



# Sun Enterprise 10000 DR Configuration Guide

---

Sun Microsystems, Inc.  
901 San Antonio Road  
Palo Alto, CA 94303-4900  
U.S.A. 650-960-1300

Part No. 806-5228-10  
January 2001, Revision A

Send comments about this document to: [docfeedback@sun.com](mailto:docfeedback@sun.com)

Copyright 2001 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or 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 other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, AnswerBook2, docs.sun.com, Sun Enterprise, SunFIDDI, Sun StorEdge, OpenBoot, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and 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.

Federal Acquisitions: Commercial Software—Government Users Subject to Standard License Terms and Conditions.

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 2001 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, Californie 94303 Etats-Unis. Tous droits réservés.

Ce produit ou document est 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 États-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, AnswerBook2, docs.sun.com, Sun Enterprise, SunFIDDI, Sun StorEdge, OpenBoot, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux États-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 États-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'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 CONTREFAÇON.



#### Sun Enterprise 10000 SSP Attributions:

This software is copyrighted by the Regents of the University of California, Sun Microsystems, Inc., and other parties. The following terms apply to all files associated with the software unless explicitly disclaimed in individual files.

The authors hereby grant permission to use, copy, modify, distribute, and license this software and its documentation for any purpose, provided that existing copyright notices are retained in all copies and that this notice is included verbatim in any distributions. No written agreement, license, or royalty fee is required for any of the authorized uses. Modifications to this software may be copyrighted by their authors and need not follow the licensing terms described here, provided that the new terms are clearly indicated on the first page of each file where they apply.

IN NO EVENT SHALL THE AUTHORS OR DISTRIBUTORS BE LIABLE TO ANY PARTY FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OF THIS SOFTWARE, ITS DOCUMENTATION, OR ANY DERIVATIVES THEREOF, EVEN IF THE AUTHORS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

THE AUTHORS AND DISTRIBUTORS SPECIFICALLY DISCLAIM ANY WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, AND NON-INFRINGEMENT. THIS SOFTWARE IS PROVIDED ON AN "AS IS" BASIS, AND THE AUTHORS AND DISTRIBUTORS HAVE NO OBLIGATION TO PROVIDE MAINTENANCE, SUPPORT, UPDATES, ENHANCEMENTS, OR MODIFICATIONS.

RESTRICTED RIGHTS: Use, duplication or disclosure by the government is subject to the restrictions as set forth in subparagraph (c) (1) (ii) of the Rights in Technical Data and Computer Software Clause as DFARS 252.227-7013 and FAR 52.227-19.

This is scotty, a simple tcl interpreter with some special commands to get information about TCP/IP networks. Copyright (c) 1993, 1994, 1995, J. Schoenwaelder, TU Braunschweig, Germany, Institute for Operating Systems and Computer Networks. Permission to use, copy, modify, and distribute this software and its documentation for any purpose and without fee is hereby granted, provided that this copyright notice appears in all copies. The University of Braunschweig makes no representations about the suitability of this software for any purpose. It is provided "as is" without express or implied warranty.



# Contents

---

## **Preface**   vii

Before You Read This Book   vii

How This Book Is Organized   vii

Using UNIX Commands   viii

Typographic Conventions   ix

Shell Prompts   ix

Related Documentation   x

Accessing Sun Documentation Online   x

Ordering Sun Documentation   x

Sun Welcomes Your Comments   xi

## **1. DR Configuration Issues**   1

`dr-max-mem` Variable   1

▼ To Enable the Kernel Cage   1

Configuration for DR Detach   2

    I/O Devices   2

    Driver Parameters   3

    Target Memory Constraints   3

    Swap Space   3

Network Devices	4
Non-Network Devices	5
Processes	5
Processors	6
Reconfiguration After a DR Operation	6
When to Reconfigure	6
Disk Devices	7
DR and AP Interaction	8
RPC Time-Out or Loss of Connection	8
System Quiescence Operation	9
Suspend-Safe/Suspend-Unsafe Devices	10
Special Handling for Tape Devices	11
Special Handling of Sun StorEdge A3000	11
DR and DDI	11
DR and DDI_DETACH	12
DR and DDI_SUSPEND/DDI_RESUME	12

# Preface

---

This guide describes the domain-side configuration of the Sun Enterprise™ 10000 server Dynamic Reconfiguration (DR) feature. For information about how to use these features, refer to the appropriate user guide listed in “Related Documentation” on page ix.

---

## Before You Read This Book

This guide is intended for the Dynamic Reconfiguration system administrator who has a working knowledge of UNIX® systems, particularly those based on the Solaris™ operating environment. If you do not have such knowledge, first read all of the books in the Solaris System Administration collection in AnswerBook2™ format provided with your server and consider UNIX system administration training.

Also read and be familiar with the *TCP/IP and Data Communications Administration Guide* that is provided with your server in AnswerBook2 format.

---

## How This Book Is Organized

This guide contains one chapter that describes on how to configure and reconfigure a Sun Enterprise 10000 domain before and after a DR operation.

---

# Using UNIX Commands

This document does not contain information on basic UNIX commands and procedures such as shutting down the system, booting the system, and configuring devices.

See one or more of the following sources for this information:

- AnswerBook2 online documentation for the Solaris 2.x software environment, particularly those dealing with Solaris system administration
- Other software documentation that you received with your system

---

# Typographic Conventions

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



---

## Shell Prompts

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

---

## Related Documentation

Application	Title	Part Number
User	<i>Sun Enterprise 10000 SSP 3.4 User Guide</i>	806-4870
	<i>Sun Enterprise 10000 InterDomain Networks User Guide</i>	806-4131
	<i>TCP/IP and Data Communications Administration Guide</i>	805-4003
Reference	<i>Sun Enterprise 10000 SSP 3.4 Reference Manual</i>	806-4871
	<i>Sun Enterprise 10000 IDN Error Messages</i>	806-5231

---

## Accessing Sun Documentation Online

The `docs.sun.comSM` web site enables you to access Sun technical documentation on the Web. You can browse the `docs.sun.com` archive or search for a specific book title or subject at:

`http://docs.sun.com`

---

## Ordering Sun Documentation

Fatbrain.com, an Internet professional bookstore, stocks select product documentation from Sun Microsystems, Inc.

For a list of documents and how to order them, visit the Sun Documentation Center on Fatrain.com at:

<http://www.fatbrain.com/documentation/sun>

---

## Sun Welcomes Your Comments

We are interested in improving our documentation and welcome your comments and suggestions. You can email your comments to us at:

[docfeedback@sun.com](mailto:docfeedback@sun.com)

Please include the part number (806-5228-10) of your document in the subject line of your email.

## DR Configuration Issues

---

This chapter describes how to configure a domain for all DR operations and capabilities.



---

**Caution** – Be careful when choosing the slot into which a board is inserted to prevent disk controller renumbering. For more information, see “Reconfiguration After a DR Operation” on page 6.

---

---

### dr-max-mem Variable

With the Solaris 7 and Solaris 8 operating environments, `dr-max-mem` is no longer used. Instead, the DR feature, specifically DR detach, must be enabled by using the `system(4)` variable `kernel_cage_enable`. A caged kernel confines the nonpageable memory to a minimal (most often one) number of systems boards. By default, the kernel cage is disabled, preventing DR detach operations.

---

**Note** – DR attach is enabled regardless of the setting of `kernel_cage_enable`.

---

### ▼ To Enable the Kernel Cage

1. Edit the `/etc/system` file so that `kernel_cage_enable` equals 1.

```
set kernel_cage_enable=1
```

2. Reboot the domain.

After the reboot completes successfully, you can verify that the kernel cage is enabled by reviewing the `/var/adm/messages` file for the following message.

```
NOTICE: DR Kernel Cage is ENABLED
```

---

## Configuration for DR Detach

This section describes how to configure DR before you perform a detach operation.

### I/O Devices

The DR detach feature works with Alternate Pathing (AP) or Solstice™ DiskSuite™ mirroring when you detach a board that hosts I/O controllers that are attached to vital system resources. If, for example, the root (`/`) or `/usr` partition is on a disk attached to a controller on the board, the board cannot be detached unless there is a hardware alternate path to the disk, and AP has been configured to take advantage of it, or the disk is mirrored. The alternate path or the mirrors must be hosted by other boards in the domain. The same applies to network controllers. The board that hosts the Ethernet controller that connects the SSP to the Sun Enterprise 10000 platform cannot be detached unless an alternate path exists to an Ethernet controller on another board for this network connection.

To enable device suspension for the `soc` and `pln` drivers, you must edit the `/etc/system` file so that the `pln_enable_detach_suspend` and `soc_enable_detach_suspend` variables are set to 1, as in the following example:

```
set pln:pln_enable_detach_suspend=1
set soc:soc_enable_detach_suspend=1
```

The domain swap space should be configured as multiple partitions on disks attached to controllers hosted by different boards. With this kind of configuration, a particular swap partition is not a vital resource because swap partitions can be added and deleted dynamically (refer to the `swap(1M)` man page for more information).

---

**Note** – When memory (`swapfs`) or swap space on a disk is detached, there must be enough memory or swap space remaining in the domain to accommodate currently running programs.

---

A board that hosts non-vital system resources can be detached whether or not there are alternate paths to the resources. All of the devices on the board must be closed before the board can be detached; all of its file systems must be unmounted; and, its swap partitions must be deleted. You may have to kill processes that have open files or devices, or place a hard lock on the file systems (using `lockfs(1M)`) before you unmount the boards.

All I/O device drivers involved with I/O devices on the board(s) must support the `DDI_DETACH` option in the detach entry-point of the driver. This option releases all system resources associated with that device or adapter.

## Driver Parameters

If you use the `ndd(1M)` command to set the configuration parameters for network drivers, the parameters may not persist after a DR Detach or DR Attach operation. Use the `/etc/system` file or the `driver.conf` file for a specific driver to set the parameters permanently.

## Target Memory Constraints

When detaching a board with nonpageable memory, DR must locate an alternative (target) memory board to which to copy the nonpageable memory. In the Solaris 7 5/99 version, if no target board is found, the detach operation is refused, and DR displays the following warning message on the system console:

```
WARNING: sfdr: sfdr_pre_release_mem: no available target for mem-  
unit (board.0)
```

## Swap Space

The domain swap configuration consists of the swap devices and `swapfs` (memory). The domain must contain enough swap space so that it can flush pageable memory. For example, if you want to remove 1 Gbyte of memory from a 2-Gbyte domain, you will need 1 Gbyte of swap space, depending on the load. Insufficient swap space prevents DR from completing the detach of a board that contains memory. If this happens, the memory drain phase does not complete, so you must abort the detach operation.

# Network Devices

DR automatically terminates usage of all network interfaces on the board that is being detached. When you complete the detach operation, the `dr_daemon(1M)` identifies all configured interfaces on the board being detached and issues the following `ifconfig(1M)` commands on each such interface.

```
ifconfig interface down
ifconfig interface unplumb
```

Additionally, if FDDI interfaces are detached, DR kills the FDDI network monitoring daemon before you perform the detach operation. DR then restarts it after the detach is complete. Note that the `/usr/sbin/nf_snmd` daemon for `nf` devices is neither started nor stopped when a board that contains a FDDI interface is attached.

DR does not execute these commands on a board that contains a network interface that fits any of the following conditions. In these cases, the detach operation fails and DR displays an error message.

- The interface is the primary network interface for the domain; that is, the interface whose IP address corresponds to the network interface name contained in the file `/etc/nodename`. Note that bringing down the primary network interface for the domain prevents network information name services from operating, which results in the inability to make network connections to remote hosts using applications such as `ftp(1)`, `rsh(1)`, `rcp(1)`, `rlogin(1)`. NFS client and server operations are also affected.
- The interface is on the same subnet as the SSP host for the system; that is, the subnet of the IP address that corresponds to the SSP host name found in `/etc/ssphostname`. Bringing down this interface interrupts communication between the host and SSP. Since DR operations are initiated on the SSP, control of the detach process would be lost. (Note that the `/etc/ssphostname` file contains the name of the SSP that controls the host; therefore, if you rename the SSP, the `/etc/ssphostname` must be manually updated.)
- The interface is the active alternate for an Alternate Pathing (AP) metadvice when the AP metadvice is plumbed. Interfaces used by AP should not be the active path when the board is being detached. AP 2.1 performs the switch automatically; however, you can manually switch the active path to an interface that is not on the board being detached. If no such path exists, manually execute the `ifconfig down` and `ifconfig unplumb` commands on the AP interface. (To manually switch an active path, use the `apconfig(1M)` command.)



---

**Caution** – Detaching network interfaces may affect NFS client systems.

---

# Non-Network Devices

All non-network devices must be closed before they are detached. In the Hostview device display and in the `drshow(1M)` I/O listing, there is an open count field that indicates how many processes have opened particular devices. To see which processes have these devices open, use the `fuser(1M)` command on the domain.

You must perform certain tasks for non-network devices. Although the following list of tasks implies a sequence of order, strict adherence to the order is not necessary.

1. If the redundancy features of Alternate Pathing or Solstice DiskSuite mirroring are used to access a device connected to the board, reconfigure these subsystems so that the device or network is accessible using controllers on other system boards. Note that for Alternate Pathing 2.1, the system automatically switches the disk devices to an alternate interface if one is available.
2. Unmount file systems, including Solstice DiskSuite metadevices that have a board-resident partition (for example, `umount /partit`).
3. Remove Alternate Pathing or Solstice DiskSuite databases from board-resident partitions. The location of Alternate Pathing or Solstice DiskSuite databases is explicitly chosen by the user and can be changed.
4. Remove any private regions used by Sun Enterprise Volume Manager™ or Veritas Volume Manager. Volume manager by default uses a private region on each device that it controls, so such devices must be removed from volume manager control before they can be detached.
5. Remove disk partitions from the swap configuration by using `swap(1M)`.
6. Either kill any process that directly opens a device or raw partition, or direct it to close the open device on the board.
7. If a detach-unsafe device is present on the board, close all instances of the device and use `modunload(1M)` to unload the driver.
8. Kill all of the real-time processes that are open if the operating environment must be suspended.



---

**Caution** – Unmounting shared file systems by using the `share(1M)` utility may affect NFS client systems.

---

## Processes

You must perform certain tasks for processes. Although the following list of tasks implies a sequence of order, strict adherence to the order is not necessary.

1. If the operating environment must be suspended, kill all of the real-time processes that are running.
2. Kill, or unbind, any processes that are bound to on-board processors.

Processes bound to the processors of a board prevent that board from being detached. You can use `pbind(1M)` to rebind them to other processors.

## Processors

The boot processor is responsible for maintaining the netcon BBSRAM buffer. Before detaching a board on which the boot processor resides, the `dr_daemon(1M)` must assign the boot processor role to another active (online) processor.

---

# Reconfiguration After a DR Operation

This section describes how to reconfigure your domain after you have attached or detached a system board.

---

**Note** – As of the Solaris 8 GA release, manual reconfiguration is no longer needed. A new DDI subsystem, `devfsadm`, completes all of the reconfiguration tasks.

---

The DR user interface enables you reconfigure the domain after a DR Attach or DR Detach operation. The reconfiguration sequence is the same as the reconfiguration boot sequence (`boot -r`):

```
drvconfig; devlinks; disks; ports; tapes;
```

When you execute the reconfiguration sequence after you attach a board, device path names not previously seen by the domain are written to the `/etc/path_to_inst` file. The same path names are also added to the `/devices` hierarchy, and links to them are created in the `/dev` directory.

## When to Reconfigure

You should reconfigure the domain if any of the following conditions occur:



- **Board Addition** – When you add a board to a domain, you must execute the reconfiguration sequence to configure the I/O devices that are associated with the board.
- **Board Deletion** – If you remove a board that is not to be replaced, you may, but do not have to, execute the reconfiguration sequence to clean up the `/dev` links.
- **Board Replacement** – If you remove a board then reinsert it in a different slot or if you replace a board with another board that has different I/O devices, you must execute the reconfiguration sequence to configure the I/O devices that are associated with the board. However, if you replace a board with another board that hosts the *same* set of I/O devices, inserting the replacement into the *same* slot, you do not need to execute the reconfiguration sequence. But, be sure to insert a replacement board into the same slot that was vacated to retain the original mapping of `/dev` links to physical names.

## Disk Devices

Disk controllers are numbered consecutively as the `disks(1M)` program encounters them. All disk partitions are assigned `/dev` names according to the disk controller number that `disks(1M)` assigns. For example, all disk partitions that are accessible using disk controller 1 are named `/dev/dsk/cXtYdZsW`

where:

`X` is the disk controller number,

`Y`, in most cases, corresponds to the disk target number,

`Z` corresponds to the logical unit number, and

`W` corresponds to the partition number.

When the reconfiguration sequence is executed after a board is detached, the `/dev` links for all of the disk partitions on that board are deleted. The remaining boards retain their current numbering. Disk controllers on a newly inserted board are assigned the next available lowest number by `disks(1M)`.

---

**Note** – The disk controller number is part of the `/dev` link name used to access the disk. If that number changes during the reconfiguration sequence, the `/dev` link name also changes. This change may affect file system tables and software, such as Solstice DiskSuite™, which use the `/dev` link names. Update `/etc/vfstab` files and execute other administrative actions to change the `/dev` link names.

---

---

## DR and AP Interaction

DR notifies the AP subsystem when system boards are attached, detached, or placed in the drain state. In addition, DR queries AP about which controllers are in the AP database and their status (active or inactive). This communication occurs between the `dr_daemon(1M)` and `ap_daemon(1M)`. If the `ap_daemon(1M)` is not present, an error message is placed in the syslog messages buffer of the domain and DR operations continue without error. To disable this interaction, use the `-a` option when you invoke `dr_daemon(1M)`. Refer to the `dr_daemon(1M)` man page in the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual*.

If you are using AP version 2.1, the operating environment automatically switches off the active disk controllers on outgoing boards during the complete-detach phase of DR. If you are using AP version 2.0, you need to manually switch off the active disk controllers before you start the complete-detach phase. For the Solaris 8 operating environment, you must upgrade to AP version 2.3. For more information about DR and AP interaction, see the *Sun Enterprise Servers Alternate Pathing 2.3 User Guide*. For more information about AP and SDS, refer to the *RAS Companion*.

---

## RPC Time-Out or Loss of Connection

The `dr_daemon(1M)`, which runs in each domain, communicates with Hostview and the `dr(1M)` shell application (both of which run on the SSP) by way of Remote Procedure Calls (RPCs). If an RPC time-out or connection failure is reported during a DR operation, check the domain. The daemon must be configured in the `/etc/inetd.conf` file of the domain. The following line (which appears on a single line) must be present in the file:

```
300326/4 tli rpc/tcp wait root \  
/platform/SUNW,Ultra-Enterprise-10000/lib/dr_daemon/ dr_daemon
```

If the DR daemon is configured in `/etc/inetd.conf`, kill the `dr_daemon(1M)` if it is currently running. In addition, send a HUP signal to the `inetd(1M)` daemon to cause it to re-read the `inetd.conf(4)` configuration file:

```
# kill dr_daemon_pid  
# kill -HUP inetd_pid
```

In the first command, *dr\_daemon\_pid* is the process ID of the DR daemon. In the second command, *inetd\_pid* is the process ID of the *inetd(1M)* daemon. You can check */var/adm/messages* for possible error messages from *inetd(1M)* if it is having trouble starting the *dr\_daemon(1M)*. The DR daemon executable file should exist in the */platform/SUNW,Ultra-Enterprise-10000/lib* directory.

At this point, try the DR operation again, starting from the beginning.

---

## System Quiescence Operation

During a DR Detach operation on a system board with nonpageable OBP or kernel memory, the operating environment is briefly quiesced; that is, all operating environment and device activity on the domain centerplane must cease during the critical phase of the operation. The quiescence only affects the target domain; other domains in the system are not affected.

Before a board can be detached, the operating environment must temporarily suspend all processes, processors, and device activities. If the operating environment cannot quiesce, it displays its reasons, which may include the following:

- Real-time processes are running in the domain.
- A device that cannot be quiesced by the operating environment (that is, a suspend-unsafe device) is open.

The conditions that cause processes not to suspend are generally temporary in nature. You can retry the operation until the quiescence succeeds.

A failure to quiesce due to real-time processes or open suspend-unsafe devices is known as a forcible condition. You have the option of performing either a retry or forced retry. When you force the quiescence, you give the operating environment permission to continue with the quiescence even if forcible conditions are still present.



---

**Caution** – Exercise care when using the *force* option.

---

If a real-time process is running, determine if suspending the process would produce an adverse effect on the functions performed by the process. If not, you can force the operating environment to quiesce. (To force a quiescence, you can either click the *Force* button within Hostview as described in “To Detach a Board With Hostview” in the *Sun Enterprise 10000 Dynamic Reconfiguration Guide* in the SSP 3.3 AnswerBook2 collection, or enter the *complete\_detach(1M)* command with the *force* option within the *dr(1M)* shell application. Otherwise, you can abort the operation and try again later.

If any suspend-unsafe device is open and cannot be closed, you can manually suspend the device, and then force the operating environment to quiesce. After the operating environment resumes, you can manually resume the device (see “Suspend-Safe/Suspend-Unsafe Devices” on page 10).

## Suspend-Safe/Suspend-Unsafe Devices

A suspend-safe device is one that does not access the domain centerplane (for example, it does not access memory or interrupt the system) while the operating environment is quiesced. A driver is considered suspend-safe if it supports operating environment quiescence (suspend/resume) and guarantees that when a suspend request is successfully completed, the device that the driver manages will not attempt to access the domain centerplane, even if the device is open when the suspend request is made. All other I/O devices are suspend-unsafe when open.

---

**Note** – At the time of this printing, the drivers released by Sun Microsystems™ that are known to be suspend-safe are `st`, `sd`, `isp`, `esp`, `fas`, `sbus`, `pci`, `pci-pci`, `qfe`, `hme` (SunFastEthernet™), `nf` (NPI-FDDI), `qe` (Quad Ethernet), `le` (Lance Ethernet), the SSA drivers (`soc`, `pln`, and `ssd`), and the Sun StorEdge™ A5000 drivers (`sf`, `socal`, `ses`).

---

To enable device suspension for the `soc` and `pln` drivers, you must edit the `/etc/system` file so that the `pln_enable_detach_suspend` and `soc_enable_detach_suspend` variables are set to 1, as in the following example:

```
set pln:pln_enable_detach_suspend=1
set soc:soc_enable_detach_suspend=1
```

The operating environment refuses a quiesce request if a suspend-unsafe device is open. If you can manually suspend the device, you can force the operating environment to quiesce. To manually suspend the device, you may have to close the device by killing the processes that have it open, ask users not to use the device, or disconnect the cables. For example, if a device that allows asynchronous unsolicited input is open, you can disconnect its cables prior to quiescing the operating environment, preventing traffic from arriving at the device and the device from accessing the domain centerplane. You can reconnect the cables after the operating environment resumes. If you cannot make a device suspend its access to the domain centerplane, you should not force the operating environment to quiesce. Doing so could cause a domain to crash or hang. Instead, postpone the DR operation until the suspend-unsafe device is no longer open.



---

**Caution** – If you attempt a forced quiesce operation while activity is occurring on a suspend-unsafe device, the domain may hang. However, if the domain hangs, it will not affect other domains that are running on the Sun Enterprise 10000 system.

---

## Special Handling for Tape Devices

For the Solaris 8 operating environment, tape devices that are natively supported by Sun Microsystems™ are suspend-safe and detach-safe (refer to the `st(7D)` man page for a list of natively-supported drives). If a system board that you are detaching contains a natively-supported tape device, you can safely detach the board without suspending the device. If you want to use a tape device that is not natively supported by Sun Microsystems, you can use it, but you should make it detach-safe. To ensure correct input/output and DR operations, you need to make a suitable entry in `/kernel/drv/st.conf` with the `ST_UNLOADABLE (0x0400)` flag set in the entry (refer to the `st(7D)` man page for more information). After you update `st.conf`, you must reboot the domain to process the new entry.

## Special Handling of Sun StorEdge A3000

The Sun StorEdge™ A3000 (formerly known as the RSM Array 2000) has dual controller paths with automatic load balancing and automatic failover. To detach a system board that has one or both of the StorEdge A3000 controllers, the controllers on the board that is being detached must be idle or offline. You can take these controllers offline manually by using the `rm6` or `rdacutil` programs before you attempt to detach the system board.

---

## DR and DDI

Not all drivers support the Sun Enterprise 10000 system Dynamic Reconfiguration (DR) feature. To support DR, a driver must be able to perform three basic DDI/DKI (Device Driver Interface/Device Kernel Interface) functions, `DDI_DETACH` and `DDI_SUSPEND/DDI_RESUME`. These functions impact DR in different ways.

## DR and DDI\_DETACH

You can detach a system board that hosts a device only if the driver for that device supports the `DDI_DETACH` interface, or is not currently loaded. `DDI_DETACH` provides the ability to detach a particular instance of a driver without impacting other instances that are servicing other devices. A driver that supports `DDI_DETACH` is called *detach-safe*; a driver that does not support `DDI_DETACH` is called *detach-unsafe*.

Detaching a detach-unsafe driver that is loaded involves the following process.

- Stopping all usage of the controller for the detach-unsafe device and all other controllers of the same type on all of the boards in the domain.

Because the detach-unsafe driver must be unloaded, you must stop usage of that controller type on *all* of the system boards in the domain. The remaining controllers can be used again after the DR Detach is complete.

- Using standard Solaris interfaces to manually close and to unload all such drivers on the board.

Refer to the `modunload(1M)` man page in the *SunOS Reference Manual*.

- Detaching the system board in the normal fashion.

If you cannot accomplish the above process, you can reboot the domain with the board blacklisted (refer to the `blacklist(4)` man page), so the board can be removed later.

---

**Note** – Many third-party drivers (those purchased from vendors other than Sun Microsystems) do not support the standard Solaris `modunload(1M)` interface. Conditions that invoke the functions occur infrequently during normal operation and the functions are sometimes missing or work improperly. Sun Microsystems suggests that you test these driver functions during the qualification and installation phases of any third-party device.

---

## DR and DDI\_SUSPEND/DDI\_RESUME

To perform a DR Detach of a board that contains nonpageable memory, the domain must be quiesced. Memory can be detached only when all of the drivers throughout the entire domain (not just on the board being detached) either support the `DDI_SUSPEND/DDI_RESUME` driver interface, or are closed. Drivers that support these DDI functions are called *suspend-safe*; drivers that do not are called *suspend-unsafe*.

The most straightforward way to quiesce a domain is to close any *suspend-unsafe* devices. For each network driver you must execute the `ifconfig(1M)` command with its `down` parameter, then again with its `unplumb` parameter (refer to the `ifconfig(1M)` man page for more information).

---

**Note** – It should be possible to unplumb all network drivers. However, this action is rarely tested in normal environments and may result in driver error conditions. If you use DR, Sun Microsystems suggests that you test these driver functions during the qualification and installation phases of any *suspend-unsafe* device.

---

If the system refuses to quiesce because a *suspend-unsafe* driver is open, you can force the operating domain to quiesce. Doing so forces the operating environment to permit the detach. Note that, although a detach can be forced to proceed when there are open *suspend-unsafe* devices in the system, it is not possible to force a detach when a *detach-unsafe* device resides on the board and its driver is loaded.

To successfully force the operating environment to quiesce, you must manually quiesce the controller. Procedures to do that, if any, are device-specific. The device must not transfer any data, reference memory, or generate interrupts during the operation. Be sure to test any procedures used to quiesce the controller while it is open prior to executing them on a production system.



---

**Caution** – Using the `force` option to quiesce the operating environment, without first successfully quiescing the controller, can result in a domain failure and subsequent reboot.

---





# Index

---

## A

- Alternate Pathing (AP) and DR, 2
- alternate pathing and vital partitions during detach, 2
- AP (Alternate Pathing) and DR, 2
- AP / DR interaction, disabling, 8
- AP and Solstice DiskSuite, 2
- attach
  - reconfiguration sequence after attach, 6
- automatically switching off active controllers during detach, 8

## B

- blacklisting, alternative for detach-unsafe devices, 12
- board added, reconfigure after, 7
- board deleted, optionally reconfigure after, 7
- board replaced, reconfigure after, 7

## C

- communication timeouts affecting Hostview and dr(1M), 8
- configuring for detach, 2
- configuring swap space I/O controllers across boards, 2
- connection, loss of, 8
- controllers (disk), number of, 7

## D

- DDI/DKI, 11
- DDI\_DETACH, 11, 12
- DDI\_DETACH support needed for detach, 3
- DDI\_RESUME, 11, 12
- DDI\_SUSPEND, 11, 12
- detach
  - configuring for detach, 2
  - detaching detach-unsafe device, 12
  - devices must be closed before detach, 3
  - file systems unmounted before detach, 3
  - I/O controllers on board being detached, 2
  - network between SSP and UE10000, and detach, 2
  - network controllers and detach, 2
  - pageable memory and swap space during detach, 3
  - RSM 2000 and detach, 11
  - Sun StorEdge A3000 and detach, 11
  - swap partitions must be deleted before detach, 3
  - swap space and detach, 2
  - switching of active controllers during detach, 8
- detach and network devices, 4
- detach and non-network devices, 5
- detach and processors, 6
- detach-safe, 12
- detach-safe tape devices, 11
- detach-unsafe, 12
- detach-unsafe devices present, cannot force detach, 13
- dev, reconfiguring /dev links after DR operation, 6

- devices must be closed before detach, 3
- disabling AP / DR interaction, 8
- disk controller numbering, 7
- disk devices, reconfiguring after DR operation, 7
- disk swap space, and detach, 2
- DR / AP interaction, disabling, 8
- DR suspend-safe device, 10
- DR suspend-unsafe device, 10
- drivers that support DR, 11
- drivers, listing of suspend-safe drivers, 10

## E

- Ethernet between SSP and UE1000, and detach, 2

## F

- file systems unmounted before detach, 3
- files
  - st.conf (ST\_UNLOADABLE flag and tape devices), 11
- force quiesce, how to, 9
- forcible conditions and quiesce failures, 9

## H

- hard lock on file systems (lockfs) before detach, 3

## I

- I/O controllers on board being detached, 2
- I/O devices, configuring for detach, 2
- I/O devices, reconfiguring after DR operation, 6

## L

- lock on file systems (lockfs) before detach, 3
- loss of connection, 8

## M

- manually suspending suspend-unsafe devices, 10

## N

- network between SSP and UE1000, and detach, 2
- network controllers and detach, 2
- network devices and detach, 4
- network drivers, suspend-unsafe, 13
- non-network devices and detach, 5
- numbering of disk controllers, 7

## P

- pageable memory and swap space, during detach, 3
- processors and detach, 6

## Q

- quiesce affects only target domain, 9
- quiesce failures and forcible conditions, 9
- quiesce OS during detach, and nonpageable memory, 9
- quiesce OS, how to force, 9
- quiesce, reasons it may fail, 9
- quiescing OS and real-time processes, 9
- quiescing OS and suspend-unsafe devices, 9

## R

- real-time processes and quiescing OS, 9
- reconfiguration sequence after attach, 6
- reconfigure after board added, 7
- reconfigure after board deleted, optionally, 7
- reconfigure after board replaced, 7
- reconfiguring disk devices after DR operation, 7
- reconfiguring domain after DR operation, 6
- reconfiguring, when to, 6
- releasing system resources (DDI\_DETACH support) needed for detach, 3
- root partition and I/O controllers during detach, 2

RPC timeout, 8  
RSM 2000 and detach, 11

**U**  
usr partition and I/O controllers during detach, 2

## **S**

Solstice DiskSuite and mirroring, 2  
ST\_UNLOADABLE flag and tape devices, 11  
StorEdge A3000 and detach, 11  
Sun StorEdge A3000 and detach, 11  
suspend affects only target domain, 9  
suspend failures and forcible conditions, 9  
suspend OS, how to force, 9  
suspend, reasons it may fail, 9  
suspending OS and real-time processes, 9  
suspending OS and suspend-unsafe devices, 9  
suspending OS during detach, and nonpageable memory, 9  
suspend-safe device, 10  
suspend-safe devices, 12  
suspend-safe drivers listing, 10  
suspend-unsafe device, 10  
suspend-unsafe device and quiescing OS, 9  
suspend-unsafe devices, 12  
suspend-unsafe devices, dealing with, 10  
suspend-unsafe devices, manually suspending, 10  
suspend-unsafe tape devices, 11  
swap partitions deleted before detach, 3  
swap space, configuring for detach, 3  
swap space, configuring I/O controllers across boards, 2  
switching off active controllers during detach, 8  
system board added, reconfigure after, 7  
system board deleted, optionally reconfigure after, 7  
system board replaced, reconfigure after, 7

## **T**

tape devices and ST\_UNLOADABLE flag, 11  
tape devices, detach-safe, 11  
tape devices, suspend-unsafe, 11  
timeout, RPC, 8  
timeouts affecting Hostview and dr(1M), 8

