's Guide (007-1458-nnn)*
- *IRIX NetWorker User's Guide (007-2304-nnn)*

These documents are viewable in IRIS InSight.

---

## Configuring, Starting, and Halting the OPS Software

This chapter explains how to configure, start, and halt the OPS system. It contains these major sections:

- “Checking OPS Software Configuration” on page 15
- “Creating System Controller Password Files” on page 18
- “Starting the Node Controller Software on the Indy or O2 Workstation” on page 18
- “Starting the OPS Software Manually on the Servers” on page 18
- “Starting the OPS Software Automatically on the Servers” on page 19
- “Starting a Server for Single-Host Operation” on page 20
- “Halting the OPS Software on the Servers” on page 20
- “Halting the OPS Software Manually” on page 20
- “Checking Log Files” on page 21

### Checking OPS Software Configuration

The OPS software is installed and configured during installation of the OPS system. This section explains how you can check that the configuration files created during this process contain the correct information.

The OPS configuration files that are edited or created during OPS system setup are

- `/usr/opscm/conf/opsconf` on each server
- `/usr/opscm/conf/sidconf` on each server
- `/usr/opsnc/conf/ncconf` on the Indy or O2 workstation

Checking each file is explained separately in a subsection.

## Checking the opscnf Configuration File

To check the contents of the *opscnf* configuration file, follow these steps:

1. Open */usr/opscm/conf/opscnf* on each server.
2. Check to see that the line

```
CLUSTER 1 test 2 ichostname
```

has the correct hostname of the workstation used for IRISconsole.

3. Define the Distributed Lock Manager (DLM) domain(s) and DLM instances, as explained in the section “OPS Instances and Domains” in Chapter 1. Check that the lines under

```
#NODE dom inst ndname ndaddress cmsvc apsvc wt
```

have accurate information on your OPS servers. For example:

```
NODE 0 0 host1 150.166.42.37 opscm opsdm 1  
NODE 0 1 host2 150.166.42.38 opscm opsdm 1
```

4. Save and close */usr/opscm/conf/opscnf*.
5. Edit the file */etc/services*. Add three services:

```
opscm newnumber1/tcp  
opsdm newnumber2/tcp  
opsnc newnumber3/tcp
```

*newnumbern* is a number not used elsewhere at this site. For example:

```
opscm 7018/tcp  
opsdm 7019/tcp  
opsnc 7020/tcp
```

6. Save and exit the file.

## Checking the sidconf Configuration File

The *sidconf* file maps each Oracle instance (sid, or Oracle system ID) to a DLM domain-instance pair. To check the *sidconf* configuration file, follow these steps:

1. Determine the sid of the Oracle database for each instance.
2. In */usr/opscm/conf/sidconf*, check that the lines
 

```
MAP sid0 domainnumber instancenumbr
MAP sid1 domainnumber instancenumbr
```

 contain accurate information for the servers. For example:
 

```
MAP finance1 0 0
MAP finance2 0 1
```
3. Save and close */usr/opscm/conf/sidconf*.
4. If necessary, change permissions on this file so that it can be read by all.

## Checking the nconf Configuration File

To check the *nconf* configuration file, follow these steps:

1. Note the ports on the multiplexer to which the Remote System Control server ports are cabled.
2. In an IRIX shell window on the workstation, open the */usr/opsnc/conf/nconf* file.
3. Check that the entries under

```
#nodename      ttyname      system controller type
ops1           /dev/ttyd031  SYSCTLR |MMSC |MSC
ops2           /dev/ttyd033  SYSCTLR
```

contain accurate information for the servers. For example:

The last digit in each *tty* entry should be the multiplexer port into which the Remote System Control for each server is cabled. The numbers 1 through 16 on the multiplexer correspond to 0 through f in the *tty* entries. In the example above, the Remote System Control ports are cabled to ports 2 and 4 on the multiplexer. SYSCTLR is used if the system is a CHALLENGE server, MMSC is used if the system is a rackmount Origin2000 server, and MSC is used if the system is a deskmount Origin2000.

4. Save and close */usr/opsnc/conf/nconf*.

## Creating System Controller Password Files

If the OPS servers are Origin servers and one or both of them have system controller passwords, the passwords must be stored in the `/usr/opsnc/conf/.sysctrlpw-*` files on the Indy or O2 workstation. This is required for the Node Controller software to work correctly.

For each OPS server that has a system controller password, create a `.sysctrlpw` file by entering this command as root on the Indy or O2 workstation:

```
/usr/opsnc/bin/spng -n OPS_hostname -d sys_ctrl_type -w password
```

`OPS_hostname` is the hostname of an OPS server. `sys_ctrl_type` is the OPS server's system controller type, either **MSC** or **MMSC**. `password` is the unencrypted system controller password.

## Starting the Node Controller Software on the Indy or O2 Workstation

To start the Node Controller software (`opsnc`) on the Indy or O2 workstation, enter this command as root:

```
/usr/opsnc/bin/opsnc
```

You must restart the Node Controller software using this command each time the workstation is rebooted.

## Starting the OPS Software Manually on the Servers

This section describes how to start the OPS software manually on the servers. You may want to do this while testing the software or before you set up the servers to run the software automatically.

The OPS software consists of the Connection Manager (CM) software (`opscm`) and Oracle Distributed Lock Manager (DLM) software (`dmmmon` and `dlmd`). Follow these steps to start them manually:

1. Follow the procedure in the section “Halting the OPS Software Manually” in this chapter to ensure that the CM and DLM software aren’t already running.
2. For each server, create a startup script containing the following lines:

```
#!/sbin/sh

ORACLE_HOME=/usr/people/oracle
ORACLE_SID=sidname
LKDOM=0
LKINST=0
USER=oracle
GROUP=dba

export ORACLE_HOME ORACLE_SID LKDOM LKINST

/usr/opscm/bin/opscm
$ORACLE_HOME/bin/lkmgrd -u $USER -g $GROUP
```

In each script, make sure that the values for `LKDOM=` and `LKINST=` are accurate for the domain and instance on that server. These values must match those in `/usr/opscm/conf/sidconf`, as explained in the section “Checking the sidconf Configuration File” in this chapter. `lkmgrd` is the Distributed Lock Manager daemon. Its options specify the user ID and group of lock database owner.

3. As root, run the startup script you created in step 2 on each server to bring up the CM and DLM software on each server.

## Starting the OPS Software Automatically on the Servers

To enable the OPS software (Connection Manager and Oracle Distributed Lock Manager) to start automatically at boot time and to start it, follow these steps on each server:

1. Edit the `/etc/init.d/opscmgr` script. This script is similar to the startup script created in step 2 in the section “Starting the OPS Software Manually on the Servers” in this chapter.
2. Run this command as root:  

```
chkconfig -f opscm on
```
3. Reboot the server. For example:  

```
reboot
```

## Starting a Server for Single-Host Operation

Occasionally, you may need to run the Connection Manager software with just one server operating, for example, while one server is turned off. To start a server for single-host operation, run the *opscm* command with the **-F** option:

```
/usr/opscm/bin/opscm -F
```

**Note:** Do not use the **-F** option for normal OPS dual-host operation.

## Halting the OPS Software on the Servers

To halt the Connection Manager and Oracle Distributed Lock Manager Software on both servers, follow these steps:

1. Back up Oracle database information as needed.
2. On each OPS server, halt Oracle database operation.
3. Enter the following command to terminate the Connection Manager gracefully:

```
killall -TERM opscm
```

## Halting the OPS Software Manually

If you are not able to gracefully halt the Connection Manager and Oracle Distributed Lock Manager Software on both servers as described in the previous section, use this procedure to halt the processes and clean up. Follow these steps on each server:

1. Check for the presence of the CM lock file in */tmp*. This filename has the suffix *.nn*, where each number stands for the DLM domain and instance, for example, *.00*. If this file exists, delete it.

2. Check to see if the CM is already running by entering

```
ps -ef | grep opscm
```

3. If the CM is running, enter the following to kill it:

```
killall -TERM opscm
```

4. Check to see if the DLM is already running by entering

```
ps -ef | grep dlm
```

5. If the DLM is running, enter the following to kill its processes:

```
killall -TERM dlmon
killall -TERM dlmd
```

6. Run *ipcs* to determine the shared memory and semaphores used on the server.

```
ipcs
```

The following is an example output:

```
IPC status from /dev/kmem as of Thu May 18 14:31:22 1995
T      ID      KEY          MODE          OWNER      GROUP
Message Queues:
Shared Memory:
m       0 0x53637444 --rw-r--r--   root      sys
m     301 0x000022bb --rw-rw----   oracle    dba
m     2202 0x0c33b7c9 --rw-r-----   oracle    dba
Semaphores:
s     2200 0x00000000 --ra-r-----   oracle    dba
```

7. If Oracle or DLM is using any shared memory segments or semaphores, save them to another location if you need them for debugging a DLM or Oracle crash; otherwise, delete them with *ipcrm*. In the example in step 6, you would use

```
ipcrm -m 301 -m 2202 -s 2200
```

## Checking Log Files

You can examine log and control files to learn about failures. The log and control files that are relevant for OPS failures are

```
/var/adm/SYSLOG
```

Node Controller and Connection Manager log information

```
/var/tmp/dlm
```

Distributed Lock Manager log





---

## Tell Us About This Manual

As a user of Silicon Graphics products, you can help us to better understand your needs and to improve the quality of our documentation.

Any information that you provide will be useful. Here is a list of suggested topics:

- General impression of the document
- Omission of material that you expected to find
- Technical errors
- Relevance of the material to the job you had to do
- Quality of the printing and binding

Please send the title and part number of the document with your comments. The part number for this document is 007-2763-003.

Thank you!

## Three Ways to Reach Us

- To send your comments by **electronic mail**, use either of these addresses:
  - On the Internet: [techpubs@sgi.com](mailto:techpubs@sgi.com)
  - For UUCP mail (through any backbone site): *[your\_site]!sgi!techpubs*
- To **fax** your comments (or annotated copies of manual pages), use this fax number: 415-965-0964
- To send your comments by **traditional mail**, use this address:

Technical Publications  
Silicon Graphics, Inc.  
2011 North Shoreline Boulevard, M/S 535  
Mountain View, California 94043-1389