System Management Services (SMS) 1.2 Administrator Guide

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara,CA 95054 U.S.A.
U.S.A. 650-960-1300

Part No. 816-2527-10
February 2002, Revision A

Send comments about this document to: docfeedback@sun.com

Contents

Figures

Preface

The *System Management Services (SMS) 1.2 Administrator Guide* describes the SMS software components of the Sun Fire™ 15K server system product line.

Before You Read This Book

This manual is intended for the Sun Fire 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, read the Solaris User and System Administrator AnswerBook2™ documentation provided with this system, and consider UNIX system administration training.

All members of the next-generation Sun Fire server family can be configured as loosely coupled clusters. However, it is currently outside of the scope of this document to address system management for Sun Fire 15K cluster configurations.

How This Book Is Organized

This guide contains the following chapters:

Chapter 1 introduces System Management Services and describes its command line interfaces.

Chapter 2 introduces security on the domains.

Chapter 3 describes SMS domain internals and explains how to use them.

Chapter 4 describes domain configuration.

Chapter 5 describes the control functions.

Chapter 6 describes network services available and how to use them.

Chapter 7 describes status monitoring.

Chapter 8 describes system controller (SC) failover.

Chapter 9 describes event monitoring.

Appendix A provides a list of SMS man pages.

Appendix B describes SMS error messages.

Glossary is a list of words and phrases and their definitions.

Using UNIX Commands

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

Refer to one or more of the following for this information:

- *Solaris Handbook for Sun Peripherals*
- AnswerBook2 online documentation for the Solaris software environment
- 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 `.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	Read Chapter 6 in the *User's Guide*. These are called *class* options. You *must* be superuser to do this.
	Command-line variable; replace with a real name or value	To delete a file, type `rm` *filename*.

Shell Prompts

Shell	Prompt
C shell	*sc_name*:*sms-user*`:>` or *domain_id*:*sms-user*`:>`
C shell superuser	*sc_name*:`#` or *domain_id*:`#`
Bourne shell and Korn shell	`>`
Bourne shell and Korn shell superuser	`#`

Related Documentation

Application	Title	Part Number
Reference (man pages)	*System Management Services (SMS) 1.2 Reference Manual*	816-2528-10
Installation	*System Management Services (SMS) 1.2 Installation Guide and Release Notes*	816-2529-10
Options	*System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*	816-3282-10
	Sun Fire 15K Dynamic Reconfiguration User Guide	816-3283-10
	IP Network Multipathing Administration Guide	816-0850-10
	OpenBoot 4.x Command Reference Manual	816-1177-10

Accessing Sun Documentation Online

A broad selection of Sun system documentation is located at:

`http://www.sun.com/products-n-solutions/hardware/docs`

A complete set of Solaris documentation and many other titles are located at:

`http://docs.sun.com`

Ordering Sun Documentation

Fatbrain.com, the Internet's most comprehensive 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 Fatbrain.com at `http://www1.fatbrain.com/documentation/sun.`

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. You can email your comments to Sun at:

`docfeedback@sun.com`

Please include the part number (816-2527-05) of your document in the subject line of your email.

CHAPTER 1

Introduction to System Management Services

This manual describes the System Management Services (SMS) 1.2 software that is available with the Sun Fire 15K server system.

Sun Fire 15K Server System

The Sun Fire 15K server is a member of the next-generation Sun Fire server family.

The system controller (SC) in the Sun Fire 15K is a multifunction, Nordica-based printed circuit board (PCB) which provides critical services and resources required for the operation and control of the Sun Fire system. In this book, the system controller is simply called the *SC*.

The Sun Fire 15K system is often referred to as the *platform*. System boards within the platform can be logically grouped together into separately bootable systems called *dynamic system domains*, or simply *domains*.

Up to 18 domains can exist simultaneously on a single platform. (Domains are introduced in this chapter, and are described in more detail in Chapter 4 “SMS Configuration”.) The system management services (SMS) software lets you control and monitor domains, as well as the platform itself.

SMS software packages are installed on the SC. In addition, SMS communicates with the Sun Fire 15K system over an Ethernet connection, see “Management Network Services” on page 114.

SMS Features

SMS software supports Sun Fire servers running the Solaris 8 02/02 operating environment with the Common Desktop Environment (CDE). SMS is compatible with Sun Fire 15K domains that are running the Solaris 8 02/02 Operating Environment. The commands provided with the SMS software can be used remotely.

Note – Graphical user interfaces for many of the commands in SMS are provided by Sun™ Management Center. For more information, see “Sun Management Center” on page 10.

SMS enables the platform administrator to perform the following tasks:

- Administrate domains by logically grouping *domain configurable units* (DCU) together. DCUs are system boards such as: CPU and I/O boards. Domains are able to run their own operating systems and handle their own workloads. See Chapter 4 “SMS Configuration”.
- Dynamically reconfigure a domain so that currently installed system boards can be *logically* attached to or detached from the operating system while the domain continues running in multiuser mode. This feature is known as *dynamic reconfiguration* and is described in the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide* (A system board can be *physically* swapped in and out when it is not attached to a domain, while the system continues running in multiuser mode.)
- Perform automatic dynamic reconfiguration of domains using a script. Refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.
- Monitor and display the temperatures, currents, and voltage levels of one or more system boards or domains.
- Monitor and control power to the components within a platform.
- Execute diagnostic programs such as power-on self-test (POST).

In addition, SMS:

- Warns you of impending problems, such as high temperatures or malfunctioning power supplies.
- Notifies you when a software error or failure has occurred.
- Monitors a dual SC configuration for single points of failure and performs an automatic failover from the main SC to the spare or from the primary control board to the spare control board, depending on the failure condition detected.
- Automatically reboots a domain after a system software failure (such as a panic).
- Keeps logs of interactions between the SC environment and the domains.
- Provides support for the Sun Fire 15K system dual grid power option.

SMS enables the domain administrator to perform the following tasks:

- Administrate domains by logically grouping *domain configurable units* (DCU) together. DCUs are system boards such as: CPU and I/O boards. Domains are able to run their own operating systems and handle their own workloads. See Chapter 4 “SMS Configuration”.
- Boot domains for which the administrator has privileges.
- Dynamically reconfigure a domain for which the administrator has privileges, so that currently installed system boards can be *logically* attached to or detached from the operating system while the domain continues running in multiuser mode. This feature is known as *dynamic reconfiguration* and is described in the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*. (A system board can be *physically* swapped in and out when it is not attached to a domain, while the system continues running in multiuser mode.)

- Perform automatic dynamic reconfiguration of domains using a script for which the administrator has privileges. Refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.
- Monitor and display the temperatures, currents, and voltage levels of one or more system boards or domains for which the administrator has privileges.
- Execute diagnostic programs such as power-on self-test (POST) for which the administrator has privileges.

The following features are provided in this release of Sun Fire 15K SMS:

- Dynamic System Domain (DSD) Configuration
- Configured Domain Services
- Domain Control Capabilities
- Domain Status Reporting
- Hardware Control Capabilities
- Hardware Status Monitoring, Reporting and Handling
- Hardware Error Monitoring, Reporting and Handling
- System Controller (SC) failover
- Configurable Administrative Privileges
- Dynamic FRUID

System Architecture

SMS architecture is best described as distributed client-server. `init`(1M) starts (and restarts as necessary) one process: `ssd`(1M). `ssd` is responsible for monitoring all other SMS processes and restarting them as necessary. See FIGURE 3-1 on page 29.

The Sun Fire 15K platform, the SC, and other workstations communicate over Ethernet. You perform SMS operations by entering commands on the SC console after remotely logging in to the SC from another workstation on the local area network. You must log in as a user with the appropriate platform or domain privileges if you want to perform SMS operations (such as monitoring and controlling the platform).

Note – The domains require at least one SC to be powered on or they will hang.

Dual controller boards are supported within the Sun Fire 15K platform. One board is designated as the primary or main controller board, and the other is designated as the spare controller board. If the main control board fails, the failover capability automatically switches to the spare control board as described in Chapter 8 “SC Failover” on page 127.

Most *domain configurable units* are active components and you need to check the system state before powering off any DCU.

Note – Circuit breakers must be on whenever a board is present, including expander boards, whether or not the board is powered on.

For details, see “Power Control” on page 103.

SMS Administration Environment

Administration tasks on the Sun Fire 15K system are secured by group privilege requirements. Upon installation, SMS installs the following 39 UNIX groups to the `/etc/group` file.

- platadmn- Platform administrator
- platoper - Platform operator
- platsvc - Platform service
- dmn[*A*...*R*]admn - domain [*domain_id*|*domain_tag*] administrator (18)
- dmn[*A*...*R*]rcfg - domain [*domain_id*|*domain_tag*] configurator (18)

`smsconfig`(1M) allows an administrator to add, remove and list members of platform and domain groups as well as set platform and domain directory privileges using the `-a`, `-r` and `-l` options.

`smsconfig` also can configure SMS to use alternate group names including NIS managed groups using the `-g` option. Group information entries can come from any of the sources for groups specified in the`/etc/nsswitch.conf` file (refer to `nsswitch.conf`(4)). For instance, if domain A was known by its domain tag as the "Production Domain," an administrator could create a NIS group with the same name and configure SMS to use this group as the domain A administrator group instead of the default, `dmnaadmn`. For more information, refer to the *System Management Services (SMS) 1.2 Installation Guide and Release Notes*, see “Administration Models” on page 12 and refer to the `smsconfig` man page.

SMS Operating Environment

You can interact with the SC and the domains on the Sun Fire 15K system by using SMS commands.

SMS provides a command line interface to the various functions and features it contains.

▼ To Begin Using the SC

1. **Boot the SC.**

   For the examples in this guide, the *sc_name* is `sc0` and *sms-user* is the *user-name* of the administrator, operator, configurator or service personnel logged onto the system.

   The privileges allotted to the user are determined by the platform or domain groups to which the user belongs. In these examples, the *sms-user* is assumed to have both platform and domain administrator privileges, unless otherwise noted.

   For more information on the function and creation of SMS user groups, refer to the *System Management Services (SMS) 1.2 Installation Guide and Release Notes* and see "Administration Models" on page 12.

Note – This procedure assumes that `smsconfig -m` has already been run. If `smsconfig -m` has not been run, you will receive the following error when SMS attempts to start and SMS will exit.

```
sms: smsconfig(1M) has not been run. Unable to start sms services.
```

2. **Log in to the SC and verify that SMS software startup has completed. Type:**

```
sc0:sms-user:> showplatform
```

3. **Wait until** `showplatform` **finishes displaying platform status:**

   At this point you can begin using SMS programs.

SMS Console Window

An SMS console window provides a command line interface from the SC to the Solaris operating environment on the domain(s).

▼ To Display a Console Window Locally

1. **Log in to the SC, if you have not already done so.**

Note – You must have domain privileges for the domain on which you wish to run `console`.

---

2. **Type**:

```
sc0:sms-user:> console -d domain_id|domain_tag <option>
```

`console` creates a remote connection to the domain's virtual `console` driver, making the window in which the command is executed a "console window" for the specified domain (*domain_id* or *domain_tag*).

The following options are available:

`-f`

Force

Opens a domain console window with "locked write" permission, terminates all other open sessions, and prevents new ones from being opened. This constitutes an "exclusive session." Use it only when you need exclusive use of the console (for example, for private debugging). To restore multiple-session mode, either release the lock (~^) or terminate the console session (~.).

`-g`

Grab

Opens a console window with "unlocked write" permission. If another session has "unlocked write" permission, the new console window takes it away. If another session has "locked" permission, this request is denied and a read-only session is started.

`-l`

Lock

Opens a console window with "locked write" permission. If another session has "unlocked write" permission, the new console window takes it away. If another session has "locked" permission, the request is denied and a read-only session is started.

`-r`

Read Only

Opens a console window in read-only mode

If `console` is invoked without any options when no other console windows are running for that domain, it comes up in exclusive “locked write” mode session.

If `console` is invoked without any options when one or more non-exclusive console windows are running for that domain, it will come up in "read-only" mode.

Locked write permission is more secure. It can only be taken away if another console is opened using `console -f` or if `~*`(tilde-asterisk) is entered from another running console window. In both cases, the new console session is an "exclusive session", and all other sessions are forcibly detached from the domain virtual console.

`console` can utilize either IOSRAM or the internal management (I1 MAN) network for domain console communication. You can manually toggle the communication path by using the `~=` (tilde-equal sign) command. Doing so is useful if the network becomes inoperable, in which case the console sessions appears to be hung.

Many console sessions can be attached simultaneously to a domain, but only one console will have write permissions; all others will have read-only permissions. Write permissions are in either "locked" or "unlocked" mode.

Tilde Usage

In a domain console window, a tilde (~) that appears as the first character of a line is interpreted as an escape signal that directs console to perform some special action, as follows:

TABLE 1-1 Tilde Usage

Character	Description
~?	Status message
~.	Disconnects console session
~#	Breaks to OpenBoot PROM or kadb
~@	Acquires unlocked write permission. See option `-g`
~^	Releases write permission

TABLE 1-1 Tilde Usage *(Continued)*

Character	Description
~=	Toggles the communication path between the network and IOSRAM interfaces. You can use `~=` only in private mode (see `~*`).
~&	Acquires locked write permission; see option `-l` . You may issue this signal during a read-only or unlocked write session.
~*	Acquires locked write permission, terminates all other open sessions, and prevent new sessions from being opened; see option `-f` . To restore multiple-session mode, either release the lock or terminate this session.

`rlogin` also processes tilde-escape sequences whenever a tilde is seen at the beginning of a new line. If you need to send a tilde sequence at the beginning of a line and you are connected using `rlogin`, use two tildes (the first escapes the second for `rlogin`). Alternatively, do not enter a tilde at the beginning of a line when running inside of an `rlogin` window.

If you use a `kill -9` command to terminate a console session, the window or terminal in which the `console` command was executed goes into raw mode, and appears hung. Type `^j`, then `stty sane`, then `^j` to escape this condition,

In the domain console window, `vi`(1) runs properly and the escape sequences (tilde commands) work as intended only if the environment variable TERM has the same setting as that of the console window.

For example:

```
sc0:sms-user:> setenv TERM xterm
```

If you need to resize the window, type:

```
sc0:sms-user:> stty rows 20 cols 80
```

For more information on domain console, see “Domain Console” on page 114 and refer to the `console` man page.

Remote Console Session

In the event that a system controller hangs and that console cannot be reached directly, SMS provides the `smsconnectsc` command to remotely connect to the hung SC. This command works from either the main or spare SC. For more information and examples, refer to the `smsconnectsc` man page.

Sun Management Center

Sun Management Center for the Sun Fire 15K is an extensible monitoring and management tool that provides a system administrator with the ability to manage the Sun Fire 15K system. Sun Management Center integrates standard SNMP based management structures with new intelligent and autonomous agent and management technology based on the client/ server paradigm.

Sun Management Center is used as the GUI and SNMP manager/agent infrastructure for the Sun Fire 15K system. The features and functions of Sun Management Center are not covered in this manual. For more information, refer to the *Sun Management Center User's Guide*.

CHAPTER 2

SMS Security

This chapter provides a brief overview of security as it pertains to SMS and the Sun Fire 15K server system.

The Sun Fire 15K platform hardware can be partitioned into one or more environments capable of running separate images of the Solaris operating environment. These environments are called *dynamic system domains* (DSD)s or *domains*.

A domain is logically equivalent to a physically separate server. The Sun Fire 15K hardware has been designed to enforce strict separation of the domain environments. This means that, except for errors in hardware shared by multiple domains, no hardware error in one domain affects another. In order for domains to act like separate servers, Sun Fire software was designed and implemented to enforce strict domain separation.

SMS provides services to all DSDs. In providing those services, no data obtained from one client DSD is leaked into data observable by another. This is particularly true for sensitive data such as buffers of console characters (including administrator passwords) or potentially sensitive data such as I/O buffers containing client DSD-owned data.

SMS limits administrator privilege to control the extent of damage that can occur due to administrator error, as well as to limit the exposure to damage caused by an external attack on a system password.

Security

The security techniques that separate the data for the different domains revolve around implicit and explicit data labeling, so that it is clear which data can be mixed (data pertaining to a single domain) and which cannot (data from different domains).

The lowest level of security that promotes the data separation techniques using data labeling is equivalent to a B1 security rating from the National Security Agency defined rating scale.

Administration Models

SMS splits domain and platform administrative privileges. It is possible to assign separate administrative privileges for system management over each domain and for system management over the entire platform. There is also a subset of privileges available for platform operator and domain configurator-class users. Administrative privileges are granted so that audits can identify the individual who initiated any action.

SMS uses site-established Solaris user accounts and grants administrative privileges to those accounts through the use of Solaris *group* memberships. This allows a site considerable flexibility with respect to creating and consolidating default privileges. For example, by assigning the same Solaris group to represent the administrator privilege for more than one domain, groups of domains can be administered by one set of domain administrators.

It also allows the site considerable flexibility in assigning multiple administrative roles to individual administrators. A single user account with group membership in the union of all configured administrative privilege groups can be set up.

The platform administrator has control over the platform hardware. Limitations have been established with respect to controlling the hardware used by a running domain, but ultimately the platform administrator can shut down a running domain by powering off server hardware.

Each domain administrator has access to the Solaris console for that domain and the privilege to exert control over the software that runs in the domain or over the hardware assigned to the domain.

Levels of each type of administrative privilege provide a subset of status and monitoring privileges to a platform operator or domain configurator.

SMS provides an administrative privilege that grants access to functions provided exclusively for servicing the product in the field.

Administrative privilege configuration can be changed at will, by the superuser, using **`smsconfig -g`** without the need to stop or restart SMS.

SMS implements Solaris access control list (ACL) software to configure directory access for SMS groups using the `-a` and `-r` options of the `smsconfig` command. ACLs restrict access to platform and domain directories providing file system security. For information on ACLs, refer to the *Solaris 8 System Administration Guide, Volume 2*.

Platform Administrator Group

The group identified as the platform administrator (`platadmn`) group provides configuration control, a means to get environmental status, the ability to assign boards to domains, power control, and other generic service processor functions. In short, the platform administrator group has all platform privileges excluding domain control and access to installation and service commands (FIGURE 2-1).

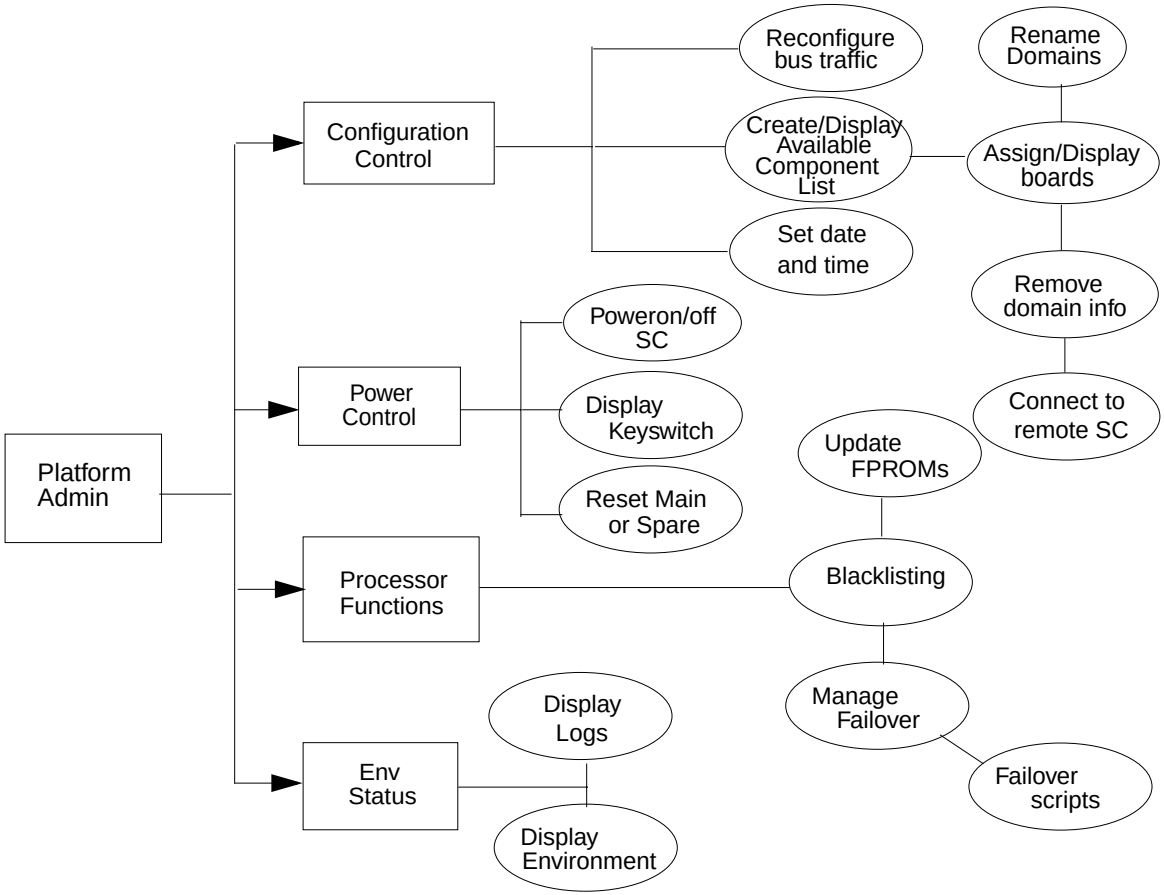


FIGURE 2-1 Platform Administrator Privileges

Platform Operator Group

The platform operator (`platoper`) group has a subset of platform privileges. This group has no platform control other than being able to perform power control. Therefore, this group is limited to platform power and status privileges (FIGURE 2-2).

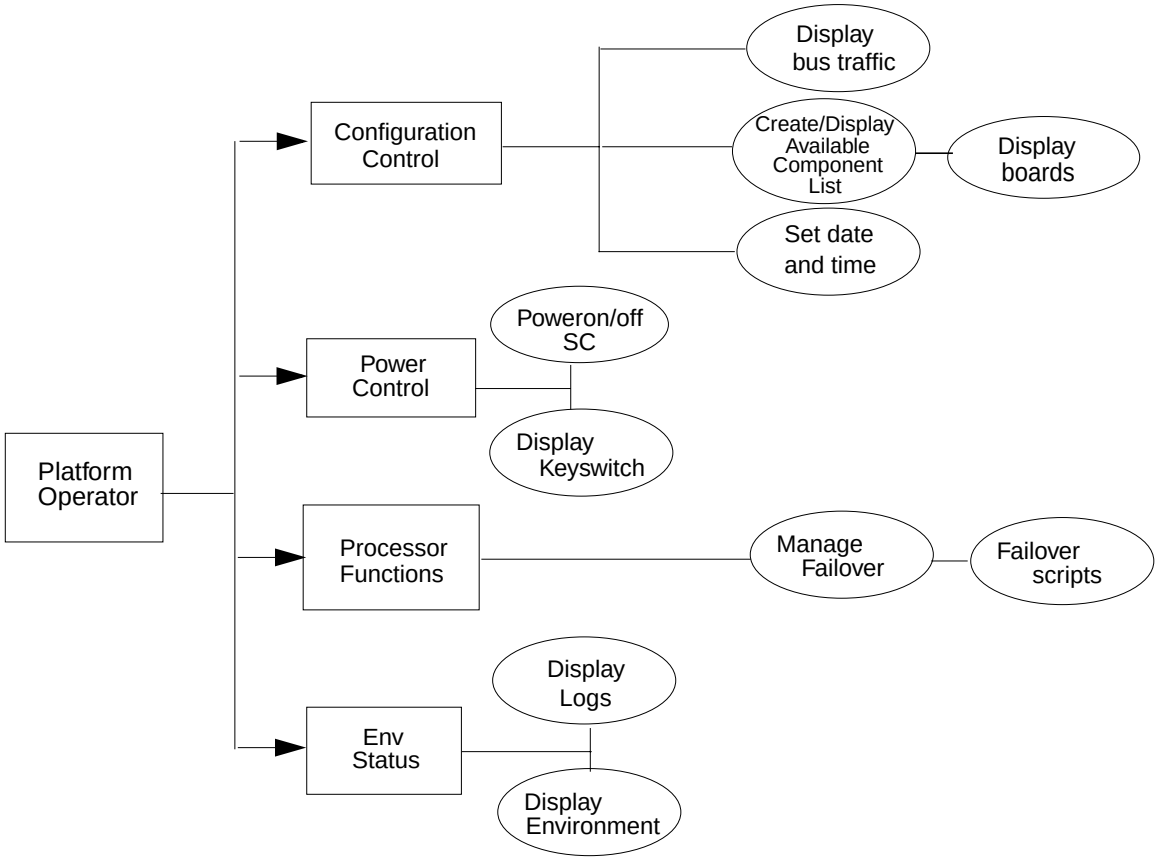


FIGURE 2-2 Platform Operator Privileges

Platform Service Group

The platform service (`platsvc`) group possesses platform service command privileges in addition to limited platform control and platform configuration status privileges (FIGURE 2-3).

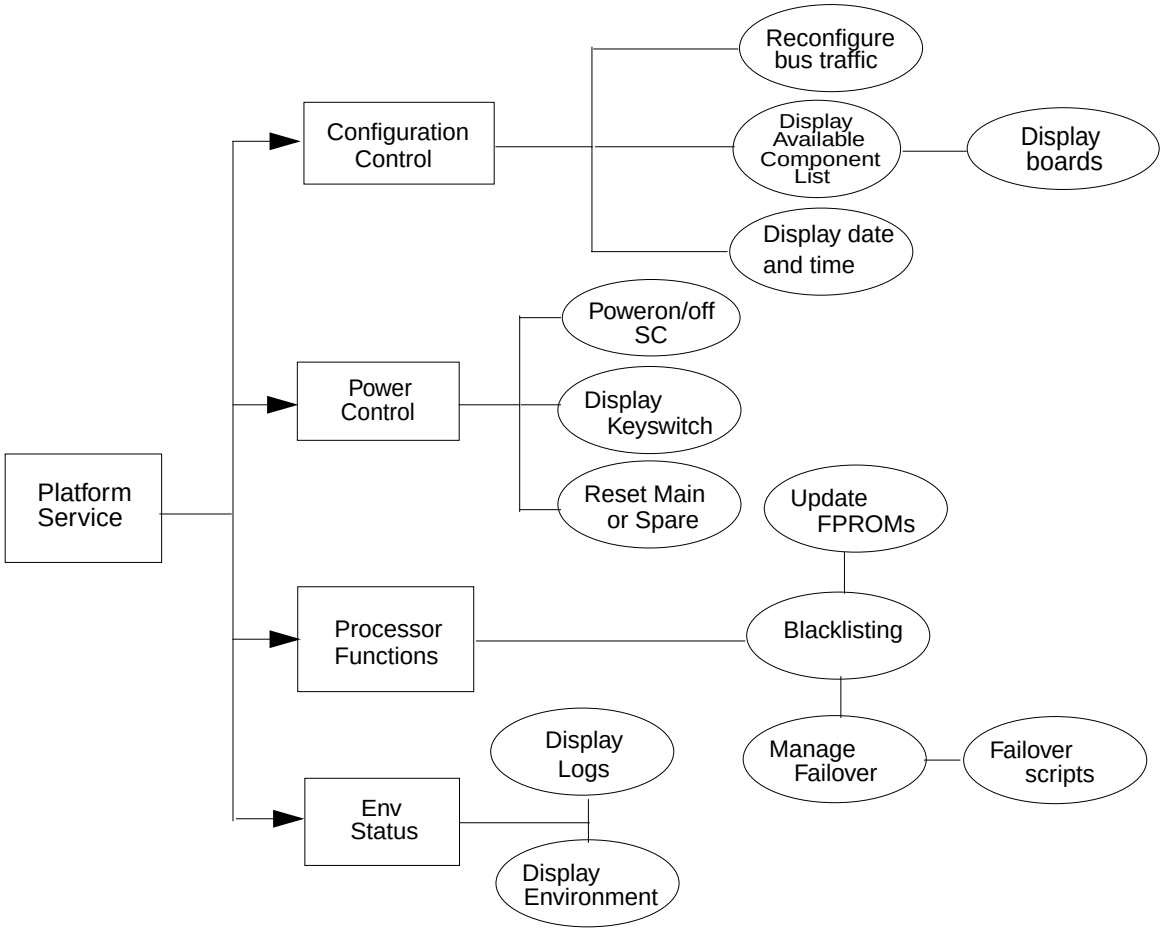
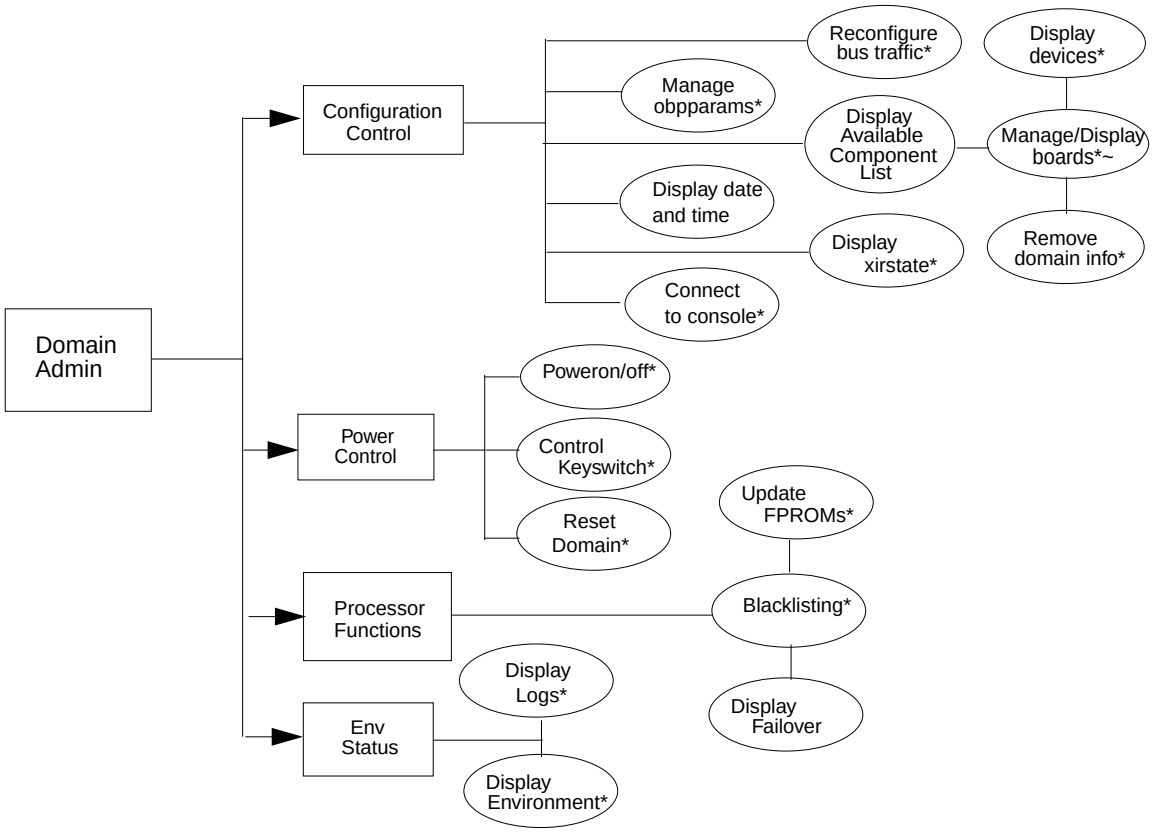


FIGURE 2-3 Platform Service Privileges

Domain Administrator Group

The domain administrator (dmn*[domain_id]*admn) group provides the ability to access the console of its respective domain as well as perform other operations that affect, directly or indirectly, the respective domain. Therefore, the domain administrator group can perform domain control, domain status, and console access, but cannot perform platform wide control or platform resource allocation (FIGURE 2-4).

There are 18 possible Sun Fire domains, A-R, identified by *domain_id*. Therefore, there are 18 Domain Administrator groups, each providing strict access over their respective domains.



* = For own domain only

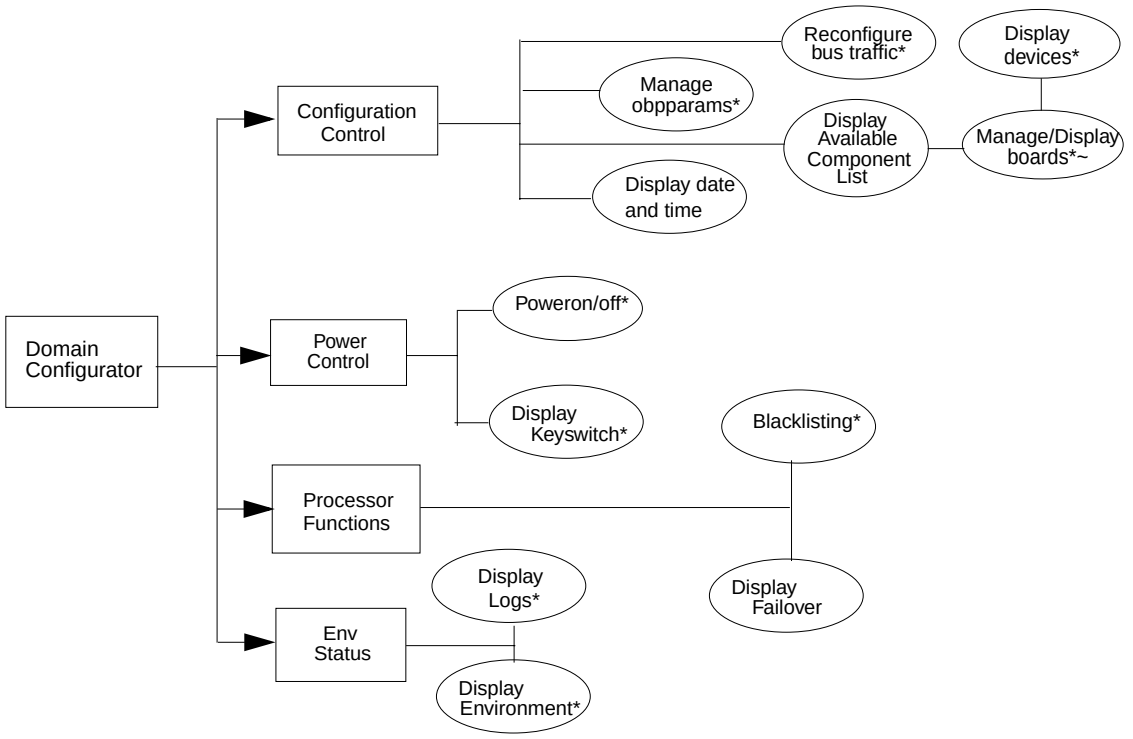
~ = board must be in the domain available component list

FIGURE 2-4 Domain Administrator Privileges

Domain Configuration Group

The domain configuration (`dmn`*[domain_id]*`rcfg`) group has a subset of domain administration group privileges. This group has no domain control other than being able to power control boards in its domain or (re)configure boards into or from its domain (FIGURE 2-5).

There are 18 possible Sun Fire domains identified by *domain_id*s. Therefore, there are 18 domain configuration groups, each allowing strict access over their respective domains.



* = For own domain only

~ = board must be in the domain available component list

FIGURE 2-5 Domain Configurator Privileges

Superuser Privileges

The superuser privileges are limited to installation, help, and status privileges (FIGURE 2-6).

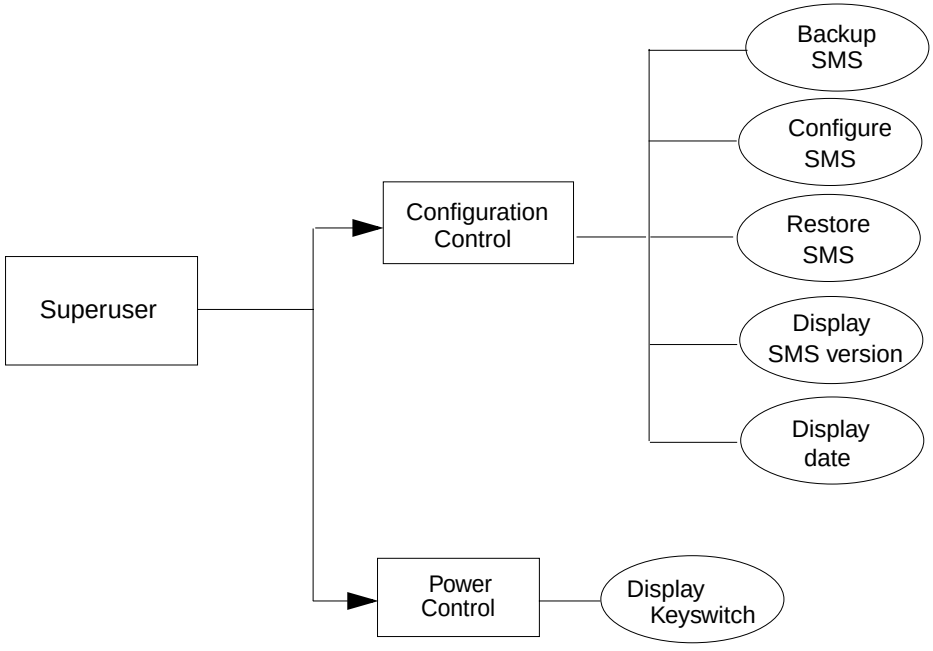


FIGURE 2-6 Superuser Privileges

All Privileges

The following is a list of all group privileges.

TABLE 2-1 All Group Privileges

Command	Group Privileges					
	Platform Administrator	Platform Operator	Domain Administrator	Domain Configurator	Platform Service	Superuser
`addboard`	A user with only platform administrator privileges can perform only the `-c` *assign*.	No	Users with only domain *X* administrator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	Users with only domain *X* configurator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	No	No
`addtag`	Yes	No	No	No	No	No
`cancelcmdsync`	Yes	Yes	Yes	Yes	Yes	No
`console`	No	No	Yes (for own domain)	No	No	No

TABLE 2-1 All Group Privileges *(Continued)*

Command	Group Privileges					
	Platform Administrator	Platform Operator	Domain Administrator	Domain Configurator	Platform Service	Superuser
`deleteboard`	A user with only platform administrator privileges can perform `-c` *unassign* only if the board(s) are in the *assign*ed state and not active in a running domain.	No	Users with only domain *X* administrator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	Users with only domain *X* configurator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	No	No
`deletetag`	Yes	No	No	No	No	No
`disablecomponent`	Yes (platform only)	No	Yes (for own domain)	Yes (for own domain)	No	No
`enablecomponent`	Yes (platform only)	No	Yes (for own domain)	Yes (for own domain)	No	No
`flashupdate`	Yes	No	Yes (for own domain)	No	No	No
`help`	Yes	Yes	Yes	Yes	Yes	Yes
`initcmdsync`	Yes	Yes	Yes	Yes	Yes	No

TABLE 2-1 All Group Privileges *(Continued)*

Command	Group Privileges					
	Platform Administrator	Platform Operator	Domain Administrator	Domain Configurator	Platform Service	Superuser
`moveboard`	A user with only platform administrator privileges can perform the `-c` *assign* only if the board is in the *assign*ed state and not active in the domain the board is being removed from.	No	Users must belong to both domains affected. If the board(s) are not already assigned to the domain the board(s) are being moved into, the board(s) must be in the available component list of that domain.	Users must belong to both domains affected. If the board(s) are not already assigned to the domain the board(s) is being moved into, the board(s) must be in the available component list of that domain.	No	No
`poweron`	Yes	No	Yes (for own domain)	Yes (for own domain)	No	No
`poweroff`	Yes	No	Yes (for own domain)	Yes (for own domain)	No	No
`rcfgadm`	A user with only platform administrator privileges can perform `-x` *assign* . The user can execute `-x` *unassign* only if the board(s) are in the *assign*ed state and not active in a running domain.	No	Users with only domain *X* administrator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	Users with only domain *X* configurator privileges can execute this command on their respective domain. If the board(s) are not already assigned to the domain, the board(s) must be in the available component list of that domain.	No	No
`reset`	No	No	Yes (for own domain)	No	No	No
`resetsc`	Yes	No	No	No	No	No
`runcmdsync`	Yes	Yes	Yes	Yes	Yes	No

TABLE 2-1 All Group Privileges *(Continued)*

Command	Group Privileges					
	Platform Administrator	Platform Operator	Domain Administrator	Domain Configurator	Platform Service	Superuser
savecmdsync	Yes	Yes	Yes	Yes	Yes	No
setbus	Yes	No	Yes (for own domain)	Yes (for own domain)	No	No
setdatasync	Yes	Yes	Yes	Yes	Yes	No
setdate	Yes	No	Yes (for own domain)	No	No	No
setdefaults	Yes	No	Yes (for own domain)	No	No	No
setfailover	Yes	No	No	No	No	No
setkeyswitch	No	No	Yes (for own domain)	No	No	No
setobpparams	No	No	Yes (for own domain)	Yes (for own domain)	No	No
setupplatform	Yes	No	No	No	No	No
showboards	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No
showbus	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No
showcmdsync	Yes	Yes	Yes	Yes	Yes	No
showcomponent	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No
showdatasync	Yes	Yes	Yes	Yes	Yes	No
showdate	Yes (platform only)	Yes (platform only)	Yes (for own domain)	Yes (for own domain)	Yes (platform only)	No
showdevices	No	No	Yes (for own domain)	Yes (for own domain)	No	No
showenvironment	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No
showfailover	Yes	Yes	No	No	Yes	No
showkeyswitch	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No

TABLE 2-1 All Group Privileges *(Continued)*

Command	Group Privileges					
	Platform Administrator	Platform Operator	Domain Administrator	Domain Configurator	Platform Service	Superuser
showlogs	Yes (platform only)	Yes (platform only)	Yes (for own domain)	Yes (for own domain)	Yes (platform only)	No
showobpparams	No	No	Yes (for own domain)	Yes (for own domain)	No	No
showplatform	Yes	Yes	Yes (for own domain)	Yes (for own domain)	Yes	No
showxirstate	No	No	Yes (for own domain)	No	No	No
smsbackup	No	No	No	No	No	Yes
smsconfig	No	No	No	No	No	Yes
smsconnectsc	Yes	No	No	No	No	No
smsrestore	No	No	No	No	No	Yes
smsversion	No	No	No	No	No	Yes

Network Connections for Administrators

The nature of the Sun Fire 15K physical architecture, with an embedded system controller, as well as the supported administrative model (with multiple administrative privileges, and hence multiple administrators) dictates that an administrator utilize a remote network connection (from a workstation) to access SMS command interfaces to manage the Sun Fire 15K system.

Since the administrators will be providing information to verify their identity (passwords) and may possibly need to display sensitive data, it is important that the remote network connection be secure. Physical separation of the administrative networks provides security on the Sun Fire 15K system. Multiple external physical network connections are available on each SC. SMS supports up to six external network communities. As of this release, two external physical network connections are supported.

For more information, see “Management Network Services” on page 114.

CHAPTER 3

SMS Internals

SMS operations are generally performed by a set of daemons and commands. This chapter provides an overview of how SMS works and describes the SMS daemons, processes, commands, and system files. For more information about daemons, commands, and system files, refer to the *System Management Services (SMS) 1.2 Reference Manual*.

Caution – Changes made to files in `/opt/SUNWSMS` can cause serious damage to the system. Only very experienced system administrators should risk changing the files described in this chapter.

Startup Flow

The events that take place when the SMS boots are as follows:

1. User powers on the Sun Fire 15K (CPU / disk, and CD-ROM). The Solaris operating environment on the SC boots automatically.

2. During the boot process, the `/etc/init.d/sms` script is called. This script, for security reasons, disables forwarding, broadcast and multicasting over the MAN network. It then starts the SMS software by invoking a background process, which starts and monitors `ssd`. `ssd` is the SMS startup daemon responsible for starting and monitoring all the SMS daemons and servers.

3. `ssd`(1M) in turn invokes: `mld`, `pcd`, `hwad`, `tmd`, `dsmd`, `esmd`, `mand`, `osd`, `dca`, `efe`, and `smnptd`.

For more information, see “SMS Daemons” on page 28, “Message Logging” on page 145. For `efe`, refer to the *Sun Management Center User’s Guide*.

4. Once the daemons are running, you can use SMS commands such as `console`.

SMS startup can take a few minutes during which time any commands run will return an error message indicating that SMS has not completed startup. The message “SMS software start-up complete” is posted to the platform log when startup is complete and can be viewed using the `showlogs`(1M) command.

SMS Daemons

The SMS 1.2 daemons play a central role on the Sun Fire 15K system. Daemons are persistent processes that provide SMS services to clients using an API.

Note – SMS daemons are started by `ssd` and should *not* be started manually from the command line.

Daemons are always running, initiated at system startup, and restarted whenever necessary. Each daemon is fully described in its corresponding man page (with the exception of `efe`, which is referenced separately in the Sun Management Center documentation).

This section looks at the SMS daemons, their relationship to one another, and includes which CLIs (if any) access them.

FIGURE 3-1 illustrates the Sun Fire 15K client server overview.

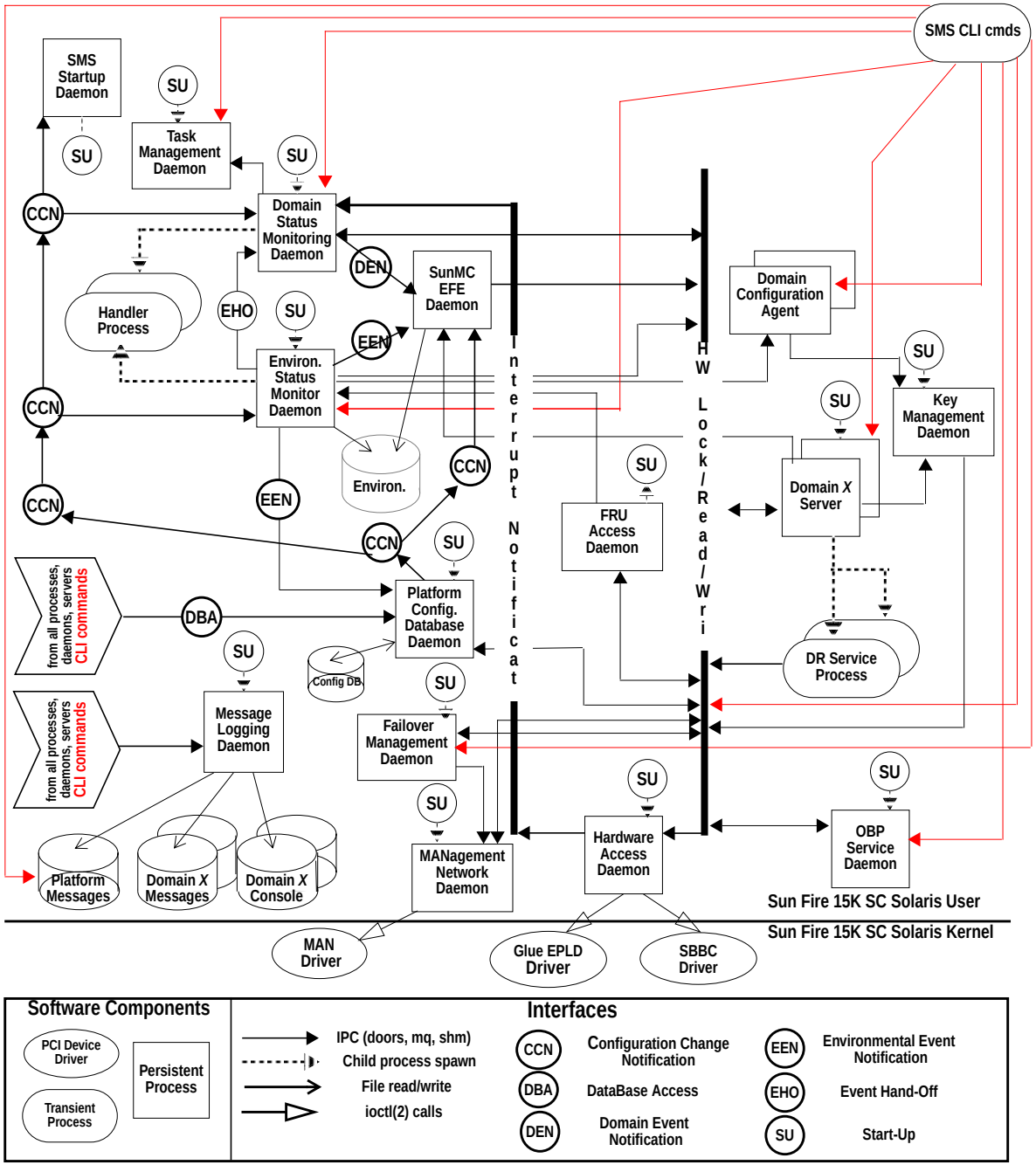


FIGURE 3-1 Sun Fire 15K Client Server Overview

Note – The domain *X* server (dxs) and domain configuration agent (dca), while not daemons, are essential server processes and included in the following table and section. There is an instance of dxs and dca running for each domain up to eighteen instances.

TABLE 3-1 Daemons and Processes

Daemon Name	Description
dca	The domain configuration agent provides a communication mechanism between the dca on the system controller and the domain configuration server (dcs) on the specified domain. There is an instance of dca for every domain up to 18 domains.
dsmd	The domain status monitoring daemon monitors domain status and OS heartbeat for up to 18 domains.
dxs	The domain *X* server provides software support for a domain. There is an instance of dxs for each domain up to 18 domains.
efe	The event front end daemon is part of Sun Management Center and acts as an intermediary between the Sun Management Center agent and SMS. It is not covered further in this manual. For more information on efe, refer to the *Sun Management Center User's Guide*
esmd	The environmental status monitoring daemon monitors system cabinet environmental conditions.
fomd	The failover monitoring daemon detects faults on the local and remote SCs and takes appropriate action (directing/taking a failover.)
frad	The FRU access daemon provides the mechanism by which SMS daemons can access any FRUs SEEPROM on the Sun Fire 15K system.
hwad	The hardware access daemon provides hardware access to SMS daemons and a mechanism for all daemons to exclusively access, control, monitor and configure the hardware.
kmd	The key management daemon manages the IPSec security associations (SAs) needed to secure the communication between the system controller (SC) and servers running on a domain.
mand	The management network daemon supports the MAN drivers, providing required network configuration.
mld	The messages logging daemon provides message logging support for the platform and domains.
osd	The OpenBoot PROM server daemon provides software support for OpenBoot PROM.
pcd	The platform configuration database daemon provides and manages controlled access to platform, domain, and system board configuration data.
ssd	The SMS startup daemon starts, stops, and monitors all the key SMS daemons and servers.
tmd	The task management daemon provides task management services, such as scheduling for SMS.

Domain Configuration Administration

dca(1M) supports remote dynamic reconfiguration (DR) by enabling communication between applications and the domain configuration server (dcs) running on a Solaris 8 domain. One dca per domain runs on the SC. Each dca communicates with its dcs over the Management Network (MAN).

ssd(1M) starts dca when the domain is brought up. ssd restarts dca if it is killed while the domain is still running. dca is terminated when the domain is shut down.

dca is an SMS application that waits for dynamic reconfiguration requests. When a DR request arrives, dca creates a dcs session. Once a session is established, dca forwards the request to dcs. dcs attempts to honor the DR request and sends the results of the operation to the dca. Once the results have been sent, the session is ended. The remote DR operation is complete when dca returns the results of the DR operation.

FIGURE 3-2 illustrates the DCA client server relationship to the SMS daemons and CLIs.

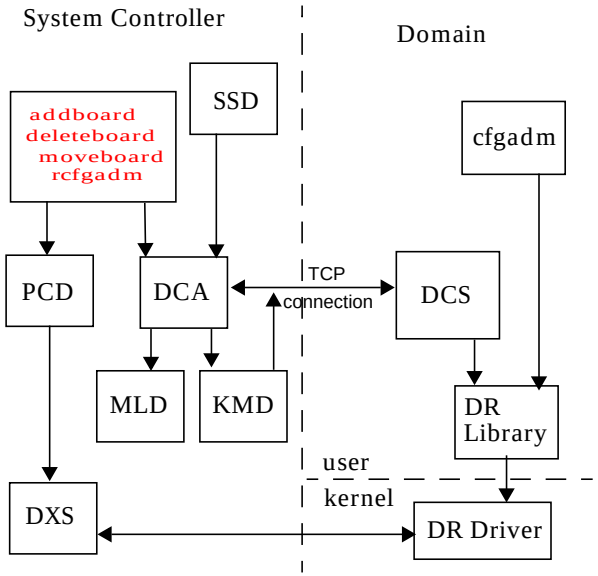


FIGURE 3-2 DCA Client Server Relationships

Domain Status Monitoring Daemon

`dsmd`(1M) monitors domain state signatures, CPU reset conditions and Solaris heartbeat for up to 18 domains. It also handles domain stop events related to hardware failure.

`dsmd` detects timeouts that can occur in reboot transition flow and panic transition flow, and handles various domain hung conditions.

`dsmd` notifies the domain *X* server (`dxs`(1M)) and Sun Management Center of all domain state changes and automatically recovers the domain based on the domain state signature, domain stop events, and automatic system recovery (ASR) Policy. ASR Policy consists of those procedures which restore the system to running all properly configured domains after one or more domains have been rendered inactive. This can be due to software or hardware failures or to unacceptable environmental conditions. For more information, see “Automatic System Recovery (ASR)” on page 95 and “Domain Stop Events” on page 160.

FIGURE 3-3 illustrates DSMD client server relationship to the SMS daemons and CLIs.

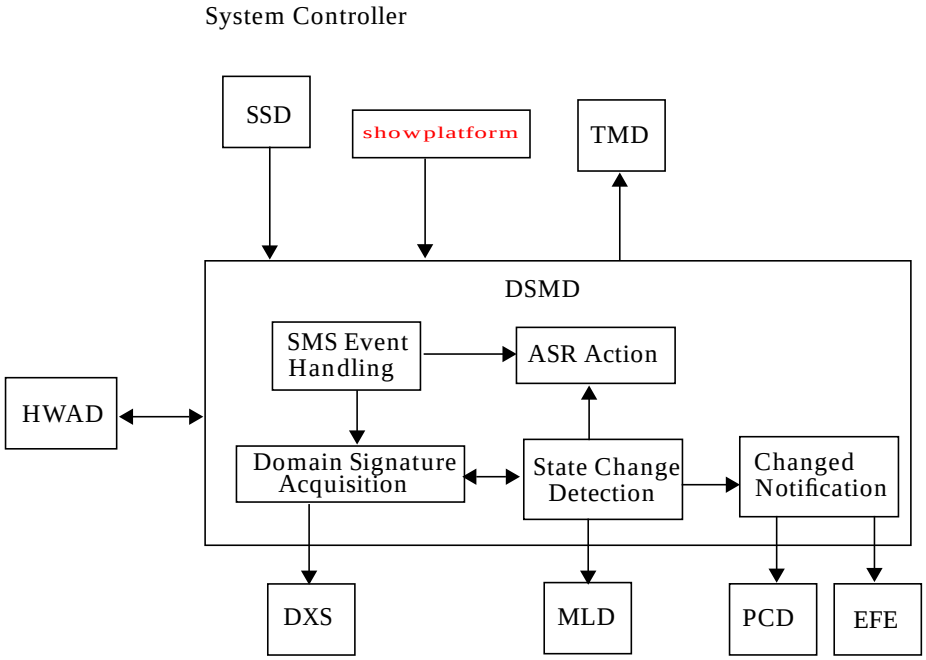


FIGURE 3-3 DSMD Client Server Relationships

Domain *X* Server

`dxs`(1M) provides software support for a running domain. This support includes virtual console functionality, dynamic reconfiguration support, and HPCI support. `dxs` handles domain driver requests and events. The virtual console functionality allows one or more users running the `console` program to access the domain's virtual console. `dxs` acts as a link between SMS console applications and the domain virtual `console` drivers.

A Sun Fire 15K system can support up to 18 different domains. Each domain may require software support from the SC, and `dxs` provides that support. The following domain related projects require `dxs` support:

- DR
- HPCI
- Virtual console

There is one domain *X* server for each Sun Fire 15K domain. `dxs` is started by `ssd` for every active domain and terminated when the domain is shut down.

FIGURE 3-5 illustrates DXS client server relationship to the SMS daemons.

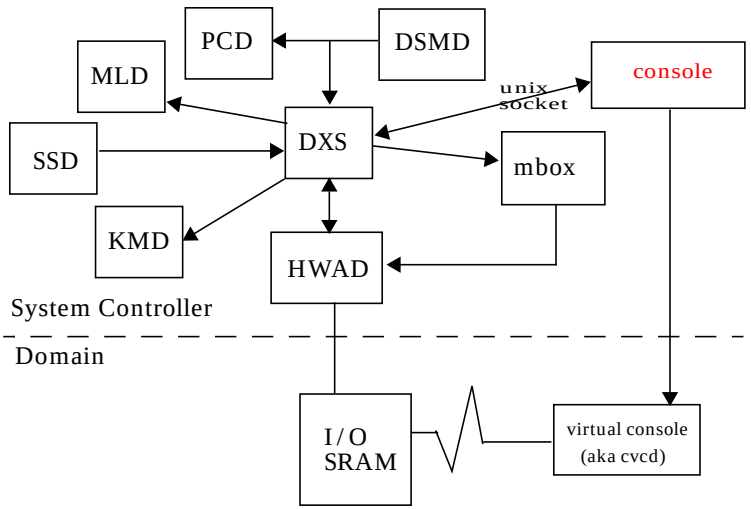


FIGURE 3-4 DXS Client Server Relationships

Environmental Status Monitoring Daemon

`esmd`(1M) monitors system cabinet environmental conditions, for example. voltage, temperature, fan tray, and power supply. `esmd` logs abnormal conditions and takes action to protect the hardware, if necessary.

See “Environmental Events” on page 156 for more information on `esmd`.

FIGURE 3-5 illustrates ESMD client server relationship to the SMS daemons.

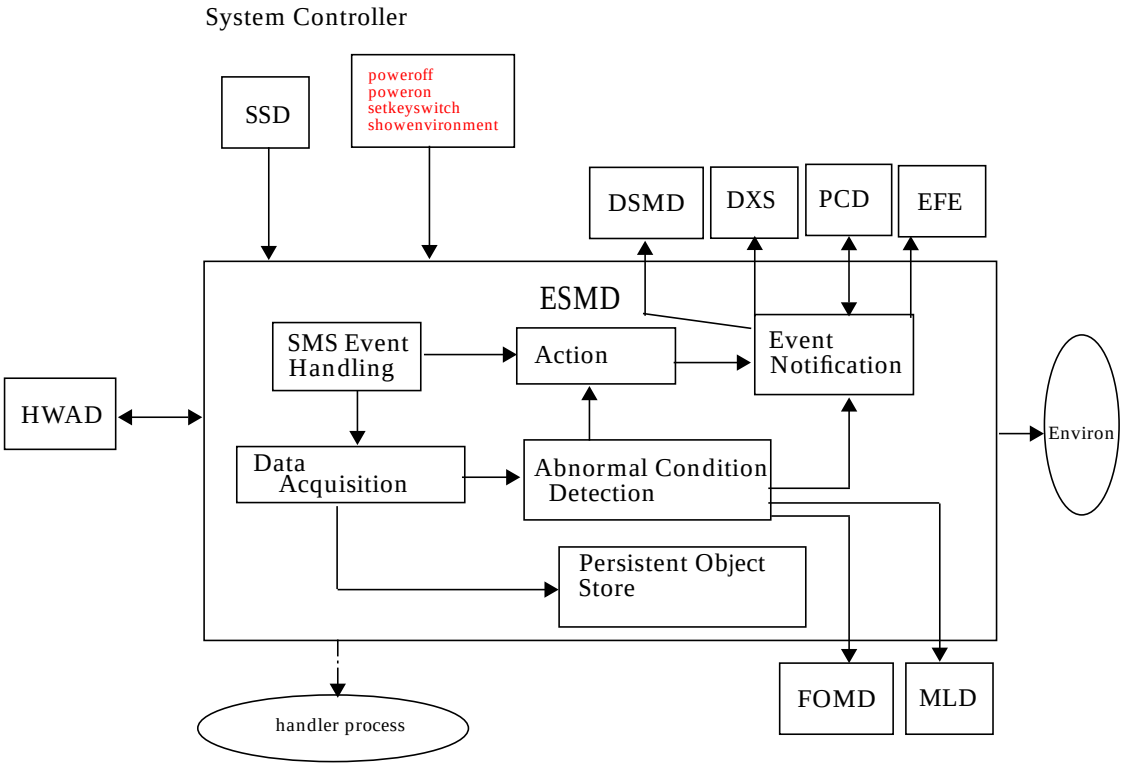


FIGURE 3-5 ESMD Client Server Relationships

Failover Management Daemon

`fomd`(1M) is the core of the SC failover mechanism. `fomd` detects faults on the local and remote SCs and takes the appropriate action (directing a failover/takeover).

`fomd` ensures that important configuration data is kept synchronized between both SCs. `fomd` runs on both the master and spare SC.

FIGURE 3-6 illustrates FOMD client server relationship to the SMS daemons.

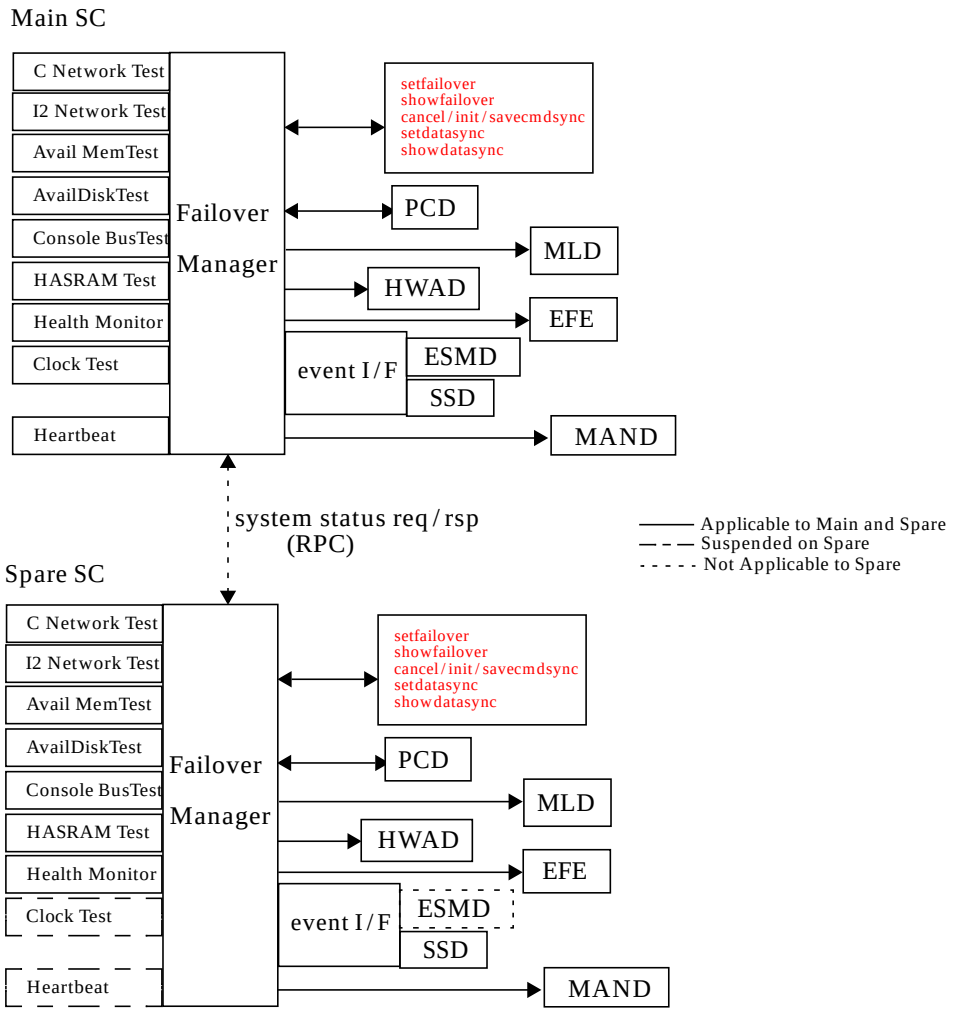


FIGURE 3-6 FOMD Client Server Relationships

FRU Access Daemon

`frad`(1M) is the field replaceable unit (FRU) access daemon for SMS. `frad` provides controlled access to any SEEPROM within the Sun Fire 15K platform that is accessible by the SC. `frad` supports dynamic FRUID which provides improved FRU data access.

`frad` is started by `ssd`.

FIGURE 3-7 illustrates FRAD client server relationship to the SMS daemons.

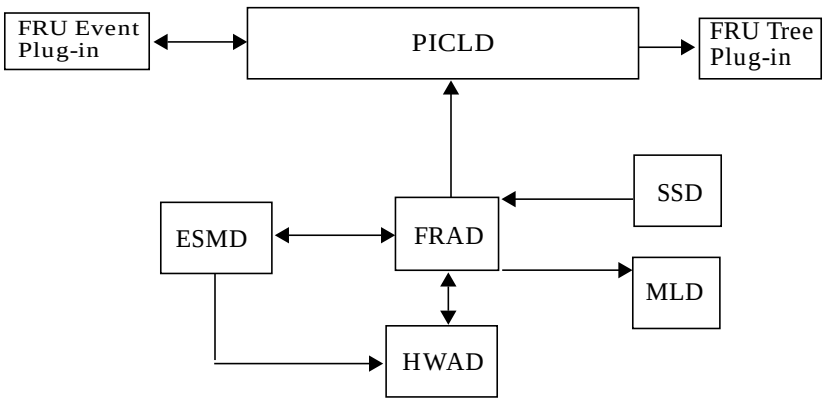


FIGURE 3-7 FRAD Client Server Relationships

Hardware Access Daemon

hwad(1M) provides hardware access to SMS daemons and a mechanism for all daemons exclusively to access, control, monitor, and configure the hardware.

hwad runs in either main or spare mode when it comes up. The failover daemon (fomd(1M)) determines which role hwad will play.

At startup, hwad opens all the drivers (sbbc, echip, gchip, and consbus) and uses ioctl(2) calls to interface with them. It reads the contents of the device presence register to identify the boards present in the system and makes them accessible to the clients. hwad also configures the local system clock and sets the clock source for each board present in the system.

IOSRAM and Mbox interfaces are also provided by hwad. This helps communication between the SC and the domain. For dynamic reconfiguration (DR), hwad directs communication to the IOSRAM (tunnel switch).

For darb interrupts, hwad notifies the dsmd(1M) if there is a dstop or rstop. It also notifies related SMS daemon(s) depending on the type of the Mbox interrupt that occurs.

hwad detects and recovers console bus and jtag errors.

Hardware access to the Sun Fire 15K system on the SC is done either by going through the PCI bus or console bus. Through the PCI bus you can access:

- SC BBC internal registers
- SC local Jtag
- Global I2C devices for clock and power control/status.

Through the Console bus you can access:

- Various ASICs internal registers
- Read/write chips
- Local I2C devices on various boards for temperature and chip level power control/status.

FIGURE 3-8 illustrates HWAD client server relationship to the SMS daemons and CLIs.

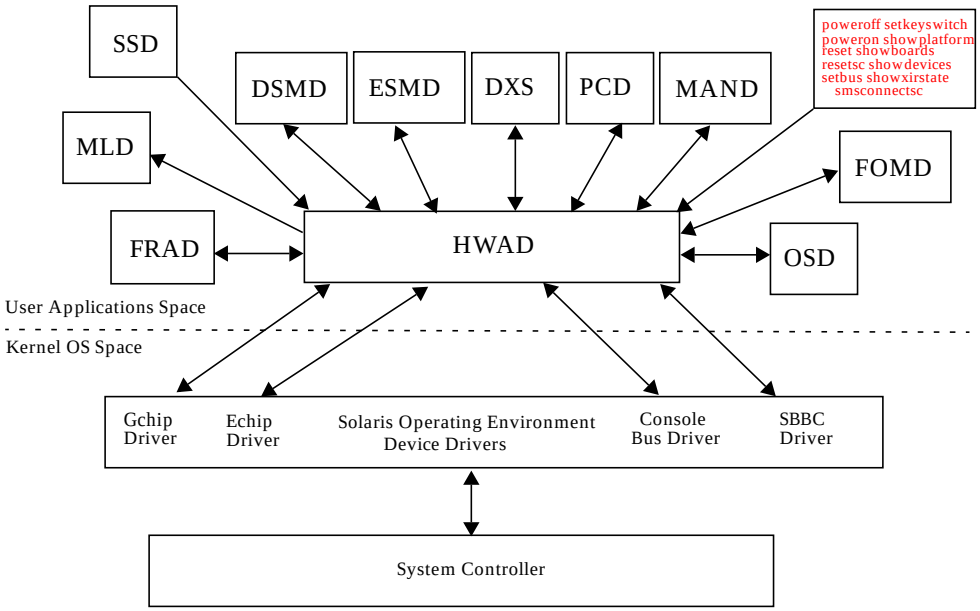


FIGURE 3-8 HWAD Client Server Relationships

Key Management Daemon

The key management daemon provides a mechanism for managing security for socket communications between the SC and the domains.

The current default configuration includes authentication policies for the dca(1M) and dxs(1M) clients on the SC, which connect to the dcs(1M) and cvcd(1M) servers on a domain.

kmd(1M) manages the IPSec security associations (SAs) needed to secure the communication between the SC and servers running on a domain.

kmd manages per-socket policies for connections initiated by clients on the SC to servers on a domain.

At system startup, kmd creates a domain interface for each domain that is active. An active domain has both a valid IOSRAM and is running the Solaris operating environment. Domain change events can trigger creation or removal of a domain kmd interface.

kmd manages shared policies for connections initiated by clients on the domain to servers on the SC. The kmd policy manager reads a configuration file and stores policies used to manage security associations. A request received by kmd is compared to the current set of policies to ensure that it is valid and to set various parameters for the request.

Static global policies are configured using ipsecconf(1M) and associated data file (/etc/inet/ipsecinit.conf). Global policies are used for connections initiated from the domains to the SC. Corresponding entries are made in the kmd configuration file. Shared security associations for domain to SC connections are created by kmd when the domain becomes active.

Note – In order to work properly, policies created by ipsecconf and kmd must match.

The kmd configuration file is used for both SC-to-domain and domain-to-SC initiated connections. The kmd configuration file resides in /etc/opt/SUNWSMS/config/kmd_policy.conf.

The format of the kmd configuration files is as follows:

```
dir:d_port:protocol:sa_type:aut_alg:encr_alg:domain:login
```

where:

- dir is identified using the sctodom or domtosc strings.
- d_port is the destination port
- protocol is identified using the tcp or udp strings.

- `sa_type` is the security association type. Valid choices are the `ah` or `esp` strings.
- `auth_alg` is the authentication algorithm. The authentication algorithm is identified using the `none` or `hmac-md5` strings or leaving the field blank.
- `encr_alg` is the encryption algorithm. The encryption algorithm is identified using the `none` or `des` strings or leaving the field blank.
- `domain` is the *domain_id* associated with the domain. Valid *domain_id*s are integers 0–17, space Using a space in the *domain_id* field defines a policy that applies to all domains. A policy for a specific domain overrides a policy applied to all domains.
- `login_name` is the login name of the user affected by the policy. Currently this includes `sms-dxs`, `sms-dca`, and `sms-mld`.

For example:

```
# Copyright (c) 2001 by Sun Microsystems, Inc.
# All rights reserved.
#
#
# This is the policy configuration file for the SMS Key Management Daemon.
# The policies defined in this file control the desired security for socket
# communications between the system controller and domains.
#
# The policies defined in this file must match the policies defined on the
# corresponding domains. See /etc/inet/ipsecinit.conf on the sun fire 15K domain.
# See also the ipsec(7P), ipsecconf(1M) and sckmd(1M) man pages.
#
# The fields in the policies are a tuple of eight fields separated by the pipe '|' #
character.
#
#<dir>|<d_port>|<protocol>|<sa_type>|<auth_alg>|<encr_alg>|<domain>|<login>|
#
# <dir>         --- direction to connect from. Values: sctodom, domtosc
# <d_port>      --- destination port
# <protocol>    --- protocol for the socket. Values: tcp, udp
# <sa_type>     --- security association type. Values: ah, esp
# <auth_alg>    --- authentication algorithm. Values: none, md5, sha1
# <encr_alg>    --- encryption algorithm. Values: none, des, 3des
# <domain>      --- domain id. Values: integers 0 - 17, space
#                   A space for the domain id defines a policy which applies
#                   to all domains. A policy for a specific domain overrides
#                   a policy which applied to all domains.
# <login>       --- login name. Values: Any valid login name
#
# ------------------------------------------------------------------------
sctodom|665|tcp|ah|md5|none| |sms-dca|
sctodom|442|tcp|ah|md5|none| |sms-dxs|
```

FIGURE 3-9 illustrates KMD client server relationship to the SMS daemons.

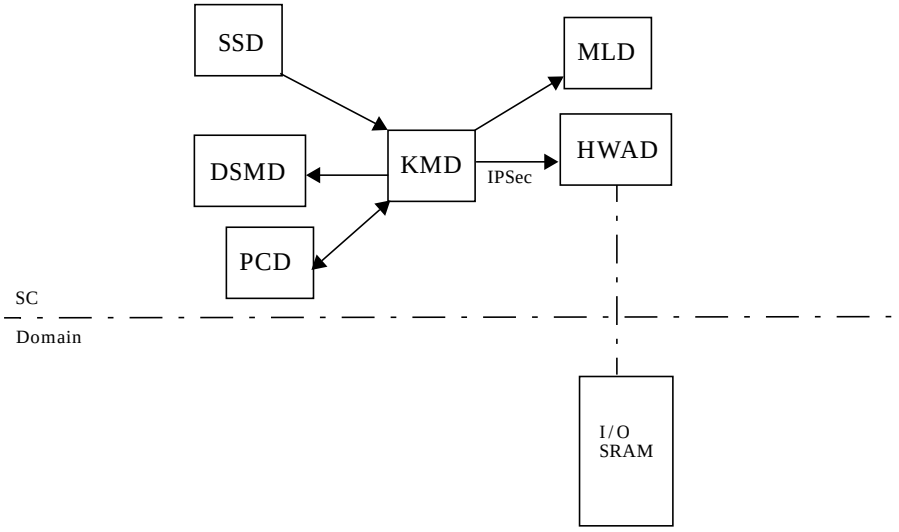


FIGURE 3-9 KMD Client Server Relationships

Management Network Daemon

mand(1M) supports the Management Network (MAN). See “Management Network Services” on page 114 mand runs in either main or spare mode when it comes up. The failover daemon (fomd(1M)) determines which role mand plays.

At system startup, mand creates the mapping between *domain_tag* and IP address in the platform configuration database (pcd), and configures the SC-to-SC private network. This information is obtained from the file /etc/opt/SUNWSMS/config/MAN.cf, which is created by the smsconfig(1M) command. mand then obtains domain configuration information from the pcd and programs the scman(7d) driver accordingly. After initializing the pcd and the scman driver, mand registers for domain keyswitch events, tracks changes in domain active board lists, tracks active Ethernet information from the dman(7d) driver and updates the scman driver, as appropriate.

mand also communicates system startup MAN information to each domain when the domain is powered on (setkeyswitch on). This information includes Ethernet and MAN IP addressing information. This information is used during the initial software installation on the domain.

FIGURE 3-10 illustrates MAND client server relationship to the SMS daemons.

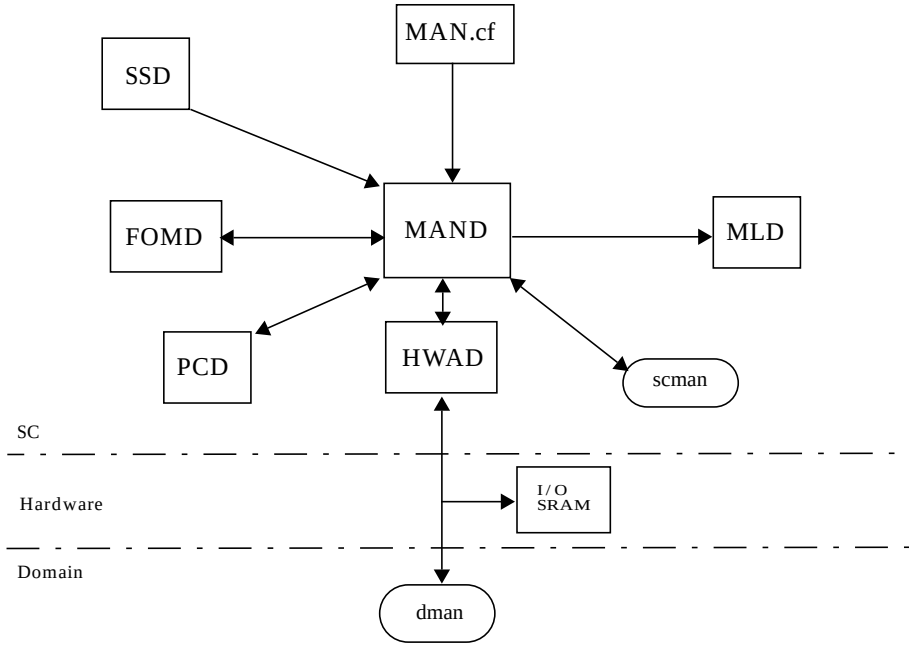


FIGURE 3-10 MAND Client Server Relationships

Message Logging Daemon

The message logging daemon, `mld`, captures the output of all other SMS daemons and processes. `mld` supports three configuration directives: File, Level, and Mode, in the `/var/opt/SUNWSMS/adm/.logger` file.

- File—Specifies the default output locations for the message files. The default is `msgdaemon` and should *not* be changed.

  Platform messages are stored on the SC in `/var/opt/SUNWSMS/adm/platform/messages`

  Domain messages are stored on the SC in `/var/opt/SUNWSMS/adm/`*domain_id*`/messages`

  Domain `console` messages are stored on the SC in`/var/opt/SUNWSMS/adm/`*domain_id*`/console`

  Domain `syslog` messages are stored on the SC in`/var/opt/SUNWSMS/adm/`*domain_id*`/syslog`.

- Level—Specifies the minimum level necessary for a message to be logged. The supported levels are `NOTICE`, `WARNING`, `ERR`, `CRIT`, `ALERT`, and `EMERG`. The default level is `NOTICE`.
- Mode—Specifies the verbosity of the messages. Two modes are available: `verbose` and `terse`. The default is `verbose`.

`mld` monitors the size of each of the message log files. For each message log type, `mld` keeps up to ten message files at a time, x.0 though x.9. For more information on log messages, see “Message Logging” on page 145

FIGURE 3-11 illustrates MLD client server relationship to the SMS daemons and CLIs.

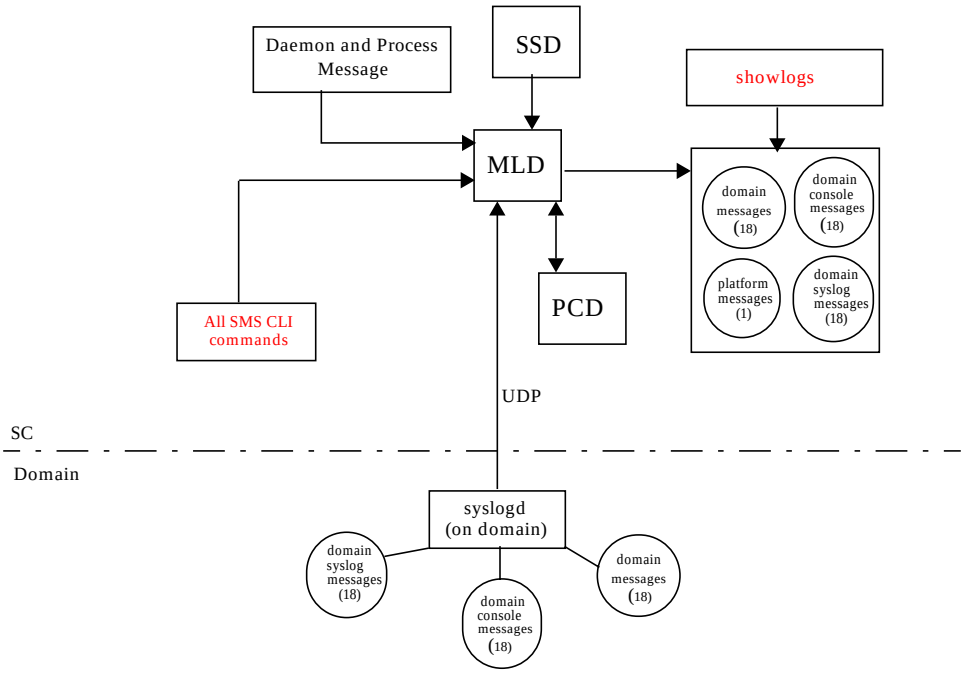


FIGURE 3-11 MLD Client Server Relationships

OpenBoot PROM Support Daemon

`osd`(1M) provides support to the OpenBoot PROM Process running on a domain. `osd` and OpenBoot PROM communication is through a mailbox that resides on the domain. The `osd` daemon monitors the OpenBoot PROM mailbox. When the OpenBoot PROM writes requests to the mailbox, `osd` executes the requests accordingly.

`osd` runs at all times on the SC even if there are no domains configured. `osd` provides virtual TOD service, virtual NVRAM, and virtual REBOOTINFO for OpenBoot PROM and an interface to `dsmd`(1M) to facilitate auto-domain recovery. `osd` also provides an interface for the following commands: `setobpparams`(1M), `showobpparams`(1M), `setdate`(1M) and `showdate`(1M). See also Chapter 4 “SMS Configuration”.

`osd` is a trusted daemon in that it will not export any interface to other SMS processes. It exclusively reads and writes from and to all OpenBoot PROM mailboxes. There is one OpenBoot PROM mailbox for each domain.

`osd` has two main tasks; to maintain its current state of the domain configuration, and to monitor the OpenBoot PROM mailbox.

FIGURE 3-12 illustrates OSD client server relationship to the SMS daemons and CLIs.

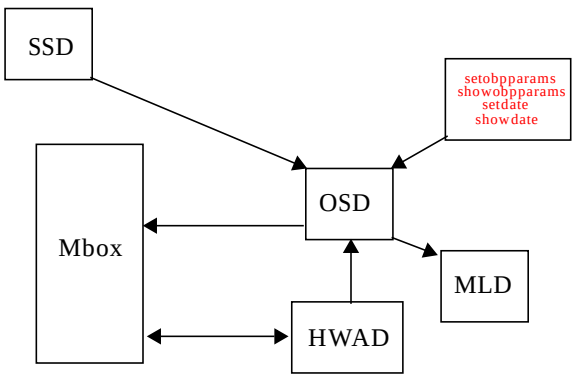


FIGURE 3-12 OSD Client Server Relationships

Platform Configuration Database Daemon

pcd(1M) is a Sun Fire 15K system management daemon that runs on the SC with primary responsibility for managing and providing controlled access to platform and domain configuration data.

pcd manages an array of information that describes the Sun Fire system configuration. In its physical form, the database information is a collection of flat files, each file appropriately identifiable by the information contained within it. All SMS applications that want to access the database information must go through pcd.

In addition to managing platform configuration data, pcd is responsible for platform configuration change notifications. When pertinent platform configuration changes occur within the system, the pcd sends out notification of the changes to clients who have registered to receive the notification.

FIGURE 3-13 illustrates PCD client server relationship to the SMS daemons and CLIs.

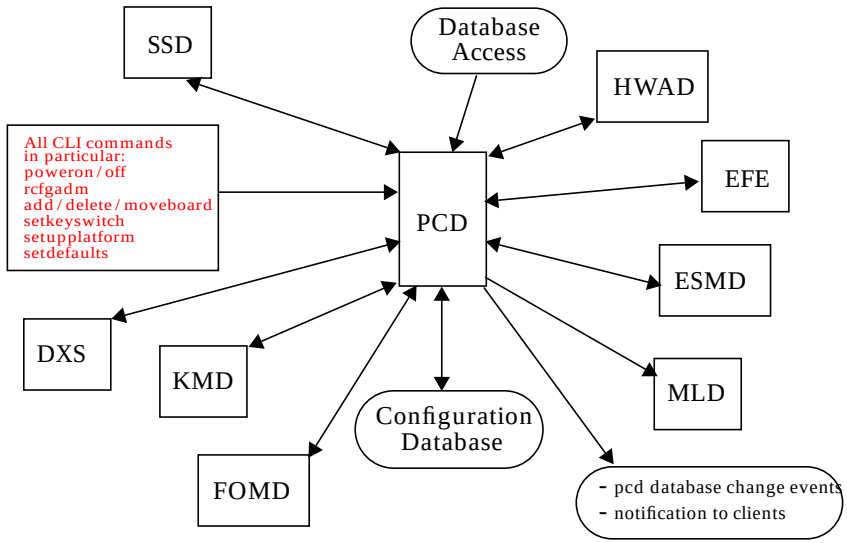


FIGURE 3-13 PCD Client Server Relationships

Platform Configuration

The following information uniquely identifies the platform:

- Platform type
- Platform name
- Rack ID
- Cacheable Address Slice Map
- System clock frequency
- System clock type
- SC IP address
- SC0 to SCI IP address
- SC1 to SC0 IP address
- SC to SC IP netmask

Domain Configuration

The following information is domain related:

- domain_id
- domain_tag
- OS version (currently not used)
- OS type (currently not used)
- Available component list
- Assigned board list
- Active board list
- Golden IOSRAM I/O board
- Virtual keyswitch setting for a domain
- Active Ethernet I/O board
- Domain creation time
- Domain dump state
- Domain bringup priority
- IP host address
- Host name
- Host netmask
- Host broadcast address
- Virtual OpenBoot PROM address
- Physical OpenBoot PROM address

System Board Configuration

The following information is related to system boards:

- Expander position
- Slot position
- Board type
- Board state
- DomainID assigned to board
- Available component list state
- Board test status
- Board test level
- Board memory clear state

SMS Startup Daemon

`ssd`(1M) is responsible for starting and maintaining all SMS daemons and domain *X* servers.

`ssd` checks the environment for availability of certain files and the availability of the Sun Fire 15K system, sets environment variables, and then starts `esmd`(1M). `esmd` monitors environmental changes by polling the related hardware components. When an abnormal condition is detected, `esmd` handles it or generates an event so that the correspondent handlers will take appropriate action and/or update their current status. Some of those handlers are: `dsmd`, `pcd` and Sun Management Center (if installed). The main objective of `ssd` is to ensure that the SMS daemons and servers are always up and running.

FIGURE 3-14 illustrates SSD client server relationship to the SMS daemons.

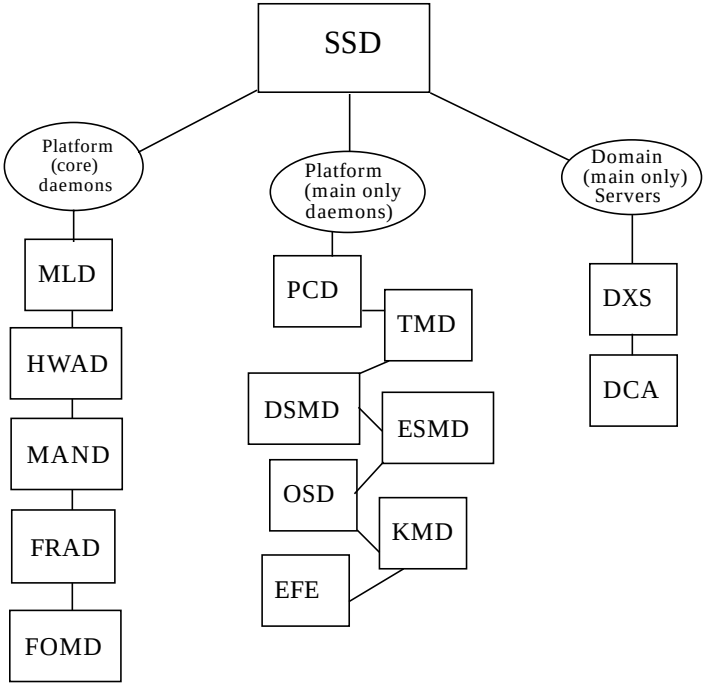


FIGURE 3-14 SSD Client Server Relationships

Scripts

`ssd` uses a configuration file, `ssd_start` to determine which components and in what order to start up the SMS software. This configuration file is located in the `/etc/opt/SUNWSMS/startup` directory.

Caution – This is a system configuration file. Mistakes in editing this file can render the system inoperable. `args` is the only field that should ever be edited in this script. Refer to the daemon man pages for specific options and pay particular attention to syntax.

`ssd_start` consists of entries in the following format:

name:args:nice:role:type:trigger:startup_timeout:shutdown_timeout:uid:start_order:stop_order

where:

name is the name of the program.

args are the valid program options or arguments. Refer to the daemon man pages for more information.

nice specifies a process priority tuning value. Do *not* adjust.

role specifies whether the daemon is platform or domain specific.

type specifies whether the program is a daemon or a server.

trigger specifies whether the program should be started automatically or upon event reception.

startup_timeout is the time in seconds `ssd` will wait for the program to startup.

stop_timeout is the time in seconds `ssd` will wait for the program to shutdown.

uid is the *user_id* the associated program will run under.

start_order is the order in which `ssd` will startup the daemons. Do *not* adjust. Changing the default values can result in the SMS daemons not working properly.

stop_order is the order in which `ssd` will shutdown the daemons. Do *not* adjust. Changing the default values can result in the SMS daemons not working properly.

Spare Mode

Each time `ssd` starts, it comes up in `spare` mode. Once `ssd` has started the platform core daemons running, it queries `fomd`(1M) for its role. If the `fomd` query returns with `spare`, `ssd` will stay in this mode. If the `fomd` returns with `main`, then `ssd` transitions to `main` mode.

After this initial query phase, `ssd` only switches between modes through events received from the `fomd`.

When in spare mode, ssd starts and monitors all of the core platform role, auto trigger programs in the ssd_start file. Currently, this list is made up of the following programs.

- mld
- hwad
- mand
- frad
- fomd

If, while in main mode, ssd receives a spare event, then ssd shuts down all programs except the core platform role and auto trigger programs found in the ssd_start file.

Main Mode

ssd will stay in spare mode until it receives a main event. At that time, ssd starts and monitors (in addition to the already running daemons) all of the platform role (main only) event trigger programs, in the ssd_start file. Currently, this list is made up of the following programs.

- pcd
- tmd
- dsmd
- esmd
- osd
- kmd
- efe

Finally, after starting all the platform role, event trigger programs, ssd queries the pcd to determine which domains are active. For each of these domains, ssd starts all the domain role, event trigger programs found in the ssd_start file.

Domain-specific Process Startup

ssd uses domain start and stop events from pcd as instructions for starting and stopping domain-specific servers.

Upon reception, ssd either starts or stops all of the domain role, event trigger programs (for the domain identified) found in the ssd_start file.

Monitoring and Restarts

Once ssd has started a process, it monitors the process and restarts in the event the process fails.

SMS Shutdown

In certain instances, such as SMS software upgrades, the SMS software needs to be shut down. ssd provides a mechanism to shut down itself and all SMS daemons and servers under its control.

ssd notifies all SMS software components under its control to shut down. After all the SMS software components have been shut down, ssd shuts itself down.

Task Management Daemon

`tmd`(1M) provides task management services such as scheduling for SMS. This reduces the number of conflicts that can arise during concurrent invocations of the hardware tests and configuration software.

Currently, the only service exported by `tmd` is the `hpost`(1M) scheduling service. In the Sun Fire 15K system, `hpost` is scheduled based on two factors.

- Restriction of `hpost`. When the platform first comes up and no domains have been configured, a single instance of `hpost` takes exclusive control of all expanders and configures the centerplane ASICs. All subsequent `hpost` invocations wait until this is complete before proceeding.

  Only a single `hpost` invocation can act on any one expander at a time. For a Sun Fire 15K system configured without split expanders, this restriction does not prevent multiple `hpost` invocations from running. This restriction does come into play however, when the machine is configured with split expanders.

- System-wide `hpost` throttle limit. There is a limit to the number of concurrent `hpost` invocations that can run at a single time without saturating the system. The ability to throttle `hpost` invocations is available using the `-t` option in `ssd_startup`.

Caution – Changing the default value can adversely affect system functionality. Do not adjust this parameter unless instructed by a Sun service representative to do so.

FIGURE 3-15 illustrates TMD client server relationship to the SMS daemons.

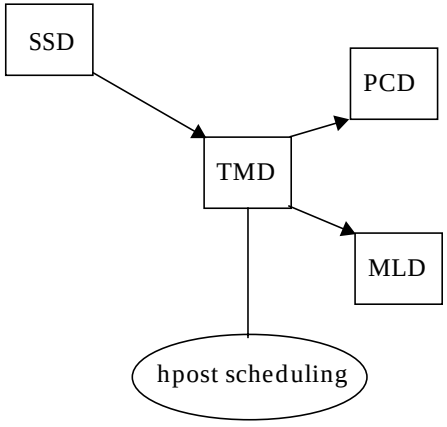


FIGURE 3-15 TMD Client Server Relationships

Environment Variables

Basic SMS environment defaults *must* be set in your configuration files to run SMS commands.

- PATH to include `/opt/SUNWSMS/bin`
- LD_LIBRARY_PATH to include`/opt/SUNWSMS/lib`
- MANPATH to include `/opt/SUNWSMS/man`

Setting other environment variables when you log in can save time. TABLE 3-2 suggests some useful SMS environment variables.

TABLE 3-2 Example Environment Variables

`SMSETC`	The path to the `/etc/opt/SUNWSMS` directory containing miscellaneous SMS-related files.
`SMSLOGGER`	The path to the `/var/opt/SUNWSMS/adm` directory containing the configuration file for message logging, `.logger`.
`SMSOPT`	The path to the `/opt/SUNWSMS` directory containing the SMS package binaries, libraries, and object files; configuration and startup files.
`SMSVAR`	The path to the `/var/opt/SUNWSMS` directory containing platform and domain message and data files.

CHAPTER 4

SMS Configuration

A *dynamic system domain* (DSD) is an independent environment, a subset of a server, that is capable of running a unique version of firmware and a unique version of the Solaris operating environment. Each domain is insulated from the other domains. Continued operation of a domain is not affected by any software failures in other domains nor by most hardware failures in any other domain.

The system controller (SC) supports commands that let you logically group system boards into *dynamic system domains*, or simply *domains*, which are able to run their own operating system and handle their own workload. Domains can be created and deleted without interrupting the operation of other domains. You can use domains for many purposes. For example, you can test a new operating system version or set up a development and testing environment in a domain. In this way, if problems occur, the rest of your system is not affected.

You can also configure several domains to support different departments, with one domain per department. You can temporarily reconfigure the system into one domain to run a large job over the weekend.

The Sun Fire 15K system allows up to 18 domains to be configured.

Domain configuration establishes mappings between the domains and the server's hardware components. Also included in domain configuration is the establishment of various system management parameters and policies for each domain. This chapter discusses all aspects of domain configuration functionality that the Sun Fire 15K system provides.

Domain Configuration Units (DCU)

A domain configuration unit (DCU) is a unit of hardware that can be assigned to a single domain; DCUs are the hardware components from which domains are constructed. DCUs that are not assigned to any domain are said to be in *no-domain*.

All DCUs are system boards and all system boards are DCUs. The Sun Fire 15K DCU's are:

- CPU/Memory board
- Sun Fire HsPCI I/O assembly (HPCI)
- Sun Fire MaxCPU board (MCPU)
- Sun Fire Link wPCI board (WPCI)

Sun Fire 15K hardware requires the presence of at least one of the two types of boards containing CPUs and memory, plus at least one of the I/O board types in each configured domain. csb, exb boards and the SC are *not* DCUs.

Domain Configuration Requirements

You can create a domain out of any group of system boards, provided the following conditions are met:

- The boards are present and not in use in another domain.
- At least one board has a CPU and memory.
- At least one is an I/O board.
- At least one board has a network interface.
- The boards have sufficient memory to support an autonomous domain.
- The name you give the new domain is unique (as specified in the `addtag`(1M) command).
- You have an `idprom.image` file for the domain that was shipped to you by the factory. If your `idprom.image` file has been accidentally deleted or corrupted and you do not have a backup, contact your Sun field support representative.
- There must be at least one boot disk connected to one of the boards that will be grouped together into a domain. Alternatively, if a domain does not have its own disk, there must be at least one network interface so that you can boot the domain from the network.

DCU Assignment

The assignment of DCUs to a domain is the result of one of three logical operations acting upon a DCU (system board):

- Adding the board (from *no-domain*) to a domain
- Removing the board from a domain (leaving the board in *no-domain*)
- Moving the board from one domain to another

Static Versus Dynamic Domain Configuration

Although there are logically three DCU assignment operations the underlying implementation is based upon four domain configuration operations:

- Adding a board to an inactive domain
- Removing a board from an inactive domain
- Adding a board to an active domain
- Removing a board from an active domain

The first two domain configuration operations apply to inactive domains, that is, to domains that are not running software. These operations are called static domain configuration operations. The latter two domain configuration operations apply to active domains, that is, those running software and are called dynamic domain configuration operations.

Dynamic domain configuration requires interaction with the domain's Solaris software to introduce or remove the DCU-resident resources such as CPUs, memory, or I/O devices from Solaris operating environment control. The Sun Fire 15K dynamic reconfiguration (DR) project provides a capability called remote DR for an external agent, such as the SC, to request dynamic configuration services from a domain's Solaris environment.

The SC command user interfaces utilize remote DR as necessary to accomplish the requested tasks. Local automatic DR allows applications running on the domain to be aware of impending DR operations and to take action, as appropriate, to adjust to resource changes. This improves the likelihood of success of DR operations, particularly those which require active resources to be removed from domain use. For more information on DR, refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.

When a domain is configured for local automatic DR, remote DR operations initiated from the SC benefit from the automation of DR operations for that domain. With local automatic DR capabilities available in Sun Fire domains, simple scripts can be constructed and placed in a `crontab`(1) file allowing simple platform reconfigurations to take place on a time schedule.

SMS provides for adding or removing boards from an active (running) domain, (dynamic domain configuration). Initiation of a remote DR operation on a domain requires administrative privilege for that domain. SMS grants the ability to initiate remote DR on a domain to individual administrators on a per-domain basis.

The remote DR interface is secure. Since invocation of DR operations on the domain itself requires superuser privilege, remote DR services are provided only to known, authenticated remote agents.

The user command interfaces that initiate DCU assignment operations are the same whether the affected domain(s) have local automatic DR capabilities or not.

SMS provides for the addition or removal of a board from an inactive domain, such as, static domain configuration using `addboard`, `deleteboard`, and `moveboard`. For more information, refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.

Global Automatic Dynamic Reconfiguration

Remote DR and local automatic DR functions are building blocks for a feature called global automatic DR. Global automatic DR introduces a framework that can be used to automatically redistribute the system board resources on a Sun Fire system. This redistribution can be based upon factors such as production schedule, domain resource utilizations, domain functional priorities, and so on. Global automatic DR accepts input from customers describing their Sun Fire resource utilization policies and then uses those policies to automatically marshal the Sun Fire 15K resources to produce the most effective utilization. For more information on DR, refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.

Configuration for Platform Administrators

This section briefly describes the configuration services available to the platform administrator.

Available Component List

Each domain (A-R) defaults to having a 0-board list of boards that are available to an administrator or configurator to assign to their respective domain(s). Boards can be added to the available component list of a domain by a platform administrator using the `setupplatform`(1M) command. Updating an available component list requires `pcd` to perform the following tasks.

- Update the domain configuration available component list
- Update the board available component list state for each board to show the domain to which it is now `available`
- Notify `dxs` of boards added to their respective domain's available component list
- `dxs` notifies the running domain of the arrival of an `available` board.

▼ To Set Up the Available Component List

`setupplatform` sets up the available component list for domains. If a *domain_id* or *domain_tag* is specified, a list of boards must be specified. If no value is specified for a parameter, it will retain its current value.

1. **In an SC window, log in as a platform administrator.**

2. **Type**:

```
sc0:sms-user:> setupplatform -d domain_id|domain_tag   location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the domain ID.

domain_tag is the name assigned to the domain using `addtag`(1M).

location is the board (DCU) location

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

The following is an example of making boards at SB0, IO1, and IO2 available to domain A:

```
sc0:sms-user:> setupplatform -d A SB0 IO1 IO2
```

At this point the platform administrator can `assign` the board to domain A using the `addboard`(1M) command or leave that up to the domain administrator.

A platform administrator has privileges for only the `-c assign` option of the `addboard` command. All other board configuration requires domain privileges. For more information refer to the `addboard` man page.

Configuring Domains

▼ To Name or Change Domain Names From the Command Line

You do not need to create domains on the Sun Fire 15K system. Eighteen domains have already been established (domains A...R, case insensitive). These domain designations are customizable. This section describes how to uniquely name domains.

Note – Before proceeding, see “Domain Configuration Requirements” on page 58. If the system configuration must be changed to meet any of these requirements, call your service provider.

---

1. **Log in to the SC.**

2. **Type**:

```
sc0:sms-user:> addtag -d domain_id|domain_tag  new_tag
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

new_tag is the new name you want to give to the domain. It should be unique among all domains controlled by the SC.

Naming a domain is optional.

The following is an example of naming Domain `A` to `dmnJ`:

```
sc0:sms-user:> addtag -d A dmnJ
```

▼ To Add Boards to a Domain From the Command Line

1. **Log in to the SC.**

Note – Platform administrators are restricted to using the `-c assign` option and only for `available` not `active` boards.

The system board must be in the `available` state to the domain to which it is being added. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> addboard -d domain_id|domain_tag  -c assign location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c` assign specifies the transition of the board from the current configuration state to the `assign`ed state.

location is the board (DCU) location. Multiple locations are accepted.

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

For example:

```
sc0:sms-user:> addboard -d C -c assign SB0 I01 SB1 I02
```

SB0, IO1, and IO2 have now gone from being `available` to domain C to being assigned to it.

`addboard` performs tasks synchronously and does not return control to the user until the command is complete. If the command fails the board does not return to its original state. A `dxs` or `dca` error is logged to the domain. If the error is recoverable you can retry the command. If it is unrecoverable, you will need to reboot the domain in order to use that board.

▼ To Delete Boards From a Domain From the Command Line

Note – Platform administrators are restricted to using the `-c unassign` option and only for `assign`ed not active boards.

---

1. **Log in to the SC.**

The system board must be in the assigned state to the domain from which it is being deleted. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> deleteboard -d domain_id|domain_tag -c unassign location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c unassign` specifies the transition of the board from the current configuration state to a new `unassigned` state.

location is the board (DCU) location. Multiple locations are accepted.

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

For example:

```
sc0:sms-user:> deleteboard -d C -c unassign SB0
```

SB0 has now gone from being assigned to domain C to being available to it.

If the command fails the board does not return to its original state. A `dxs` or `dca` error is logged to the domain. If the error is recoverable you can retry the command. If it is unrecoverable, you will need to reboot the domain in order to use that board.

▼ To Move Boards Between Domains From the Command Line

Note – Platform administrators are restricted to the `-c assign` option and only for assigned not active boards.

---

1. **Log in to the SC.**

The system board must be in the assigned state to the domain from which it is being deleted. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> moveboard -d domain_id|domain_tag -c assign location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c assign` specifies the transition of the board from the current configuration state to an `assign`ed state.

location is the board (DCU) location

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

`moveboard` performs tasks synchronously and does not return control to the user until the command is complete. You can only specify one *location* when using `moveboard`.

For example:

```
sc0:sms-user:> moveboard -d C -c assign SB0
```

SB0 has been moved from its previous domain and assigned to domain C.

If the command fails the board does not return to its original state. A `dxs` or `dca` error is logged to the domain. If the error is recoverable you can retry the command. If it is unrecoverable, you will need to reboot the domain in order to use that board.

▼ To Obtain Board Status

1. **Log in to the SC.**

Platform administrators can obtain board status for all domains.

2. **Type**:

```
sc0:sms-user:> showboards [-d domain_id|domain_tag]
```

The board status is displayed.

The following partial example shows the board information for a user with platform administrator privileges. All domains are visible.

```
sc0:sms-user:> showboards
Retrieving board information. Please wait.
.........
Location    Pwr    Type of Board   Board Status  Test Status   Domain
--------    ---    -------------   ------------  -----------   ------
SB0         On     CPU             Active        Passed        domainC
SB1         On     CPU             Active        Passed        A
SB2         On     CPU             Active        Passed        A
SB3         On     CPU             Active        Passed        engB
SB4         On     CPU             Active        Passed        engB
SB5         On     CPU             Active        Passed        engB
SB6          -     Empty Slot      Available         -         Isolated
SB7         On     CPU             Active        Passed        domainC
SB8         Off    CPU             Available     Unknown       Isolated
SB9         On     CPU             Active        Passed        dmnI
SB10        Off    CPU             Available     Unknown       Isolated
SB11        Off    CPU             Available     Unknown       Isolated
SB12         -     Empty Slot      Available         -         Isolated
SB13         -     Empty Slot      Available         -         Isolated
SB14        Off    CPU             Assigned      Failed        domainC
SB15        On     CPU             Active        Passed        J
SB16        On     CPU             Active        Passed        K
SB17         -     Empty Slot      Assigned          -         dmnL
IO0          -     Empty Slot      Available         -         Isolated
IO1         On     HPCI            Active        Passed        A
IO2         On     HPCI            Active        Passed        engB
IO3         On     HPCI            Active        Passed        domainC
IO4         On     HPCI            Available     Degraded      domainC
IO5         Off    HPCI            Available     Unknown       Isolated
IO6          -     Empty Slot      Available         -         Isolated
IO7         On     wPCI            Active        Passed        domainD
IO8         On     wPCI            Active        Passed        domainE
IO9         On     wPCI            Active        Passed        domainF
IO10        On     wPCI            Assigned      Unknown       domainG
IO11        On     wPCI            Assigned      Unknown       domainH
```

▼ To Obtain Domain Status

1. **Log in to the SC.**

Platform administrators can obtain domain status for all domains.

2. **Type**:

```
sc0:sms-user:> showplatform -d domain_id|domain_tag
```

The status listing is displayed.

The following partial example shows the domain information for a user with platform administrator privileges. All domains are visible.

```
sc0:sms-user:> showplatform
...
Domain   Solaris Hostname   Domain Status
newA     sun15-b0            Powered Off
engB     sun15-b1            Keyswitch Standby
domainC  sun15-b2            Running OBP
eng1     sun15-b3            Loading Solaris
E        sun15-b4            Running Solaris
domainF  sun15-b5            Running Solaris
dmnG     sun15-b6            Running Solaris
H        sun15-b7            Solaris Quiesced
I        sun15-b8            Powered Off
dmnJ     sun15-b9            Powered Off
K        sun15-b10           Booting Solaris
L        sun15-b11           Powered Off
M        sun15-b12           Powered Off
N        sun15-b13           Keyswitch Standby
O        sun15-b14           Powered Off
P        sun15-b15           Running Solaris
Q        sun15-b16           Running Solaris
dmnR     sun15-b17           Running Solaris
```

Virtual Time of Day

The Solaris environment uses the functions provided by a hardware time of day (TOD) chip to support Solaris system date/time. Typically Solaris software reads the current system date/time at boot using a get TOD service. From that point forward, Solaris software either uses a high resolution hardware timer to represent current date/time or, if configured, uses Network Time Protocol (NTP) to synchronize current system date/time to a (presumably more accurate) time source.

The SC is the only computer on the platform that has a realtime clock. The virtual TOD for domains is stored as an offset from that realtime clock value. Each domain can be configured to use NTP services instead of `setdate` (1M) to manage the running system date/time, however, the SC should *not* use NTP to set its clock. If the SC uses NTP, no offset adjustments will be made and the virtual TOD values stored on the domain will get skewed. For more information on NTP, see "Configuring NTP" on page 71 or refer to the `xntpd`(1M) man page in the *man Pages(1M): System Administration Commands* section of the Solaris 8 02/02 Reference Manual Collection.

Note – NTP is a separate package that must be installed and configured on the domain in order to function as described. Use `setdate` on the domain prior to installing NTP.

However system date/time is managed while Solaris software is running, an attempt is made to keep the boot-time TOD value accurate by setting the TOD when variance is detected between the current TOD value and the current system date/time.

Since the Sun Fire 15K hardware provides no physical TOD chip for Sun Fire domains, SMS provides the time-of-day services required by the Solaris environment for each domain. Each domain is supplied with a TOD service that is logically separate from that provided to any other domain. This difference allows system date/time management on a Sun Fire 15K domain to be as flexible as that provided by standalone servers. In the unlikely event that a domain needs to be set up to run at a time other than real world time, the Sun Fire 15K TOD service allows that domain to be configured without affecting the TOD values supplied to other domains running real world time.

Time settings are implemented using `setdate`(1M). You must have platform administrator privileges in order to run `setdate`. See "All Privileges" on page 20 for more information.

Setting the Date and Time

`setdate` (1M) allows the SC platform administrator to set the system controller date and time values. After setting the date and time, `setdate`(1M) displays the current date and time for the user.

▼ To Set the Date on the SC

1. **Log in to the SC.**
2. **Type:**

```
sc0:sms-user:> setdate 020210302000.00
System Controller: Wed Feb 2 10:30 2000 US/Pacific
```

Optionally, `setdate`(1M) can set a domain TOD. The domain's keyswitch must be in the `off` or `standby` position. You must have platform administrator privileges to run this command on the domain.

▼ To Set the Date for Domain eng2

1. **Log in to the SC.**
2. **Type:**

```
sc0:sms-user:> setdate -d eng2 020210302000.00
Domain eng2: Wed Feb 2 10:30 2000 US/Pacific
```

`showdate`(1M) displays the current SC date and time.

▼ To Display the Date on the SC

1. **Log in to the SC.**
2. **Type:**

```
sc0:sms-user:> showdate
System Controller: Wed Feb 2 10:30 2000 US/Pacific
```

Optionally, `showdate`(1M) can display the date and time for a specified domain. Superuser or any member of a platform or domain group can run `showdate`.

▼ To Display the Date on Domain eng2

1. **Log in to the SC.**
2. **Type**:

```
sc0:sms-user:> showdate -d eng2
Domain eng2: Wed Feb 2 10:30 2000 US/Pacific
```

Configuring NTP

The NTP daemon, x`ntpd(1M)` for the Solaris 8 02/02 operating environment, provides a mechanism for keeping the time settings synchronized between the SC and the domains. The OpenBoot PROM obtains the time from the SC when the domain is booted, and NTP keeps the time synchronized on the domain from that point on.

NTP configuration is based on information provided by the system administrator.

Note – Do not use NTP to set the SC clock. It should only be used on domains.

The NTP packages are compiled with support for a local reference clock. This means that your system can poll itself for the time instead of polling another system or network clock. The poll is done through the network loopback interface. The numbers in the IP address are 127.127.1.0. This section describes how to set the time on the SC using `setdate`, and then to set up the SC to use its own internal time-of-day clock as the reference clock in the `ntp.conf` file.

NTP can also keep track of the drift (difference) between the SC clock and the domain clock. NTP corrects the domain clock if it loses contact with the SC clock, provided that you have a drift file declaration in the `ntp.conf` file. The drift file declaration specifies to the NTP daemon the name of the file that stores the error in the clock frequency computed by the daemon. See the following procedure for an example of the drift file declaration in an `ntp.conf` file.

If the `ntp.conf` file does not exist, create it as described below. You must have an `ntp.conf` file on both the SC and the domains.

▼ To Create the `ntp.conf` File

1. **Log in to the main SC as superuser.**

2. **Change to the** `/etc/inet` **directory and copy the NTP *server* file to the NTP configuration file:**

```
sc0:# cd /etc/inet
sc0:# cp ntp.server ntp.conf
```

3. **Using a text editor, edit the** `/etc/inet/ntp.conf` **file created in the previous step.**

   The `ntp.conf` file for the Solaris 8 02/02 operating environment is located in `/etc/inet`.

   The following is an example of server lines in the `ntp.conf` file on the main SC, to synchronize clocks.

```
peer spare_sc_hostname
server 127.127.1.0
fudge 127.127.1.0 stratum 13
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

4. **Save the file and exit.**

5. **Stop and restart the NTP daemon:**

```
sc0:# /etc/init.d/xntpd stop
sc0:# /etc/init.d/xntpd start
```

6. **Log in to the spare SC as superuser.**

7. **Change to the** `/etc/inet` **directory and copy the NTP *server* file to the NTP configuration file:**

```
sc1:# cd /etc/inet
sc1:# cp ntp.server ntp.conf
```

8. **Using a text editor, edit the** `/etc/inet/ntp.conf` **file created in the previous step.**

   The `ntp.conf` file for the Solaris 8 02/02 operating environment is located in `/etc/inet`.

   The following is an example of server lines in the `ntp.conf` file on the spare SC, to synchronize clocks.

```
server main_sc_hostname prefer
server 127.127.1.0
fudge 127.127.1.0 stratum 13
driftfile /var/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

9. **Stop and restart the NTP daemon**:

```
sc0:# /etc/init.d/xntpd stop
sc0:# /etc/init.d/xntpd start
```

10. **Log in to each domain as superuser.**

11. **Change to the** `/etc/inet` **directory and copy the NTP *client* file to the NTP configuration file**:

```
domain_id:# cd /etc/inet
domain_id:# cp ntp.client ntp.conf
```

12. **Using a text editor, edit the** `/etc/inet/ntp.conf` **file created in the previous step.**

    The `ntp.conf` file for the Solaris 8 02/02 operating environment is located in `/etc/inet`.

    For the Solaris 8 02/02 operating environment, you can add lines similar to the following to the `/etc/inet/ntp.conf` on the domains:

```
server main_sc_hostname prefer
server spare_sc_hostname
```

Note – If you do not have any server lines in the `ntp.conf` file other than the local clock, it might take up to 25 minutes for NTP to synchronize at boot time. For a workaround, refer to Bug 4325813.

13. **Save the file and exit.**

14. **Change to the initialization directory, stop and restart the NTP daemon on the domain:**

```
domain_id:# /etc/init.d/xntpd stop
domain_id:# /etc/init.d/xntpd start
```

NTP is now installed and running on your domain. Repeat Step 10 through Step 14 for each domain.

For more information on the NTP daemon, refer to the `xntpd`(1M) man page in the *man Pages(1M): System Administration Commands* section of the Solaris 8 02/02 Reference Manual Collection.

Virtual ID PROM

Each configurable domain has a virtual ID PROM that contains identifying information about the domain such as hostID and domain Ethernet address. The hostID is unique among all domains on the same platform. The Ethernet address is world unique.

Sun Fire 15K system management software provides a virtual ID PROM for each configurable domain containing identifying information that can be read, but not written, from the domain. The information provided meets the requirements of the Solaris environment.

Flashupdate Command

SMS provides the `flashupdate`(1M) command to update the Flash PROM in the system controller (SC), and the Flash PROMs in a domain's CPU and MaxCPU boards after SMS software upgrades or applicable patch installation. `flashupdate` displays both the current Flash PROM and the flash image file information prior to any updates.

Note – Once you have updated the FPROM(s) you must reset the SC using the `reset-all` command at the OpenBoot PROM (`ok`) prompt.

For more information and examples, refer to the `flashupdate` man page.

Configuration for Domain Administrators

This section briefly goes over the configuration services available to the domain administrator.

Configuring Domains

The domain administrator has been given far more capability with regard to the `addboard`, `deleteboard`, and `moveboard` commands.

▼ To Add Boards to a Domain From the Command Line

1. **Log in to the SC as a domain administrator for that domain.**

Note – In order for the domain administrator to add a board to a domain, that board must appear in the domain available component list.

The system board must be in the available or assigned state to the domain to which it is being added. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> addboard -d domain_id|domain_tag  -c function  location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c` *function* specifies the transition of the board from the current configuration state to a new configuration state.

Configuration states are: `assign`, `connect`, or `configure`. If the `-c` *function* option is not specified, the default expected configuration state is `configure`.

location is the board (DCU) location. Multiple locations are accepted.

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

For example:

```
sc0:sms-user:> addboard -d C -c assign SB0 I01 SB1 I02
```

SB0, IO1, and IO2 have now gone from being available to domain C to being assigned to it.

`addboard` performs tasks synchronously and does not return control to the user until the command is complete. If the board is not powered on or tested, specify the `-c connect|configure` option, then the command will power on the board and test it.

▼ To Delete Boards From a Domain From the Command Line

1. **Log in to the SC as a domain administrator for that domain.**

The system board must be in the assigned or active state to the domain from which it is being deleted. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> deleteboard -d domain_id|domain_tag  -c function location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c` *function* specifies the transition of the board from the current configuration state to a new configuration state.

Configuration states are: `unconfigure`, `disconnect`, or `unassign`. If the `-c` option is not specified, the default expected configuration state is `unassign`.

location is the board (DCU) location. Multiple locations are accepted.

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

For example:

```
sc0:sms-user:> deleteboard -d C -c unassign SB0
```

SB0 has now gone from being assigned to domain C to being available to it.

Note – A domain administrator can unconfigure and disconnect a board but is not allowed to delete a board from a domain unless the `deleteboard` [*location*] field appears in the domain's available component list.

▼ To Move Boards Between Domains From the Command Line

Note – You must have domain administrator privileges for both domains involved.

---

1. **Log in to the SC as a domain administrator for that domain.**

The system board must be in the assigned or active state to the domain from which it is being deleted. Use the `showboards` (1M) command to determine a board's state.

2. **Type**:

```
sc0:sms-user:> moveboard -d domain_id|domain_tag -c function location
```

where:

`-d` indicates either a domain ID or tag.

domain_id is the current domain ID.

domain_tag is the current name assigned to the domain using `addtag`(1M).

`-c` *function* specifies the transition of the board from the current configuration state to a new configuration state.

Configuration states are: `assign`, `connect`, or `configure`. If the `-c` option is not specified, the default expected configuration state is `configure`.

location is the board (DCU) location

The following *location* forms are accepted:

SB(0...17)

IO(0...17)

`moveboard` performs tasks synchronously and does not return control to the user until the command is complete. If the board is not powered on or tested specify `-c connect|configure`, then the command will power on the board and test it. You can only specify one *location* when using `moveboard`.

▼ To Obtain Board Status

1. **Log in to the SC.**

Domain administrators can obtain board status only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> showboards [-d domain_id|domain_tag]
```

The board status is displayed.

The following partial example shows the board information for a user with domain administrator privileges for domain A.

```
sc0:sms-user:> showboards -d A

Location    Pwr    Type of Board   Board Status  Test Status   Domain
----        ---    -------------   ------------  -----------   ------
SB1         On     CPU             Active        Passed        A
SB2         On     CPU             Active        Passed        A
IO1         On     HPCI            Active        Passed        A
```

▼ To Obtain Domain Status

1. **Log in to the SC.**

Domain administrators can obtain domain status only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> showplatform -d domain_id|domain_tag
```

The status listing is displayed.

The following partial example shows the domain information for a user with domain administrator privileges for domains `newA`, `engB` and `domainC`.

```
sc0:sms-user:> showplatform
...
Domain   Solaris Hostname   Domain Status
newA     sun15-b0            Powered Off
engB     sun15-b1            Keyswitch Standby
domainC  sun15-b2            Running OBP
```

▼ To Obtain Device Status

1. **Log in to the SC.**

Domain administrators can obtain board status only for those domains for which they have privileges.

2. **Type**:

```
sc0:sms-user:> showdevices [-d domain_id|domain_tag]
```

The device status is displayed.

The following partial example shows the device information for a user with domain administrator privileges for domain A.

```
sc0:sms-user:> showdevices IO1

IO Devices
----------
domain    location  device  resource              usage
A         IO1       sd3     /dev/dsk/c0t3d0s0     mounted filesystem "/"
A         IO1       sd3     /dev/dsk/c0t3s0s1     dump device (swap)
A         IO1       sd3     /dev/dsk/c0t3s0s1     swap area
A         IO1       sd3     /dev/dsk/c0t3d0s3     mounted filesystem "/var"
A         IO1       sd3     /var/run              mounted filesystem "/var/run"
```

Virtual Keyswitch

Each Sun Fire 15K domain has a virtual keyswitch. Like the Sun™ Enterprise servers physical keyswitch, the Sun Fire 15K domain virtual keyswitch controls whether the domain is powered on or off, whether increased diagnostics are run at boot, whether certain operations (for example, flash PROM updates and domain `reset` commands) are permitted.

Only domains configured with their virtual keyswitch powered on are booted, monitored, and subject to automatic recovery actions, should they fail.

Virtual keyswitch settings are implemented using `setkeyswitch`(1M). You must have domain administrator privileges for the specified domain in order to run `setkeyswitch`. See “All Privileges” on page 20 for more information.

Setkeyswitch

`setkeyswitch` (1M) changes the position of the virtual key switch to the specified value. `pcd` (1M) maintains the state of each virtual key switch between power cycles of the SC or physical power cycling of the power supplies.

`setkeyswitch`(1M) is responsible for loading all the configured processors’ bootbus SRAM. All the processors are started, with one processor designated as the boot processor. `setkeyswitch`(1M) loads OpenBoot™ PROM into the memory of the Sun Fire 15K system domain and starts OpenBoot PROM on the boot processor.

The primary task of OpenBoot PROM is to boot and configure the operating system from either a mass storage device or from a network. OpenBoot PROM also provides extensive features for testing hardware and software interactively.

The `setkeyswitch`(1M) command syntax follows:

```
sc0:sms-user:> setkeyswitch -d domain_id|domain_tag [-q -y|-n]
on|standby|off|diag|secure
```

where:

`-d` *domain_id* - ID for a domain. Valid domain_ids are A...R and case insensitive.

`-d` *domain_tag* - Name assigned to a domain using `addtag` (1M).

`-h` - Help. Displays usage descriptions.

`-q` - Quiet. Suppresses all messages to `stdout` including prompts. When used alone `-q` defaults to the `-n` option for all prompts. When used with either the `-y` or the `-n` option, `-q` suppresses all user prompts, and automatically answers with either Y or N based on the option chosen.

`-n` - Automatically answers no to all prompts. Prompts are displayed unless used with `-q` option.

`-y` - Automatically answers yes to all prompts. Prompts are displayed unless used with `-q` option.

The following operands are supported:

- `on`

From the `off` or `standby` position, `on` powers on all boards assigned to the domain (if not already powered on). Then the domain is brought up. Depending on the size, configuration and diagnostic settings of the domain, this will take approximately 20 minutes.

From the `diag` position, `on` is nothing more than a position change, but upon the next reboot of the domain power on self-test (POST) is not invoked with verbosity and the diag level is set to its maximum.

From the `secure` position, `on` restores write permission to the domain.

- `standby`

From the `off` position, `standby` powers on all boards assigned to the domain (if not already powered on).

From the `on`, `diag`, or `secure` position, `standby` optionally causes an `Are you sure?` prompt and the domain gracefully shuts down. The boards remain fully powered.

- `off`

From the `on`, `diag`, or `secure` position, `off` optionally causes an `Are you sure?` prompt, and all boards are put into low-power mode.

From the `standby` position, `off` puts all boards into low-power mode.

- `diag`

From the `off` or `standby` position, `diag` powers on all boards assigned to the domain (if not already powered on). Then the domain is brought up just as in the on position, except that POST is invoked with verbosity and the diag level is set to its maximum.

From the `on` position, `diag` is nothing more than a position change, but upon the next reboot of the domain, POST is invoked with verbosity and the diag level set to its maximum.

From the `secure` position, `diag` restores write permission to the domain and upon the next reboot, POST is invoked with verbosity and the diag level is set to its maximum.

- `secure`

From the `off` or `standby` position, `secure` powers on all boards assigned to the domain (if not already powered on). Then the domain is brought up just as in the `on` position, except that the `secure` position removes write permission to the domain. For example, `flashupdate` and `reset` will not work.

From the `on` position, `secure` removes write permission to the domain (as described above). From the `diag` position, `secure` removes write permission to the domain (as described above), and on the next reboot of the domain POST is invoked with verbosity and the diag level set to its normal values.

▼ To Set the Virtual Keyswitch On in Domain A

1. **Log in to the SC.**

   Domain administrators can set the virtual keyswitch only for those domains for which they have privileges.

2. **Type**:

```
sc0:sms-user:> setkeyswitch -d A on
```

`showkeyswitch` (1M) displays the position of the virtual keyswitch of the specified domain. The state of each virtual keyswitch is maintained between power cycles of the SC or physical power cycling of the power supplies by the `pcd` (1M). Superuser or any member of a platform or domain group can run `showkeyswitch`.

▼ To Display the Virtual Keyswitch Setting in Domain A

1. **Log in to the SC.**

   Domain administrators can obtain keyswitch status only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> showkeyswitch -d A
Virtual keyswitch position: ON
```

Virtual NVRAM

Each domain has a virtual NVRAM containing OpenBoot PROM data such as the OpenBoot PROM variables. OpenBoot PROM is a binary image stored on the SC in `/opt/SUNWSMS/hostobjs` which `setkeyswitch` downloads into domain memory at boot time. There is only one version of OpenBoot PROM for all domains.

SMS software provides a virtual NVRAM for each domain and allows OpenBoot PROM full read / write access to this data.

For most NVRAM variables, the only interface available to read or write them is OpenBoot PROM. The exceptions are those OpenBoot PROM variables which must be altered in order to bring OpenBoot PROM up in a known working state or to diagnose problems that hinder OpenBoot PROM bring up. These variables are not a replacement for the OpenBoot PROM interface.

These limited number of OpenBoot PROM variable values in the domain NVRAM are readable and writable from SMS using `setobpparams`(1M). You must have domain administrator privileges to run `set/showobpparams`. If you change variables for a running domain, you must reboot the domain in order for the changes to take effect.

Note – Only experienced system administrators, familiar with OpenBoot PROM commands and their dependencies should attempt to use `setobpparams` in any manner other than that described.

Setting the OpenBoot PROM Variables

`setobpparams`(1M) sets and gets a subset of a domain's virtual NVRAM variables and REBOOTINFO data using the following syntax.

```
sc0: sms-user:> setobpparams -d domain_id|domain_tag  param=value...
```

where *param=value is :*

Variables	=	Default Value	Comment
`diag-switch?`	=	false	When set to false, the default boot device is specified by `boot-device` and the default boot file by `boot-file`. If set to true, OpenBoot PROM runs in diagnostic mode and you need to set either `diag-device` or `diag-file` to specify the correct default boot device or file. These default boot device and file settings cannot be set using `setobpparams`. Use `setenv`(1) in OpenBoot PROM.
`auto-boot?`	=	false	When set to true, the domain boots automatically after power-on or reset-all. The boot device and boot file used are based on the settings for `diag-switch` (see above). Neither `boot-device` nor `boot file` can be set using `setobpparams`. In the event the `OK` prompt is unavailable, such as a repeated panic, use `setobpparams` to set `auto-boot?` to `false`. When the `auto-boot?` variable is set to false using `setobpparams`, the reboot variables are invalidated, the system will not boot automatically and will stop in Open Boot PROM where new NVRAM variables can be set. See "To Recover From a Repeated Domain Panic" on page 86.

Variables	=	Default Value	Comment
security-mode	=	none	Firmware security level. Valid variable values for security-mode are: • none - No password required (default). • command - All commands except for boot(1M) and go require the password. • full - All commands except for go require the password.
use-nvramrc?	=	false	When set to true, this variable executes commands in NVRAMRC during system start-up.
fcode-debug?	=	false	When set to true, this variable includes name fields for plug-in device FCodes.

The following is an example of how setobpparams can be useful.

▼ To Recover From a Repeated Domain Panic

Say domain A encounters repeated panics caused by a corrupted default boot disk.

1. **Log in to the SC with domain administrator privileges.**
2. **Stop automatic reboot.**

```
sc0:sms-user:> setkeyswitch -d A standby
sc0:sms-user:> setobpparams -d A auto-boot? false
```

3. **Repost the domain.**

```
sc0:sms-user:> setkeyswitch -d A off
sc0:sms-user:> setkeyswitch -d A on
```

4. **Once the domain has come up to the OK prompt set NVRAM variables to a new non-corrupted boot-device.**

```
ok setenv boot-device bootdisk_alias
```

where:

bootdisk_alias is a user-defined alias you created. The boot device must correspond to the a bootable disk on which you have installed the operating environment.

5. **Now that you have set up a new alias for your boot device, boot the disk by typing:**

```
ok boot
```

For more information on OpenBoot variables refer to the *OpenBoot 4.x Command Reference Manual*.

▼ To Set the OpenBoot PROM Security Mode Variable in Domain A

1. **Log in to the SC.**

   Domain administrators can set the OpenBoot PROM variables only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> setobpparams -d A security-mode full
```

`security-mode` has been set to full. All commands except `go` require a password on domain A. You must reboot a running domain in order for the change to take effect.

▼ To See the OpenBoot PROM Variables

1. **Log in to the SC.**

   Domain administrators can set the OpenBoot PROM variables only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> showobpparams -d domain_id|domain_tag
```

SMS NVRAM updates are supplied to OpenBoot PROM at OpenBoot PROM initiation (or domain reboot time). For more information refer to the *OpenBoot PROM 4.x Command Reference Manual.*

Degraded Configuration Preferences

In most situations, hardware failures that cause a domain crash are detected and eliminated from the domain configuration either by POST or OpenBoot PROM during the subsequent automatic recovery boot of the domain. However, there can be situations where failures are intermittent or the boot-time tests are inadequate to detect failures that cause repeated domain failures and reboots. In those situations, Sun Fire 15K system management software uses configurations or configuration policies supplied by the domain administrator to eliminate hardware from the domain configuration in an attempt to get a stable domain environment running.

The following commands can be run by either platform or domain administrators. Domain administrators are restricted to the domains for which they have privileges.

Setbus

`setbus`(1M) dynamically reconfigures bus traffic on active expanders in a domain to use either one centerplane support board (CSB) or both. Using both CSBs is considered *normal* mode. Using one CSB is considered *degraded* mode.

You must have platform administrator privileges or domain privileges for the specified domain in order to run `setbus`.

This feature allows you to swap out a CSB without having to power off the system. Valid buses are:

- `a` - configures the address bus
- `d` - configures the data bus
- `r` - configures the response bus

▼ To Set All Buses on All Active Domains to Use Both CSBs

1. **Log in to the SC.**

   Domain administrators can set the bus only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> setbus -c CS0,CS1
```

For more information on reconfiguring bus traffic, refer to the `setbus`(1M) man page.

Showbus

`showbus`(1M) displays the bus configuration of expanders in active domains. This information defaults to displaying configuration by slot order 0–17. Any member of a platform or domain group can run `showbus`.

▼ To Show All Buses on All Active Domains

1. **Log in to the SC.**

   Domain administrators can set the bus only for those domains for which they have privileges.

2. **Type**:

```
sc0:sms-user:> showbus
```

   For more information on reconfiguring bus traffic, refer to the `showbus`(1M) man page.

CHAPTER 5

Domain Control

This chapter addresses the functions that provide control over domain software as well as server hardware. Control functions are invoked at the discretion of an administrator. They are also useful to SMS for providing automatic system recovery (ASR).

Domain control functionality provides control over the software running on a domain. It includes those functions that allow a domain to be booted and interrupted. Only the domain administrator can invoke the domain control functions.

Domain Boot

This section describes the various aspects of booting the Solaris software environment in a domain running SMS software.

`setkeyswitch`(1M) is responsible for initiating and sequencing a domain boot. It powers on the domain hardware as required and invokes POST to test and configure the hardware in the logical domain into a Sun Fire 15K physical hardware domain. It downloads and initiates OpenBoot PROM as required to boot the Solaris operating environment on the domain.

Only domains that have their virtual keyswitch set appropriately are subject to boot control. See “Virtual Keyswitch” on page 81.

OpenBoot PROM boot parameters are stored in the domain’s virtual NVRAM. `osd`(1M) provides those parameter values to OpenBoot PROM, which adapts the domain boot as indicated.

Certain parameters, in particular those that may not be adjustable from OpenBoot PROM itself when a domain is failing to boot, can be set by `setobpparams`(1M) so that they take effect at the next boot attempt.

Keyswitch On

The domain keyswitch control (“Virtual Keyswitch” on page 81) manually initiates domain boot.

`setkeyswitch` boots a properly configured domain when its keyswitch control is moved from the `off` or `standby` position to one of the `on` positions. This takes approximately 20 minutes.

`setobpparams`(1M) provides a method by which a manually initiated (keyswitch control) domain boot sequence can be stopped in OpenBoot PROM. For more information see “Setting the OpenBoot PROM Variables” on page 85 and refer to the `setobpparams` man page.

Power

SMS boots all properly configured domains when the Sun Fire 15K chassis is powered on using the `poweron`(1M) command. SMS shuts down all properly configured domains when the chassis is powered off using the `poweroff` command.

SMS checks the power state of components to determine if they are on or off and enables or disables console bus ports (where appropriate) when boards are powered on/off. `poweron` checks to see if a component is physically present. `poweroff` unconfigures DCUs from the expander and changes the expander from split-slot to non-split-slot when appropriate. `poweroff` unconfigures the expander from the centerplane when the expander is powered off and checks for voltage reading tolerances to help determine if the board is on or off.

The following components can be power controlled using the `poweron` and `poweroff` commands.

- Bulk power supply
- Fan tray
- Centerplane support board
- Expander board
- CPU/Memory board
- Standard PCI board
- Hot-pluggable PCI assembly
- MaxCPU board
- wPCI board
- System controller (spare only; `poweroff` only. `resetsc` is used to power on the spare.)

▼ To Power System Boards On and Off From the Command Line

Platform administrators are allowed to power control the entire system and can execute these commands without a *location* option. Domain administrators can power control any system board assigned to their domain(s). Users with only domain privileges must supply the *location* option.

1. **To power on a system component, type:**

```
sc0:sms-user:> poweron location
```

where

location is the location of the system component you wish to power on and, if you are a domain administrator, for which you have privileges.

For more information, refer to the `poweron`(1M) man page.

2. **To power off a system component, type:**

Note – If you are powering off a component to replace it, use the `poweroff`(1M) command. Do not use the breakers to power off the component; this can cause a Domain Stop.

```
sc0:sms-user:> poweroff location
```

where:

location is the location of the system component you wish to power off and if you are a domain administrator, for which you have privileges.

For more information, refer to the `poweroff`(1M) man page.

If you try to power off the system while any domain is actively running the operating system, the command will fail and display a message in the message panel of the window. In that case, issuing a `setkeyswitch` *domain_id* `standby` command for the active domain(s) will gracefully shut down the processors. Then, you can reissue the command to power off.

If the platform loses power due to a power outage, `pcd` records and saves the last state of each domain before power was lost.

▼ To Recover From Power Failure

If you lose power only on the SC, switch on the power to the SC. Sun Fire 15K domains are not affected by the loss of power to the SC. If you lose power to both the SC and the domains, use the following procedure to recover from the power failure. For switch locations refer to the *Sun Fire 15K System Site Planning Guide*

1. **Manually switch off the bulk power supplies on the Sun Fire 15K system as well as the power switch on the SC.**

   This prevents power surge problems that can occur when power is restored.

2. **After power is restored, manually switch on the bulk power supplies on the Sun Fire 15K system.**

3. **Manually switch on the SC power.**

   This boots the SC and starts the SMS daemons. Check your SC platform message file for completion of the SMS daemons.

4. **Wait for the recovery process to complete.**

   Any domain that was powered on and running the Solaris operating environment returns to the operating environment run state. Domains at OpenBoot PROM eventually return to an OpenBoot PROM run state.

   The recovery process must finish before any SMS operation is performed. You can monitor the domain message files to determine when the recovery process has completed.

Domain-requested

SMS reboots domains upon request from the domain software (Solaris software or `dsmd`). The domain software requests reboot services in the following situations.

- Upon execution of a user reboot request, for example, Solaris `reboot`(1M) or the OpenBoot PROM boot command, `reset-all`
- Upon Solaris software panic
- Upon trapping the CPU-detected RED_mode or Watchdog Reset conditions

Automatic System Recovery (ASR)

Automatic system recovery (ASR) consists of those procedures that restore the system to running all properly configured domains after one or more domains have been rendered inactive due to software or hardware failures or due to unacceptable environmental conditions.

SMS software supports a software-initiated reboot request as part of ASR. Every domain that crashed is automatically rebooted by `dsmd`.

Situations that require ASR are domain boots requested by domain software upon detecting failures that crash the domain (for example, panic).

There are other situations, such as detection of domain software hangs as described in “Solaris Software Hang Events” on page 154, where SMS initiates a domain boot as part of the recovery process.

`dsmd` ignores the OpenBoot PROM parameter, `auto-boot?`, which on systems without a service processor can prevent the system from automatically rebooting in power-on-reset situations. `dsmd` does *not* ignore keyswitch control. If the keyswitch is set to `off` or `standby`, the keyswitch setting will be honored in determining whether a domain is subject to ASR reboot actions.

Fast Boot

In general a fast domain reboot is possible in situations where:

- No serious error has been attributed to hardware since the last boot
- No change has been made to the configuration of hardware assigned to the domain since the last boot

Because SMS is responsible for monitoring the hardware, detecting, and responding to errors, SMS decides whether or not to request a fast reboot based upon its record of hardware errors since the last boot.

Because POST controls the hardware configuration based upon a number of inputs including, but not limited to, the blacklist data, POST decides whether or not the hardware configuration has changed so as to preclude a fast reboot. If system management has requested a fast reboot, POST will verify that the hardware configuration implied by its current inputs matches the hardware configuration used for the last boot; if it does not, POST will fail the fast-POST operation. The system management software is prepared to recover from this type of POST failure by requesting a full-test (slow) domain boot.

Sun Fire 15K system management software minimizes the elapsed time taken by the part of the domain boot process that it can control.

Domain Abort/Reset

Certain error conditions can occur in a domain that require aborting the domain software or issuing a reset to the domain software or hardware. This section describes the domain abort/reset functions that are provided by `dsmd`.

`dsmd` provides a software-initiated mechanism to abort a domain Solaris OS, requesting that it panic to take a core image. No user intervention is needed.

SMS provides the `reset`(1M) command to allow the user to abort the domain software and issue a reset to the domain hardware.

Control is passed to OpenBoot PROM after the `reset` command is issued. In the case of a user-interface-issued `reset` command, OpenBoot PROM uses its default configuration to determine whether the domain is booted to the Solaris environment. In the case of a `dsmd`-issued `reset` command, OpenBoot PROM provides parameters that force the domain to be booted to the Solaris operating environment.

`reset` normally sends a signal to all CPU ports of a specified domain. This is a hard reset and clears the hardware to a clean state. Using the `-x` option, however `reset` can send an XIR signal to the processors in a specified domain. This is done in software and is considered a soft reset. An error message is given if the virtual key switch is in the secure position. An optional `"Are you sure?"` prompt is given by default. For example:

```
sc0:sms-user:> reset -d C
Do you want to send RESET to domain C? [y|n]:y
RESET to processor 4.1.0 initiated.
RESET to processor 4.1.1 initiated.
RESET initiated to all processors for domain: C
```

For more information refer to the `reset` man page.

For information on resetting the main or spare SC see “SC Reset and Reboot” on page 106.

SMS software illuminates or darkens the indicator LEDs on LED-equipped hot-pluggable units (HPU) as necessary to reflect the correct state when the HPU is given a power-on reset.

Hardware Control

Hardware control functions are those that configure and control the platform hardware. Some functions are invoked on the domain.

Power-On Self-Test (POST)

System management services software invokes POST in two contexts.

1. At domain boot-time, POST is invoked to test and configure all functional hardware available to the domain.

   POST eliminates all hardware components that fail self-test and attempts to build a bootable domain from the functionally working hardware.

   POST provides extensive diagnostics to report hardware test results to help analyze failures. POST may be requested only to verify a domain configuration, and not test it, in situations where the domain is being rebooted with no indications that a hardware failure was the cause.

2. Before a DR operation to add a system board to a domain begins, POST is invoked to test and configure the system board components.

   If POST indicates that the candidate system board is functional, the DR operation can safely incorporate the system board into the physical (hardware) domain.

   Although POST is generally invoked automatically, there are user visible interfaces that affect automatic POST invocations:

   - The level of diagnostics testing that is performed by POST is increased from a nominal to maximum level using domain keyswitch control, `setkeyswitch`(1M) as described in “Virtual Keyswitch” on page 81.
   - You can add or remove components that you want POST to exclude from the hardware configuration by using blacklist files. These editable files are described in “Blacklist Editing” on page 98.

     This gives you finer-grained control over the hardware components that are used in a domain than is allowed by the standard domain configuration interfaces that operate on DCUs, such as system boards.
   - `setkeyswitch` invokes POST to test and configure a domain. Nominal and maximum diagnostic test level settings are provided for use in booting the domain.
   - `addboard` and moveboard invoke POST to test and configure a system board in support of a DR operation to add that board to a running Solaris domain.

- LED-equipped FRUs with components that fail POST will have the fault LED illuminated on the FRU.

Blacklist Editing

SMS supports three blacklists: one for the platform, one for the domains; and the internal automatic system recovery (ASR) blacklist.

Platform and Domain Blacklisting

The editable blacklist files specify that certain hardware resources are to be considered unusable by POST. They will not be probed for, tested, or configured in the domain interconnect.

Usually these blacklist files are empty, and are not required to be present.

Blacklist capability in this context is used for resource management purposes.

Blacklisting temporarily limits the system configuration to less than all the hardware present. This has several applications, such as benchmarking, limiting memory use to make DR detach of the board faster, and varying the configuration for troubleshooting.

Sun Fire 15K POST supports two editable canonical blacklist files, one for the platform, one for the domain, located in:

`/etc/opt/SUNWSMS/config/platform/blacklist`

and

`/etc/opt/SUNWSMS/config/`*`domain_id`*`/blacklist`

The two files are considered logically concatenated.

Note – The blacklist file specifies resources based on physical location. If the component is physically moved, any corresponding blacklist entries must be changed accordingly.

Blacklist specifies blacklisted components logically, for example, by specifying their position and the blacklist remains on the component position through a hot-swap operation rather than following a specific component.

▼ To Blacklist a Component

1. **Log in to the SC.**

   You must have platform administrator or domain administrator or configurator privileges to edit the blacklist files.

2. **Type**:

```
sc0:sms-user:> disablecomponent [-d domain_id|domain_tag] location
```

where:

`-d` *domain_id* - ID for a domain. Valid domain_ids are A...R and case insensitive.

`-d` *domain_tag* - Name assigned to a domain using `addtag` (1M).

location - List of component locations comprised of:

board_loc/proc/bank/logical_bank,

board_loc/proc/bank/all_dimms_on_that_bank

board_loc/proc/bank/all_banks_on_that_proc

board_loc/proc/bank/all_banks_on_that_board

board_loc/proc

board_loc/cassette or *board_loc/bus*

All component locations are separated by forward slashes. The *location* forms are optional and are used to specify particular components on boards in specific locations.

Multiple *location* arguments are permitted separated by a space.

location	Valid form
board_loc	SB(0...17) IO(0...17) CS(0 | 1) EX(0...17)
Processor / Processor Pair (*proc*)	P(0...3) PP(0 | 1)
bank	B
logical_bank	L(0 | 1)
all_dimms_on_that_bank	D
all_banks_on_that_proc	B
all_banks_on_that_board	B
cassette	C(3 | 5)V(0 | 1)
bus	ABUS | DBUS | RBUS (0 | 1)

Processor locations indicate single processors or processor pairs. There are four possible processors on a CPU / Mem board. Processor pairs on that board are: procs 0 and 1, and procs 2 and 3.

The MaxCPU has two processors,: procs 0 and 1, and only one proc pair (PP0). `disablecomponent` exits and displays an error message if you use PP1 as a location for this board.

The HsPCI assemblies contain hot-swappable cassettes.

There are three bus locations: address, data and response.

▼ To Remove a Component From the Blacklist,

1. **Log in to the SC.**

2. **Type:**

```
sc0:sms-user:> enablecomponent [-d domain_id|domain_tag] location
```

where:

`-d` *domain_id* - ID for a domain. Valid domain_ids are A...R and case insensitive.

`-d` *domain_tag* - Name assigned to a domain using `addtag` (1M).

location - List of component locations comprised of:

board_loc/proc/bank/logical_bank,

board_loc/proc/bank/all_dimms_on_that_bank

board_loc/proc/bank/all_banks_on_that_proc

board_loc/proc/bank/all_banks_on_that_board

board_loc/proc

board_loc/cassette or *board_loc/bus*

All component locations are separated by forward slashes. The *location* forms are optional and are used to specify particular components on boards in specific locations.

Multiple *location* arguments are permitted separated by a space.

location	Valid form
board_loc	SB(0...17) IO(0...17) CS(0 | 1) EX(0...17)
Processor / Processor Pair (*proc*)	P(0...3) PP(0 | 1)
bank	B
logical_bank	L(0 | 1)
all_dimms_on_that_bank	D
all_banks_on_that_proc	B
all_banks_on_that_board	B
cassette	C(3 | 5)V(0 | 1)
bus	ABUS | DBUS | RBUS (0 | 1)

Processor locations indicate single processors or processor pairs. There are four possible processors on a CPU / Mem board. Processor pairs on that board are: procs 0 and 1, and procs 2 and 3.

The MaxCPU has two processors,: procs 0 and 1, and only one proc pair (PP0). `disablecomponent` exits and displays an error message if you use PP1 as a location for this board.

The HsPCI assemblies contain hot-swappable cassettes.

There are three bus locations: address, data and response.

For more information, refer to the `enablecomponent`(1M) and `disablecomponent`(1M) man pages.

ASR Blacklist

Hardware that has failed repeatedly, perhaps intermittently, needs to be excluded from subsequent domain configurations for many reasons. It may be some time before the component can be physically replaced. The failed component might be a subcomponent such as one processor on a CPU board. You do not want to lose the services of the rest of the component by powering it down until it can be replaced. If the failure is hard, you do not want to waste time having POST discover that every time it runs. If the failure is intermittent, you do not want POST to pass it, only to have it fail when the OS is running.

To this end, `esmd` creates and edits the ASR blacklist file. Components that have been powered off due to environmental conditions are automatically listed and excluded from POST. `poweron`, `setkeyswitch`, `addboard`, and `moveboard` query the ASR blacklist and you are prompted for action when an excluded component is found. If the excluded component has been replaced or repaired, you can choose to remove it from the blacklist and continue bringing up the component. For more information refer to the `enablecomponent`(1M), `disablecomponent`(1M,)and `showcomponent`(1M) man pages.

Power Control

The main SC has power control over the following components in the Sun Fire 15K rack:

- Sun Fire 15K system boards, see “Domain Configuration Units (DCU)” on page 58
- HsPCI adaptor slots on the Sun Fire 15K HsPCI I/O assembly
- CPU pairs
- System controllers (power off only)
- Centerplane support boards
- wPCI boards
- Expander boards
- 48V power supplies
- AC bulk power modules

- Fan trays

See “HPU LEDs” on page 106 for a description of power control in the Sun Fire 15K I/O racks.

SMS supports the domain Solaris command interface (`cfgadm`(1M)) by providing the `rcfgadm`(1M) command to request power on or off of the HPCI adaptor slots in a Sun Fire 15K HsPCI I/O assembly. For more information refer to the `rcfgadm` man page.

The keyswitch control interface, `setkeyswitch`, as described in “Virtual Keyswitch” on page 81 allows the user to power on or off the hardware assigned to a domain.

All power operations are logged by the power control software.

The power control software conforms to all hardware requirements for powering on or off components. For example, SMS checks for adequate power available before powering on components. The power control interfaces will not perform a user-specified power on or power off operation if it violates a hardware requirement. Power operations that are performed contrary to hardware requirements or hardware recommended procedures are noted in the message logs.

By default, the power control software refuses to perform power operations that will affect running software. The power control user interfaces include methods to override this default behavior and forcibly complete the power operation at the cost of crashing running software. The use of these forcible overrides on power operations are noted in the message logs.

As described in “HPU LEDs” on page 106, SMS illuminates or darkens the indicator LEDs on LED-equipped HPUs, as necessary, to reflect the correct state when the HPU is powered on or off.

Fan Control

`esmd` provides the fan speed control for Sun Fire 15K fans. In general, fan speeds are set to the lowest speed that provides adequate cooling so as to minimize noise levels.

Hot-Swap

Hot-swap refers to the ability to physically insert or remove a board from a powered-on platform, actively running one or more domains without affecting those domains. During a hot-swap operation, the board is isolated from all domains.

The term for a hardware component that may be hot-swapped is hot-pluggable unit (HPU). The `OK to remove` indicator LED on an HPU is illuminated when it can be safely unplugged; see “HPU LEDs” on page 106 for more information about the `OK to remove` LEDs. Board presence registers indicate whether an HPU is present or absent and sense an HPU plug or unplug.

The Sun Fire 15K HsPCI I/O assemblies are equipped with `OK to remove` indicator LEDs associated with the slots into which HsPCI I/O assemblies are plugged. Each slot is equipped with a hot-plug controller that controls power to the slot and can detect presence of an adaptor in the slot. However, unlike SMS support for other Sun Fire 15K HPUs, the software that controls hot-swap for the HsPCI I/O assemblies is part of the Solaris environment on the domain.

SMS allows you to power on and off the adaptor slots.

SMS software provides software interfaces, invocable from the domain, to control hardware devices associated with the adaptor slots on I/O boards.

For the purposes of the remaining hot-swap discussion in this section, HPUs do not include hot-swappable I/O adaptors.

SMS software provides support as necessary to allow hot-swap servicing of all HPUs in the Sun Fire 15K rack.

Once an HPU is isolated from all domains the only software support required for hot-swap is power-off control.

Dynamic reconfiguration (DR) is used to isolate DCUs (system boards) from a domain by DR detaching the DCU.

Hot-Unplug

When an HPU is unplugged, the presence indicator for the HPU detects its absence resulting in a change in hardware configuration status as described in “Hardware Configuration” on page 123.

The expected mode of user interaction during hot unplug is as follows:

- Go directly to the HPU you wish to unplug. If the HPU indicator LEDs show that it is NOT `OK to remove`, request that the HPU be powered off using the `poweroff` command. If the power-off function discovers that the HPU is in use by a domain, the power-off function will fail, indicating that you first must use DR to remove the HPU from active use. Refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide* for more information.

Hot-Plug

The presence of a newly inserted HPU will be detected and reported as a change in hardware configuration status as described in “Hardware Configuration” on page 123.

SC Reset and Reboot

The SC supports software-initiated resets for the main and spare, providing the same functionality as external reset buttons on the system controller. Typically, an SC might be reset after failover. It is possible for the main SC software to reset the spare SC, if present, and vice versa. An SC cannot reset itself.

▼ To Reset the Main or Spare SC

`resetsc` (1M) sends a reset signal to the other SC. If the other SC is not present, `resetsc` exits with an error.

1. **Type**:

```
sc0:sms-user:> resetsc
“About to reset other SC. Are you sure you want to continue?” (y
or [n])? y
```

For more information, refer to the `resetsc` man page.

HPU LEDs

The LEDs reflect the status of the hot-pluggable units (HPU). LEDs come in groups of three:

- The operating indicator LED is illuminated when power is on.
- The `OK to remove` LED is illuminated when an HPU can be unplugged.
- The fault LED is illuminated when a hardware fault has been discovered in an HPU.

This section describes the LED control policies that are followed by SMS software for the HPUs.

Except for the system controllers, all Sun Fire 15K HPUs are powered on and tested under control of the SMS software that runs on the main system controller.

To a certain extent, the design of the LEDs, especially their initial state upon power-on-reset, is based upon the assumption that POST is automatically initiated at power-on-reset. The only Sun Fire 15K HPUs that meet this assumption are the system controllers. Powering on a system controller causes the processor to begin executing SC-POST code from PROM.

For all other HPUs, some are tested by POST and some are tested (or monitored) by SMS software, and although it is generally the case that testing follows shortly after power on, it is not always so.

Furthermore, it is possible that POST can be run multiple times on a power-on HPU that is being dynamically reconfigured from one domain to another. It is also possible that POST and SMS can both detect faults on the same physical HPU. These differences in power and test control between the system controllers and other Sun Fire 15K HPUs result in different policies proposed to manage them.

The system controller provides three sets of HPU LEDs:

- The state of the SC as a whole
- The state of the CP1500 slot
- The state of the SC spare slot

When the Sun Fire 15K rack is powered on, power is supplied to the system controllers. The operating indicator LED, and the `OK to remove` indicator LEDs are, appropriately, initialized by the hardware. All three fault LEDs are illuminated so that the fault LEDs correctly reflect a fault, should there be a problem that prevents SC-POST from running.

SMS software, upon powering off the spare system controller, extinguishes the operating indicator LED, and illuminates the `OK to remove` indicator LEDs on the spare system controller. SMS software cannot adjust the operating indicator or `OK to remove` indicator LEDs after powering off the main SC, where the software is running.

SC-POST does the following:

- Upon completing testing, the SC with no faults found, extinguishes the SC fault indicator LED.
- Upon completing testing the HPCI slot with no faults found, SC-POST extinguishes the SC spare slot fault LED.
- Upon completing testing the control board with no faults found at the control board, the SC main, or the SC spare slot, SC-POST extinguishes the SC fault LED.

SC-OpenBoot PROM firmware and SMS software illuminate the proper fault LED(s) on the system controller after detecting a hardware error.

The following policies are used to manage LEDs on HPUs other than the system controllers.

- On every LED-equipped non-SC HPU within the Sun Fire 15K rack, SMS assures that the operating indicator LED is steadily illuminated when power is applied to the HPU.
- On every LED-equipped non-SC HPU within the Sun Fire 15K, SMS assures that the `OK to remove` indicator LED is steadily illuminated only when the HPU can be safely unplugged. Safety considerations apply both to the person unplugging the HPU and to preserving the correct and continuing operation of Sun Fire 15K hardware and any running software.

Note – Sun Fire 15K system correctly illuminates the operating indicator LED and correctly darkens the `OK to remove` indicator LEDs when HPUs are powered on or given a power-on-reset.

- The management of the fault LEDs and their user-visible behavior differs most between the SC and non-SC HPUs.

  On the SC, the fault LEDs are illuminated at power on, maintained on during testing, and then extinguished if no fault is found.

  Faults detected after SC-POST can cause later fault LED illumination.

  So, except for the brief period when the SC is being tested by POST, the fault LEDs on the SC indicate that a fault has occurred since power on. The same holds true (an illuminated fault LED indicates that a fault has been detected since power on) for non-SC HPUs. Every LED-equipped non-SC HPU within the Sun Fire 15K system, SMS, upon power on or power on reset applied to that HPU, ensures that the fault indicator LED is extinguished.
- When directed to do so by POST, “Power-On Self-Test (POST)” on page 97, or the hardware monitoring software, “Environmental Events” on page 156, “Hardware Error Events” on page 159, and “SC Failure Events” on page 161, SMS steadily illuminates the fault indicator LED on an HPU. The fault indicator remains illuminated until the next power on or power-on-reset clears it, as described above in “HPU LEDs”.

CHAPTER 6

Domain Services

Sun Fire 15K system hardware incorporates internal, private point-to-point Ethernet connections between the SC and each domain. This network, called the Management Network (MAN), is used to provide support services for each domain. This chapter describes those services.

Management Network Overview

The Management Network (MAN) function maintains the private point-to-point network connections between the SC and each domain. No packets addressed to one domain can be routed along the network connection between the SC and another domain (FIGURE 6-1).

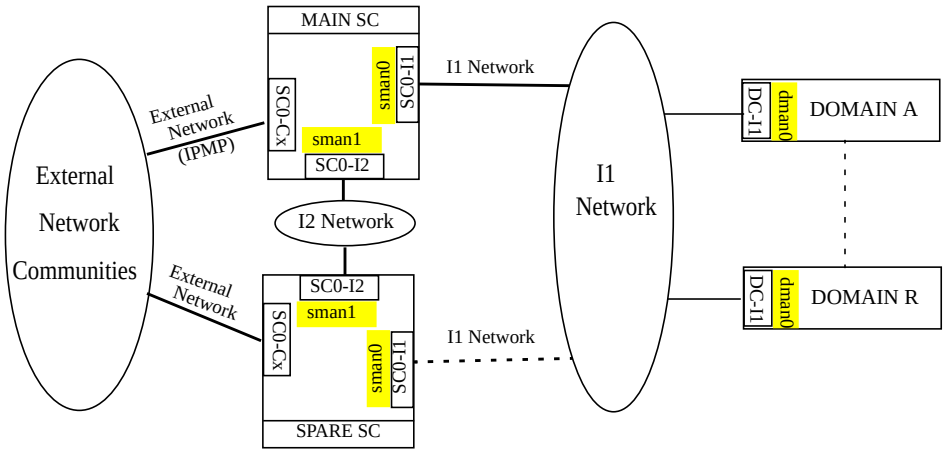


FIGURE 6-1 Management Network Overview

I1 Network

The hardware built into the Sun Fire 15K chassis to support MAN is complex. It includes 18 network interfaces (NICs) on each SC that are connected in a point-to-point fashion to NICs located on each of the 18 expander IO slots. Using this design, the number of point-to-point Ethernet links between an SC and a given DSD vary based on the number of IO boards configured in that DSD. Each NIC from the SC connects to a hub and NIC on the I/O board. The NIC is an internal part of the I/O board and not a separate adapter card. Likewise, the Ethernet hub is on the I/O board. The hub is intelligent and can collect statistics. All of these point-to-point links are collectively called the I1 network. Since there can be multiple I/O boards in a given domain, multiple redundant network connections from the SC to a domain are possible.

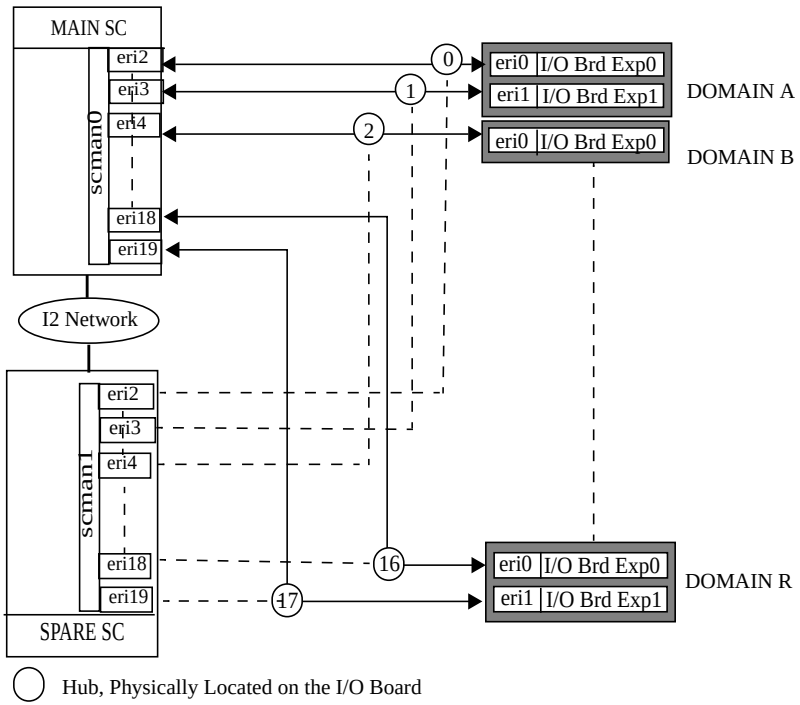


FIGURE 6-2 I1 Network Overview

Note – The I1 MAN network is a *private* network, not a general purpose network. No external IP traffic should be routed across it. Access to MAN is restricted to the system controller and the domains.

On the SC, MAN software creates a meta-interface for the I1 network presenting to the Solaris operating environment a single network interface, `scman0`. For more information refer to the Solaris `scman`(1M) man page.

MAN software detects communication errors and automatically initiates a path switch, provided an alternate path is available. MAN software also enforces domain isolation of network traffic on the I1 network. Similar software operates on the domain side.

I2 Network

There is also an internal network between the two system controllers consisting of two NICs per system controller. This network is called the I2 network. It is a private SC-to-SC network and is entirely separate from the I1 network.

MAN software creates a meta-interface for the I2 network as well. This interface is presented to the Solaris software as `scman1`. As with the I1 network, I2 has a mechanism for detecting path failure and switching paths, providing an alternative is available.

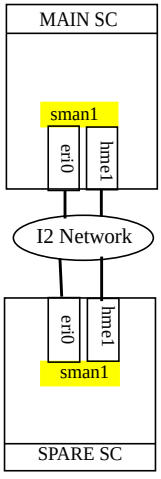


FIGURE 6-3 I2 Network Overview

The virtual network adapter on the SC presents itself as a standard network adapter. It can be managed and administered just like any other network adapter (for example, qfe, hme). The usual system administration tools such as `ndd`(1M), `netstat`(1M), and `ifconfig`(1M), can be used to manage the virtual network adapter. Certain operations of these tools (for example, changing the Ethernet address) should not be allowed for security reasons.

MAN operates and is managed as an IP network with special characteristics (for example, IP forwarding is disallowed by the MAN software). As such, the MAN operation is the same as any other IP network, with the noted exception documented above. Domains can be connected to your network depending on your site configuration and security requirements. Connecting domains is not within the scope of this document, refer to the *Solaris 8 System Administration Guide, Volume 3* .

External Network Monitoring

External Network Monitoring for the Sun Fire 15K system provides highly available network connections from the SCs to customer networks called communities. This feature is built on top of the IP Network Multipathing (IPMP) framework provided in Solaris 8. For more information on IPMP refer to the *IP Network Multipathing Administration Guide.*

External networks are made up of two communities. During installation, user communities are connected by physical cable to the RJ45 jacks on the SC connecting a node to the network.

For more information on connecting external networks, refer to the *Sun Fire 15K Site Planning Guide.*

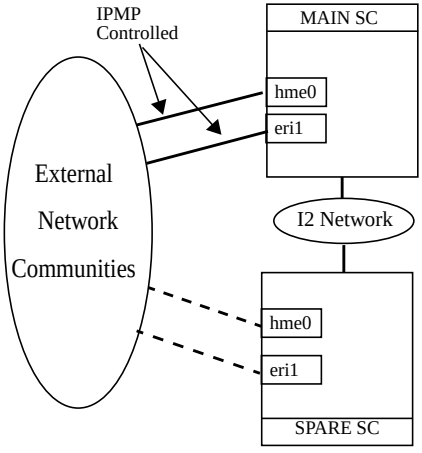


FIGURE 6-4 External Network Overview

The term *community* refers to an IP network at your site. For example, you may have an engineering community and an accounting community. A *community name* is used as the *interface group name.* An *interface group* is a group of network interfaces that attach to the same community.

Configuring External Network Monitoring requires allocating several additional IP addresses for each system controller.

The addresses can be categorized as follows:

Test Addressees - These are IP addresses that are assigned to the external network interfaces on each system controller (`hme0` and `eri1`). Each IP test address is used to test the health of the particular network interface to which it is assigned. There is one IP test address assigned to each network interface. They are permanently associated with a particular network interface. If the network interface fails, the IP test address associated with the network interface becomes unreachable.

Note – *In the case of IPv6, test addresses are not used. The link local address is used.*

Failover Addresses - There are two types of failover addresses:

SC Path Group Specific Addresses - These are IP addresses that are assigned to a particular interface group on each system controller. These IP addresses are used to provide highly available IP connectivity to a particular system controller for a given community. The SC path group specific address is reachable as long as one or more network interfaces in the interface group is functioning.

Note – An SC path group specific address is not needed if there is only one network interface in an interface group. Since there is no other network interface in the group to failover to, only the test addresses and the community failover addresses are required.

Community Failover Addresses- These are IP addresses that are assigned to a particular community on the MAIN SC (e.g. Community C1). They are used to provide IP connectivity to the MAIN SC, regardless of whether that is SC 0 or SC1.

All external software should reference the community failover address when communicating with the SC. This address will always connect to the MAIN SC. That way, if a failover occurs, external clients do not need to alter their configuration to reach the SC. For more information on SC failover, see “SC Failover” on page 127.

MAN Daemons and Drivers

For more information on the MAN daemon and device drivers, refer to the `mand`(1M), and Solaris `scman`(1M) and `dman`(1M) man pages. See also “Management Network Daemon” on page 42.

Management Network Services

The primary network services that MAN provides between the SC and the domains are:

- Domain consoles
- Message Logging
- Dynamic Reconfiguration (DR)
- Network boot/Solaris installation
- System Controller (SC) heartbeats

Domain Console

The software running in a domain (OpenBoot PROM, `kadb`, and Solaris software) uses the system console for critical communications.

The domain console supports a login session and is secure, since the default configuration of the Solaris environment allows only the console to accept `superuser` logins. Domain console access is provided securely to remote administrators over a possibly public network.

The behavior of the console reflects the health of the software running in the domain. Character echo for user entries are nearly equivalent to that of a 9600 baud serial terminal attached to the domain. Output characters that are not echoes of user input are typically either the output from an executed command, a command interpreter or unsolicited log messages from the Solaris software. Activity on other domains or SMS support activity for the domain do not noticeably alter user entry echo response latency.

You can run `kadb` on the domain's Solaris software from the domain console. Interactions with OpenBoot PROM running on a domain use the domain console. The console can serve as the destination for log messages from the Solaris software; refer to `syslog.conf`(4). The console is available when software (Solaris, OpenBoot PROM, `kadb`) is running on the domain.

You can open multiple connections to view the domain console output. However, the default is an exclusive *locked* connection.

For more information see "SMS Console Window" on page 6.

A domain administrator can forcibly break the domain console connection held by another.

You can forcibly break into OpenBoot PROM or `kadb` from the domain console, however it is not recommended. (This is a replacement for the physical `L1-A` or `STOP-A` key sequence available on a Sun SPARC™ system with a physical console.) SMS captures console output history for subsequent analysis of domain crashes. A snapshot of the last console output for every domain is available in `/var/opt/SUNWSMS/adm/`*domain_id*`/console`.

The Sun Fire 15K system provides the hardware to either implement a shared-memory console or implement an alternate network data path for console. The hardware utilized for shared-memory console imposes less direct latency upon console data transfers, but is also used for other monitoring and control purposes for all domains, so there is a risk of latency introduced by contention for the hardware resources.

MAN provides private network paths to securely transfer domain console traffic to the SC, see “Management Network Services” on page 114. The console has a dual-pathed nature so that at least one path provides acceptable console response latency when the Solaris software is running. The dual-pathed console is robust in the face of errors. It detects failures on one domain console path and fails over to the other domain console path automatically. It supports user-directed selection of the domain console path to use.

`smsconfig`(1M) is the SC configuration utility that initially configures or later modifies the hostname, IP address, and netmask settings used by management network daemon, `mand`(1M). See “Management Network Daemon” on page 42.

`mand` initializes and updates these respective fields in the platform configuration database (pcd).

`mand` is automatically started by `ssd`. The management network daemon runs on the main SC in *main* mode and on the spare SC in *spare* mode.

For more information, refer to the SMS `console`(1M), `mand`(1M), and `smsconfig`(1M) man pages and the Solaris `dman` and (1M), `scman`(1M) man pages.

Message Logging

When configured to do so, MAN transports copies of important syslog messages from the domains to disk storage on the SC in order to facilitate failure analysis for crashed or unbootable domains. For more information, see “Log File Maintenance” on page 146.

Dynamic Reconfiguration

Support for automatic network reconfiguration for a DR operation that removes the I/O board network connection endpoint from the domain, is initiated from the SC.

The MAN software layer is used to simplify the interface to the MAN hardware. MAN software handles the aspects of dynamic reconfiguration (DR) used by a DSD without requiring network configuration work by the domain or platform administrator.

Software in the domains using MAN need not be aware of which SC is currently the main SC. For more information on dynamic reconfiguration, refer to the *System Management Services (SMS) 1.2 Dynamic Reconfiguration User Guide*.

Network Boot and Solaris Software Installation

The SC provides network Solaris boot services to each domain.

Note – This is not intended to imply that diskless Sun Fire 15K domains can be supported entirely by network services from the SC; the SC network boot service is intended primarily for recovery after a catastrophic disk failure on the domain.

When Solaris software is first installed on a domain, the network interface connecting it to the MAN is automatically created for subsequent system reboots. There are no additional tasks required by the domain administrator to configure or use MAN.

MAN is configured as a private network. A default address assignment for the management network is provided using the IP address space reserved for private networks. You can override the default address assignment for MAN to handle the case where the Sun Fire 15K is connected to a private customer network that already uses the selected MAN default IP address range.

The SC supports simultaneous network boots of domains running at least two different versions of Solaris software.

The SC provides software installation services to, at most, one domain at a time.

SC Heartbeats

The I2 network is the intersystem controller communication. This is also called the heartbeat network. SMS failover mechanisms on the main use this network as one means of determining the health of the spare SC. For more information see Chapter 8 “SC Failover”.

CHAPTER 7

Domain Status

Status functions return measured values that characterize the state of the server hardware or software. As such, these functions are used both to provide values for status displays and input to monitoring software that periodically polls status functions and verifies that the values returned are within normal operational limits. Monitoring and event detection functions that use the status functions are described in this chapter.

Software Status

The software state consists of status information provided by the software running in a domain. The identity of the software component currently running (for example, POST, OpenBoot PROM, or Solaris software) is available. Additional status information is available (booting, running, panicking).

SMS software provides the following command(s) to display the status of the software, if any, currently running in a domain.

- `showboards`
- `showdevices`
- `showenvironment`
- `showobpparams`
- `showplatform`
- `showxirstate`

Status Commands

showboards Command

showboards(1M) displays the assignment information and status of the DCUs. These include the following: Location, Power, Type of board, Board status, Test status and Domain.

If no options are specified, showboards displays for the platform administrator, all DCUs including those that are assigned or available. For the domain administrator or configurator, showboards displays only those DCUs for those domains for which the user has privileges, including those boards that are assigned or available and in the domain's available component list.

If *domain_id* | *domain_tag* is specified, this command displays which DCUs are assigned or available to the given domain. If the -a option is used, showboards displays all boards including DCUs.

For examples and more information, see "To Obtain Board Status" on page 66 and refer to the showboards manpage.

showdevices Command

showdevices(1M) displays configured physical devices on system boards and the resources made available by these devices. Usage information is provided by applications and subsystems that are actively managing system resources. The predicted impact of a system board DR operation may be optionally displayed by performing an offline query of managed resources.

showdevices gathers device information from one or more Sun Fire 15K domains. The command uses the dca(1M) as a proxy to gather the information from the domains.

For examples and more information, see "To Obtain Device Status" on page 80 and refer to the showdevices manpage.

showenvironment Command

showenvironment(1M) displays environmental data including: Location, Device, Sensor, Value, Unit, Age, Status. For fan trays, Power, Speed and Fan Number are displayed. For bulk power, the Power, Value and Unit and Status are shown.

If a domain *domain_id* | *domain_tag* is specified, environmental data relating to the domain is displayed, providing that the user has domain privileges for that domain. If a domain is not specified, all domain data permissible to the user will be displayed.

DCUs (for example, CPU, I/O) belong to a domain and you must have domain privileges to view their status. Environmental data relating to such things as fan trays, bulk power, or other boards are displayed without domain permissions. You can also specify individual reports for: temperatures, voltages, currents, faults, bulk power status, and fan tray status with the `-p` option. If the `-p` option is not present, all reports will be shown.

For examples and more information, see “Environmental Status” on page 124 and refer to the `showenvironment` man page.

`showobpparams` Command

`showobpparams`(1M) displays OpenBoot PROM bringup parameters. `showobpparams` allows a domain administrator to display the virtual `NVRAM` and `REBOOT` parameters passed to OpenBoot PROM by `setkeyswitch`(1M).

For examples and more information, see “Setting the OpenBoot PROM Variables” on page 85 and refer to the `showobpparams` man page.

`showplatform` Command

`showplatform`(1M) displays the available component list and domain state of each domain.

A domain is identified by a *domain_tag* if one exists. Otherwise it is identified by the *domain_id*, a letter in the set A - R. The letter set is case insensitive. The Solaris *hostname* is displayed if one exists. If a *hostname* has not been assigned to a domain, `Unknown` is printed.

The following is a list of domain statuses:

- Unknown
- Powered Off
- Keyswitch Standby
- Running Domain POST
- Loading OBP
- Booting OBP
- Running OBP
- In OBP Callback
- Loading Solaris

- Booting Solaris
- Domain Exited OBP
- OBP Failed
- OBP in sync Callback to OS
- Exited OBP
- In OBP Error Reset
- Solaris Halted in OBP
- OBP Debugging
- Environmental Domain Halt
- Booting Solaris Failed
- Loading Solaris Failed
- Running Solaris
- Solaris Quiesce In-Progress
- Solaris Quiesced
- Solaris Resume In-Progress
- Solaris Panic
- Solaris Panic Debug
- Solaris Panic Continue
- Solaris Panic Dump
- Solaris Halt
- Solaris Panic Exit
- Environmental Emergency
- Debugging Solaris
- Solaris Exited
- In Recovery

For examples and more information, see “To Obtain Domain Status” on page 67 and refer to the `showplatform` man page.

`showxirstate` Command

`showxirstate`(1M) displays CPU dump information after sending a reset pulse to the processors. This save state dump can be used to analyze the cause of abnormal domain behavior. `showxirstate` creates a list of all active processors in that domain and retrieves the save state information for each processor.

`showxirstate` data resides, by default, in `/var/opt/SUNWSMS/adm/`*domain_id*`/dump`.

For examples and more information, refer to the `showxirstate` man page.

Solaris Software Heartbeat

During normal operation, the Solaris environment produces a periodic heartbeat indicator readable from the SC. `dsmd` detects the absence of heartbeat updates for a running Solaris system as a hung Solaris. Hangs are not detected for any software components other than the Solaris software.

Note – The Solaris software heartbeat should not be confused with the SC-to-SC (hardware) heartbeat or the heartbeat network, both used to determine the health of failover. For more information see, “SC Heartbeats” on page 117.

The only reflection of the Solaris heartbeat occurs when `dsmd` detects a failure to update the Solaris heartbeat of sufficient duration to indicate that the Solaris software is hung. Upon detection of a Solaris software hang, `dsmd` will conduct an ASR.

Hardware Status

The hardware status functions report information about the hardware configuration, hardware failures detected, and platform environmental state.

Hardware Configuration

The following hardware configuration status is available from the Sun Fire 15K system management software:

- Hardware components physically present on each board (as detected by POST)
- Hardware components not in use because they failed POST
- Presence or absence of all hot-pluggable units (HPUs) (for example, system boards)
- Hardware components not in use because they were on the blacklist when POST was invoked (see “Power-On Self-Test (POST)” on page 97)
- Contents of the SEEPROM for each FRU including the part number and serial number

Note – The hardware configuration status available to SMS running on the SC includes no information about the I/O configuration; such as, where I/O adaptors are plugged in and what devices are attached to those I/O adaptors. Such information is available only to the software running on the domain that owns the I/O adaptor.

The hardware configuration supported by functions described in this section exclude I/O adaptors and I/O devices. `showboards` displays all hardware components that are present.

As described in "Blacklist Editing" on page 98, the current contents of the component blacklist(s) can always be viewed and altered.

Environmental Status

The following hardware environmental measurements are available:

- Temperatures
- Power voltage and amperage
- Fan status (stopped, low-speed, high-speed, failed)
- Power status
- Faults

`showenvironment` displays every environmental measurement that can be taken within the Sun Fire 15K rack.

▼ To Display the Environment Status for Domain A

1. **Log in to the SC.**

   Platform administrators can view any environment status on the entire platform. Domain administrators can see the environment status only for those domains for which they have privileges.

2. **Type:**

```
sc0:sms-user:> showenvironment -d A
```

As described in "HPU LEDs" on page 106, the operating indicator LEDs on Sun Fire 15K HPUs visibly reflect that the HPUs are powered on and the `OK to remove` indicator LEDs visibly reflect those that can be unplugged.

Hardware Error Status

`dsmd` monitors the Sun Fire 15K hardware operational status and reports errors. The occurrence of some errors are directly reported to the SC (for example, the error register(s) in every ASIC propagate to the SBBC on the SC that provides an error summary register). Although the occurrence of some errors is indicated by an interrupt delivered to the SC, some error states may require the SC to monitor hardware registers for error indications. When a hardware error is detected, `esmd` follows the established procedures for collecting and clearing the hardware error state.

The following types of errors can occur on Sun Fire 15K hardware:

- Domain stops, fatal hardware errors that terminate all hardware operations in a domain
- Record stops that cause the hardware to stop collecting transaction history when a data transfer error (for example, parity) occurs
- SPARC processor error conditions such as RED_state / watchdog reset
- Nonfatal ASIC-detected hardware failures

Hardware error status is generally not reported as a status. Rather, event handling functions perform various actions when hardware errors occur such as logging errors, initiating ASR, and so forth. These functions are discussed in Chapter 9 “Domain Events”.

Note – As described in“HPU LEDs” on page 106, the fault LEDs, after POST completion, identify Sun Fire 15K HPUs in which faults have been discovered since last powered on or submitted to a power on reset.

SC Hardware and Software Status

Proper operation of SMS depends upon proper operation of the hardware and the Solaris software on the SC. The ability to support automatic failover from the main to the spare system controller requires properly functioning hardware and software on the spare. SMS software running on the main system controller must either be functioning sufficiently to diagnose a software or hardware failure in a manner that can be detected by the spare or it must fail in a manner that can be detected by the spare.

SC-POST determines the status of system controller hardware. It tests and configures the system controller at power-on or power-on reset.

The SC will not boot if the SC fails to function.

If the control board fails to function, the SC will normally boot, but without access to the control board devices. The level of hardware functionality required to boot the system controller is essentially the same as that required for a standalone SC.

SC-POST writes diagnostic output to the SC console serial port (TTY-A). Additionally, SC-POST leaves a brief diagnostics status summary message in an NVRAM buffer that can be read by a Solaris driver and logged and/or displayed when the Solaris software boots.

SC firmware and software display information to identify and service SC hardware failures.

SC firmware and software provide a software interface that verifies that the system controller hardware is functional. This selects a working system controller as main in a high-availability SC configuration.

The system controller LEDs provide visible status regarding power and detected hardware faults as described in “HPU LEDs” on page 106.

Solaris software provides a level of self-diagnosis and automatic recovery (panic and reboot). Solaris software utilizes the SC hardware watchdog logic to trap hang conditions and force an automatic recovery reboot.

There are three hardware paths of communication between the SCs (two Ethernet connections, the heartbeat network, and one SC-to-SC heartbeat signal) that are used in the high-availability SC configuration by each SC to detect hangs or failures on the other SC.

SMS practices self-diagnosis and institutes automatic failure recovery procedures; even in non-high-availability SC configurations.

Upon recovery, SMS software either takes corrective actions as necessary to restore the platform hardware to a known, functional configuration or reports the inability to do so.

SMS software records and logs sufficient information to allow engineering diagnosis of single-occurrence software failures in the field.

SMS software takes a noticeable interval to initialize itself and become fully functional. The user interfaces behave predictably during this interval. Any rejections of user commands are clearly identified as due to system initialization with advice to try again after a suitable interval.

SMS software implementation uses a distributed client/server architecture. Any errors encountered during SMS initialization, due to attempts to interact with a process that has not yet completed initialization, are dealt with silently.

CHAPTER 8

SC Failover

SC failover maximizes Sun Fire 15K system uptime by adding high availability features to its administrative operations. A Sun Fire 15K system contains two SCs. Failover provides software support to a high-availability two SC system configuration.

The main SC provides all resources for the entire Sun Fire 15K system. If hardware or software failures occur on the main SC or on any hardware control path (for example, console bus interface or Ethernet interface) from the main SC to other system devices, then SC failover software automatically triggers a failover to the spare SC. The spare SC then assumes the role of the main and takes over all the main SC responsibilities. In a high-availability two SCs system configuration, SMS data, configuration, and log files are replicated on the spare SC. Active domains are not affected by this switch.

Overview

In the current high-availability SC configuration, one SC acts as a “hot spare” for the other.

Failover eliminates the single point of failure in the management of the Sun Fire 15K system. `fomd` identifies and handles as many multiple points failure as possible. Some failover scenarios are discussed in “Failure and Recovery” on page 138.

At anytime during SC failover, the failover process does not adversely affect any configured or running domains except for temporary loss of services from the SC.

In a high-availability SC system:

- If a fault (software or hardware) is detected on the main SC, `fomd` automatically fails over to the spare SC.

- If the spare SC detects that the main SC has stopped communicating with it, it initiates a takeover and assumes the role of main.

The failover management daemon (`fomd`(1M)) is the core of the SC failover mechanism. It is installed on both the main and spare SC.

The `fomd`:

- Determines an SC's role; main or spare.
- Requests the general health status of the remote SC hardware and software in the form of a periodic health status message request sent over the SMS management network (MAN) that exists between the two SCs.
- Checks and/or handles recoverable and unrecoverable hardware and software faults.
- Makes every attempt to eliminate the possibility of split-brain condition between the two SCs. [A condition is considered split-brain when both the SCs think they are the main SC.]
- Provides a recovery time from a main SC failure of between five and eight minutes. The recovery time includes the time for `fomd` to detect the failure, reach an agreement on the failure, and assume the main SC responsibilities on the spare SC.
- Logs an occurrence of an SC failover in the platform message log.

Services that would be interrupted during a SC failover include:

- All network connections
- Any SC-to-domain and domain-to-SC IOSRAM/Mailbox communication

You do not need to know the hostname of the main SC in order to establish connections to it. As part of configuring SMS (refer to the `smsconfig`(1M) man page), a logical hostname was created which will always be active on the main SC. Refer to the *Sun Fire™ 15K System Site Planning Guide* and the *System Management Services (SMS) 1.2 Installation Guide and Release Notes* for information on the creation of the logical hostnames in your network database.

Operations interrupted by an SC failover can be recovered after the failover completes. Reissuance of the interrupted operation causes the operation to resume and continue to completion.

All automated functions provided by `fomd` resume without operator intervention after SC failover. Any recovery actions interrupted before completion by the SC failover will restart.

Fault Monitoring

There are three types of failovers:

1. Main initiated

   A main initiated failover is where the `fomd` running on the main SC yields control to the spare SC in response to either an unrecoverable local hardware/software failure or an operator request.

2. Spare initiated (takeover)

   A spare initiated failover (takeover) is where the `fomd` running on the spare determines that the main SC is no longer functioning properly.

3. Indirect triggered takeover

   If the I2 network path between the SCs is down and there is a fault on the main, then the main switches itself to the role of spare and upon detecting this, the spare SC assumes the role of main.

In the last two scenarios, the spare `fomd` eliminates the possibility of a "split-brain" condition by resetting the main SC. A "split-brain" condition can occur if the spare thinks that the main SC is dead, when it is not.

Note – The OpenBoot PROM variable `auto-boot` must be set to `false` in order for this to work. (`False` is the SMS default.)

The spare SC resets the main. The reset main comes up to and stays at the `ok` prompt after POST runs. Since SMS does not startup, there is no possibility that the former main SC can come up as main while the other SC is main.

When a failover occurs, either software controlled or user forced, `fomd` deactivates the failover mechanism. This eliminates the possibility of repeatedly failing over back and forth between the two SCs.

File Propagation

One of the purposes of the `fomd` is propagation of data from the main SC to the spare SC through the interconnects that exist between the two SCs. This data includes configuration, data, and log files.

`fomd`:

- Propagates all native SMS files from the main to the spare SC at startup. These include all the domain data directories, the pcd configuration files, the `/etc/opt/SUNWSMS/config` directory, the `/var/opt/SUNWSMS/adm` platform and domain files, and the `.logger` files. Any user-created application files are not propagated unless specified in the cmdsync scripts.
- Monitors files that have been modified between propagation, and propagates those files.
- In the event of a failover, propagates all modified SMS files before the spare SC assumes its role as main.
- In the event of a spare SC takeover, `fomd` transfers all modified files over to the new main SC before the SC takes over.

  The I2 network must be operative for the transfer of data to occur.

Note – Any changes made to the network configuration on one SC using `smsconfig -m` must be made to the other SC as well. Network configuration is not automatically propagated.

Should both interconnections between the two SCs fail, failover can still occur provided main and spare SC accesses to the high-availability srams (HASram) remain intact. Due to the failure of both interconnections, propagation of SMS data can no longer occur, creating the potential of stale data on the spare SC. In the event of a failover, `fomd` on the new main keeps the current state of the data, logs the state and provides other SMS daemons/clients information of the current state of the data.

When either of the interconnects between the two SCs is healthy again, data is pulled over depending on the timestamp of each SMS files. If the timestamp of the file is earlier than the one on the now spare SC, it gets transferred over. If the timestamp of the file is later than the one on the spare SC, no action is taken.

Failover cannot occur when both of the following conditions are met:

- Both interconnects between the two SCs fail
- Access to both HASrams fail

This is considered a quadruple fault, and failover will be disabled until at least one of the links is restored.

Failover Management

Startup

For the failover software to function, both SCs must be present in the system. The determination of main and spare roles are based in part on the SC number. This slot number does not prevent a given SC from assuming either role, it is only meant to control how it goes about doing so.

If SMS is started on one SC first, it will become main. If SMS starts up on both SCs at essentially the same time, whichever SC first determines that the other SC either is not main or is not running SMS will become main.

There is one case during start-up where, if SC0 is in the middle of the start-up process, and it queries SC1 for its role and the SC1 role cannot be confirmed, SC0 will try to become main. SC0 will reset SC1 in this process. This is done to prevent both SCs from assuming the main role; a condition known as “split brain.” The reset will occur even if the failover mechanism is deactivated.

Main SC

Upon startup, the `fomd` running on the main SC begins periodically testing the hardware and network interfaces. Initially the failover mechanism is disabled (internally) until at least one status response has been received from the remote (spare) SC indicating that it is healthy.

If a local fault is detected by the main `fomd` during initial startup, failover occurs when all of the following conditions are met:

1. The I2 network was not the source of the fault.
2. The remote SC is healthy (as indicated by the health status response).
3. The failover mechanism has not been deactivated.

Spare SC

Upon startup, `fomd` runs on the spare SC and begins periodically testing the software, hardware, and network interfaces.

If a local fault is detected by the `fomd` running on the spare SC during initial startup, it informs the main `fomd` of its debilitated state.

Failover CLIs

`setfailover` Command

`setfailover` modifies the state of the SC failover mechanism. The default state is on. You can set failover to:

State	Definition
on	Enables failover for systems that previously had failover disabled due to a failover or an operator request. This option instructs the command to attempt to re-enable failover only. If failover cannot be re-enabled, subsequent use of the `showfailover` command indicates the current failure that prevented the enable.
off	Disables the failover mechanism. This prevents a failover until the mechanism is re-enabled.
force	Forces a failover to the spare SC. The spare SC must be available and healthy.

Note – In the event a patch must be applied to SMS 1.2, failover must be disabled before the patch is installed. Refer to *System Management Services (SMS) 1.2 Installation Guide and Release Notes.*

For more information and examples refer to the `setfailover` man page.

showfailover Command

showfailover allows you to monitor the state and display the current status of the SC Failover mechanism. The -v option displays the current status of all monitored components.

```
sc0:sms-user:> showfailover -v
SC Failover Status:       ACTIVE
Clock Phase Locked: ..... Yes
HASRAM Status (by location):
    HASRAM (CSB at CS0): .....Good
    HASRAM (CSB at CS1): .....Good
Status of xc12-sc0:
Role:.....MAIN
System Clock: .....Good
X1 Network:
    hme0: . .....Good
I2 Network: .....Good
System Memory: .....0.5%
Disk Status:
    /: .....1.4%
Console Bus Status:
    EXB at EX3: .....Good
    EXB at EX6: .....Good
    EXB at EX12: .....Good
    EXB at EX15: .....Good
Status of xc12-sc1:
Role: .....SPARE
System Clock: .....Good
X1 Network:
    hme0: .....Good
I2 Network: .....Good
System Memory: .....0.6%
Disk Status:
    /: .....1.4%
Console Bus Status:
    EXB at EX3: .....Good
    EXB at EX6: .....Good
    EXB at EX12: .....Good
    EXB at EX15: .....Good
```

The `-r` option displays the SC role: main, spare or unknown. For example:

```
sc0:sms-user:> showfailover -r
MAIN
```

If you do not specify an option, then only the state information is displayed:

```
sc0:sms-user:> showfailover
SC Failover: state
```

The failover mechanisms may be in one of three states: ACTIVE, DISABLED, and FAILED.

TABLE 8-1 Failover Mechanisms

State	Definition
ACTIVE	Identifies the failover mechanism as being enabled and functioning normally.
DISABLED	Identifies that the failover mechanism has been disabled due to the occurrence of a failover or an operator request (`setfailover` off)
FAILED	Identifies that the failover mechanism has detected a failure that prevents a failover from being possible.

In addition `showfailover` displays the state of each of the network interface links monitored by the failover processes. The display format is as follows:

```
network i/f device name: [GOOD | FAILED]
```

`showfailover` returns a failure string describing the failure condition. Each failure string has a code associated with it. The following table defines the codes, and associated failure strings.

TABLE 0-1 `Showfailover` Failure Strings

String	Explanation
None	No failure
M-SC / S-SC EXT NET	The main and spare SC external network interfaces have failed
S-SC CONSOLE BUS	A fault has been detected on the spare SC console bus path(s).
S-SC LOC CLK	The spare SC local clock has failed.
S-SC CLK NOT PHASE LOCKED	The spare SC clock is not phase locked with the main.
S-SC DISK FULL	The spare SC system is full.
TS-SC MEM EXHAUSTED	The spare SC memory / swap space has been exhausted.
S-SC SMS DAEMON	At least one SMS daemon could not be started / restarted on the spare SC.
NO CSBS POWERED ON	At least one CSB must be powered on.

For examples and more information, refer to the `showfailover` man page.

Command Synchronization

If an SC failover occurs during the execution of a command, you can restart the same command on the new main SC.

All commands and actions are recorded to do the following:

- Mark the start of a command or action.
- Remove or indicate the completion of a command or action.
- Keep any state transition and / or pertinent data which SMS can use to resume the command.

`fomd` provides:

- Command sync support for `dsmd`(1M) to automatically resume ASR reboots of any or all affected domain(s) after a failover.
- Command sync support for `hwad` to recover the old tunnel and the potential new tunnel after the failover.

- Command sync support for all SMS DR related daemons and CLIs to recover the last DR operation after a failover.

The four CLIs in SMS that require command sync support are `addboard`, `deleteboard`, `moveboard`, and `rcfgadm`.

Cmdsync CLIs

The *cmdsync* commands provide the ability to initialize a script or command with a cmdsync descriptor, update an existing cmdsync descriptor execution point, or cancel a cmdsync descriptor from the spare SC's list of recovery actions. Commands or scripts can also be run in a cmdsync envelope.

In the case of an SC failover to the spare, initialization of a cmdsync descriptor on the spare SC allows the spare SC to restart or resume the target script or command from the last execution point set. These commands only execute on the main SC, and have no effect on the current cmdsync list if executed on the spare.

Commands or scripts invoked with the cmdsync commands when there is no enabled spare SC will result in a no-op operation. That is, command execution will proceed as normal, but a log entry in the platform log will indicate that a cmdsync attempt has failed.

`initcmdsync` Command

`initcmdsync`(1M) creates a cmdsync descriptor. The target script or command and its associated parameters are saved as part of the cmdsync data. The exit code of the `initcmdsync` command provides a cmdsync descriptor that can be used in subsequent cmdsync commands to reference the action. Actual execution of the target command or script is not performed. For more information, refer to the `initcmdsync` (1M) man page.

`savecmdsync` Command

`savecmdsync`(1M) saves a new execution point in a previously defined cmdsync descriptor. This allows a target command or script to restart execution at a location associated with an identifier. The target command or script supports the ability to be restarted at this execution point, otherwise the restart execution is at the beginning of the target command or script. For more information, refer to the `savecmdsync` (1M) man page.

cancelcmdsync Command

cancelcmdsync(1M) removes a cmdsync descriptor from the spare restart list. Once this command is run, the target command or script associated with the cmdsync descriptor is not restarted on the spare SC in the event of a failover. Take care to insure that all target commands or scripts contain a initcmdsync command sequence as well as a cancelcmdsync sequence after the normal or abnormal termination flows. For more information, refer to the cancelcmdsync (1M) man page.

runcmdsync Command

runcmdsync(1M) executes the specified target command or script under a cmdsync wrapper. You cannot restart at execution points other than the beginning. The target command or script is executed through the system command after creation of the cmdsync descriptor. Upon termination of the system command, the cmdsync descriptor is removed from the cmdsync list, and the exit code of the system command returned to the user. For more information, refer to the runcmdsync (1M) man page.

showcmdsync Command

showcmdsync(1M) displays the current cmdsync descriptor list. For more information, refer to the showcmdsync (1M) man page.

Failure and Recovery

In a high-availability configuration, `fomd` manages the failover mechanism on the local and remote SCs. `fomd` detects the presence of local hardware and software faults and determines the appropriate action to take.

`fomd` is responsible for detecting faults in the following categories:

a	All relevant hardware buses that are local to the SC Control board (CB)/CPU board
b	The external network interfaces
c	The I2 network interface between the SCs
d	Unrecoverable software failures. This category is for those cases where an SMS software component (daemon) crashes and cannot be restarted after three attempts; the file system is full; the heap is exhausted and so forth.

FIGURE 8-1 illustrates the failover fault categories.

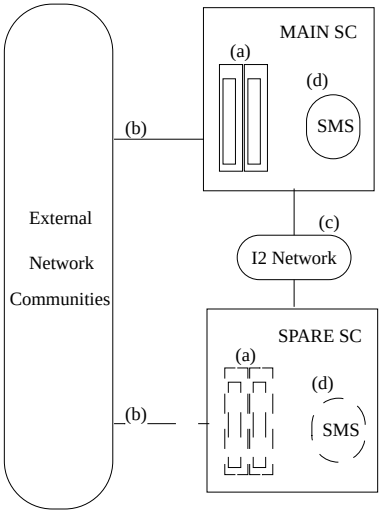


FIGURE 8-1 Failover Fault Categories

The following table illustrates how faults in the above-mentioned categories affect the failover mechanism. Assume that the failover mechanism is activated.

TABLE 8-2 High-Level Fault Overview

Failure Point	Main SC	Spare SC	Failover	Notes
a	X		X	Failover to spare occurs.
a		X	disables	No effect on the main SC, but the spare SC has suffered a hardware fault so failover is disabled.
b	X			Failover to spare.
b		X	No effect	The fact that the spare SC external network interfaces have failed does not affect the failover mechanism.
c			No effect	Main and spare SC log the fault.
d	X		X	Failover to the spare SC assuming that it is healthy.
d		X	Disables	Failover is disabled because the spare SC is deemed unhealthy at this point.

Failover on Main SC (Main Controlled Failover)

The following lists events for the **main** `fomd` during SC failover in order.

1. Detects the fault.
2. Stops generating heartbeats.
3. Tells the remote failover software to start a takeover timer. The purpose of this timer is to provide an alternate means for the remote (spare) SC to takeover if for any reason the main hangs up and never reaches Number 10.
4. Starts the SMS software in spare mode.
5. Removes the logical IP interface.
6. Enables the console bus caging mechanism.
7. Triggers propagation of any modified SMS files to the spare SC/HASrams.
8. Stops file propagation monitoring.
9. Logs a failover event.

10. Notifies remote (spare) failover software that it should assume the role of main. If the takeover timer expires before the spare is notified, the remote SC will takeover on its own.

The following lists the order of events for the **spare** `fomd` during failover.

1. Receives message from the main `fomd` to assume main role, or the takeover timer expires. If the former is true, then the takeover timer is stopped.
2. Resets the old main SC.
3. Notifies `hwad` to configure itself in the main role.
4. Assumes the role of main.
5. Starts generating heartbeat interrupts.
6. Configures the logical IP interface.
7. Disables the console bus caging mechanism.
8. Starts the SMS software in **main** mode.
9. Logs a role reversal event, spare to main.
10. The spare SC is now the main and `fomd` deactivates the failover mechanism.

Fault on Main SC (Spare Takes Over Main Role)

In this scenario the spare SC takes main control in reaction to the main SC going away. The most important aspect of this type of failover is the prevention of the split-brain condition. Another assumption is that the failover mechanism is not deactivated. If this is not the case, then no takeover can occur.

The spare `fomd` does the following:

- Notices the main SC is not healthy.

  From the spare `fomd` perspective, this phenomenon can be caused by two conditions; the main SC is truly dead, and/or the I2 network interface is down.

  In the former case, a failover is needed (provided that the failover mechanism is activated) while in the latter it is not. To identify which is the case, the spare `fomd` polls for the presence of heartbeat interrupts from the main SC to determine if the main SC is still up and running. The polling period for this is configurable. As long as there are heartbeat interrupts being received, and/or the failover mechanism is deactivated and/or disabled, no failover occurs. In the case where no interrupts are detected, but the failover mechanism is deactivated, the spare `fomd` does not attempt to take over unless the operator manually activates the

failover mechanism using the CLI command, `setfailover`. Otherwise, if the spare SC is healthy, the spare `fomd` proceeds to take over the role of main as listed.

- Initiates a takeover by resetting the remote (main) SC.

The following lists the events for the **spare** `fomd`, in order, during failover.

1. Reconfigures itself as main. This includes taking over control of the I2C bus, configuring the logical main SC IP address, and starting up the necessary SMS software daemons.
2. Starts generating heartbeat interrupts.
3. Configures the logical IP interface.
4. Disables console bus caging.
5. Starts the SMS software in main mode.
6. Configures the darb interrupts.
7. Logs a takeover event.
8. The spare `fomd`, now the main, deactivates the failover mechanism.

I2 Network Fault

The following lists the events, in order, that occur after an I2 network fault.

1. The **main** `fomd` detects the I2 network is not healthy.
2. The **main** `fomd` stops propagating files and checkpointing data over to the spare SC.

1. The **spare** `fomd` detects the I2 network is not healthy. From the spare `fomd` perspective, this phenomenon can be caused by two conditions; the main SC is truly dead, and/or the I2 network interface is down. In the former case, the corrective action is to fail over, while in the latter it is not. To identify which is the case, the `fomd` starts polling for the presence of heartbeat interrupts from the main SC to determine if the main SC is still up and running. If heartbeat interrupts are present, then the `fomd` keeps the spare as spare.
2. The **spare** `fomd` clears out the checkpoint data on the local disk.

Fault on Main SC (I2 Network Is Also Down)

The following lists the events, in order, that occur after a fault on the **main** SC.

1. The **main** `fomd` detects the fault.

   If the last known state of the spare SC was good, then the main `fomd` stops generating heartbeats. Otherwise failover does not continue.

   If the access to the console bus is still available, main failover software finishes propagating any remaining critical files to HASram and flushes out any or all critical state information to HASram.

2. The **main** `fomd` reconfigures the SMS software into spare mode.

3. The **main** `fomd` removes the logical main SC IP address.

4. The **main** `fomd` stops generating heartbeat interrupts.

Fault Recovery and Reboot

I2 Fault Recovery

The following lists the events, in order, that occur during an I2 network fault recovery.

1. The **main** `fomd` detects the I2 network is healthy.

   If the spare SC is completely healthy as indicated in the health status response message, the `fomd` enables failover and, assuming that the failover mechanism has not been deactivated by the operator, does a complete re-sync of the log files and checkpointing data over to the spare SC.

2. The **spare** `fomd` detects the I2 network is healthy.

   The spare `fomd` disables failover and clears out the checkpoint data on the local disk.

Reboot and Recovery

The following lists the events, in order, that occur during a reboot and recovery. A reboot and recovery scenario happens in the following two cases.

Main SC Receives a Master Reset or Its UltraSPARC Receives a Reset

1. Assume SSCPOST passed without any problems. If SSCPOST failed and OS cannot be booted, the **main** is inoperable.
2. Assume all SSC Solaris drivers attached without any problems. If the SBBC driver fails to attach, see “Fault on Main SC (Spare Takes Over Main Role)” on page 140, if any other drivers fail to attach, see “Failover on Main SC (Main Controlled Failover)” on page 139.
3. The **main** `fomd` is started.
4. If the `fomd` determines that the remote SC has already assumed the main role, then see Number 5 in “Spare SC Receives a Master Reset or Its UltraSPARC Receives a Reset”. Otherwise proceed to Number 5 in this list.
5. The `fomd` configures the logical main IP address and starts up the rest of the SMS software.
6. SMS daemons start in recovery mode if necessary.
7. **Main** `fomd` starts generating heartbeat interrupts.
8. At this point, the **main** SC is fully recovered.

Spare SC Receives a Master Reset or Its UltraSPARC Receives a Reset

1. Assume SSCPOST passed without any problems. If SSCPOST failed and OS cannot be booted, the **spare** is inoperable.
2. Assume all SSC Solaris drivers attached without any problems. If the SBBC driver fails to attach, or any other drivers fail to attach, the **spare** SC is deemed inoperable.
3. The `fomd` is started.
4. The `fomd` determines that the SC is the preferred spare and assumes **spare** role.
5. The `fomd` starts checking for the presence of heartbeat interrupts from the remote (initially presumed to be **main**) SC. If after a configurable amount of time no heartbeat interrupts are detected, then the failover mechanism state is checked. If enabled and activated, `fomd` initiates a take over. Now refer to Number 5 of “Main SC Receives a Master Reset or Its UltraSPARC Receives a Reset”. Otherwise, `fomd` continues monitoring for the presence of heartbeat interrupts and the state of the failover mechanism.
6. The `fomd` starts periodically checking the hardware/software and network interfaces.
7. The `fomd` configures the local **main** SC IP address.

8. At this point, the **spare** SC is fully recovered.

Client Failover Recovery

The following lists the events that occur during a client failover recovery. A recovery scenario happens in the following two cases.

Fault on Main SC— Recovering From the Spare SC

Clients with any operations in progress are manually recovered by checkpointing data unless they are non-recurring.

Fault on Main SC (With I2 Network Down)— Recovering From the Spare SC

Since the I2 network is down, all checkpointing data are removed. Clients cannot perform any recovery.

Reboot Main SC (With Spare SC Down)

Same as“Fault on Main SC— Recovering From the Spare SC”

Reboot of Spare SC

No recovery necessary.

Security

All failover specific network traffic (such as health status request/response messages and file propagation packets) are sent *only* over the interconnect network that exists between the two SCs.

CHAPTER 9

Domain Events

Event monitoring periodically checks the domain and hardware status to detect conditions that need to be acted upon. The action taken is determined by the condition and can involve reporting the condition or initiating automated procedures to deal with it. This chapter describes the events that are detected by monitoring and the requirements with respect to actions taken in response to detected events.

Message Logging

SMS logs all significant events, such as those that require taking actions other than logging or updating user monitoring displays in message files on the SC. Included in the log is information to support subsequent servicing of the hardware or software.

SMS writes log messages for significant hardware events to the platform log file located in `/var/opt/SUNWSMS/adm/platform/messages`.

The actions taken in response to events that crash domain software systems include automatic system recovery (ASR) reboots of all affected domain(s), provided that the domain hardware (or a bootable subset thereof) meets the requirements for safe and correct operation.

SMS logs all significant actions other than logging or updating user monitoring displays taken in response to an event. Log messages for significant domain software events and their response actions are written to the message log file for the affected domain located in `/var/opt/SUNWSMS/adm/`*`domain_id`*`/messages`.

SMS writes log messages to `/var/opt/SUNWSMS/adm/`*`domain_id`*`/messages` for significant hardware events that can visibly affect one or more domains of the affected domain(s).

SMS also logs domain console, syslog, post and dump information and as well as manages `sms_core` files.

Log File Maintenance

SMS software maintains SC-resident copies of logs of all server information that it logs. Use the `showlogs`(1M) command to access log information.

The platform message log file can be accessed only by administrators for the platform using:

```
sc0:sms-user:> showlogs
```

SMS log information relevant to a configured domain can be accessed only by administrators for that domain. SMS maintains separate log files for each domain using:

```
sc0:sms-user:> showlogs -d domain_id|domain_tag
```

SMS maintains copies of domain `syslog` files on the SC in `/var/opt/SUNWSMS/adm/`*domain_id*`/syslog`. `syslog` information can be accessed only by administrators for that domain using:

```
sc0:sms-user:> showlogs -d domain_id|domain_tag -s
```

Solaris console output logs are maintained to provide valuable insight into what happened before a domain crashed. Console output is available on the SC for a crashed domain in `/var/opt/SUNWSMS/adm/`*domain_id*`/console`. `console` information can be accessed only by administrators for that domain using:

```
sc0:sms-user:> showlogs -d domain_id|domain_tag -c
```

XIR state dumps, generated by the reset command, can be displayed using `showxirstate`. For more information refer to the `showxirstate` man page.

Domain post logs are for service diagnostic purposes and are not displayed by `showlogs` or any SMS CLI.

The `/var/tmp/sms_core.`*daemon* files are binaries and not viewable.

The availability of various log files on the SC supports analysis and correction of problems that prevent a domain or domains from booting. For more information refer to the `showlogs` man page.

Note – Panic dumps for panicked domains are available in the `/var/crash` logs on the domain and not on the SC.

The following table lists the SMS log information types, and their descriptions.

TABLE 9-1 SMS Log Type Information

Type	Description
Firmware Versioning	Unsuitable configuration of firmware version at firmware invocation is automatically corrected and logged.
Power On Self Test	LED fault; Platform and domain messages detailing why a fault LED was illuminated.
Power Control	*A*ll power operations are logged.
Power Control	Power operations that violate hardware requirements or hardware recommended procedures.
Power Control	Use of override to forcibly complete a power operation.
Domain Console	Automatic logging of console output to a standard file.
Hardware Configuration	Part numbers are used to identify board type in message logs.
Event Monitoring and Actions	All significant environmental events (those that require taking action).
Event Monitoring and Actions	All significant actions taken in response to environmental events.
Domain Event Monitoring and Actions	All significant domain software events and their response actions.
Event Monitoring and Actions	Significant hardware events written to the platform log.
Domain Event Monitoring and Actions	Significant hardware events that visibly affect one or more domains are written to the domain(s) log.
Domain Boot Initiation	Initiation of each boot and the passage through each significant stage of booting a domain is written to the domain log.
Domain Boot Failure	Boot failures are logged to the domain log.
Domain Boot Failures	All ASR recovery attempts are logged to the domain log.

TABLE 9-1 SMS Log Type Information *(Continued)*

Type	Description
Domain Panic	Domain panics are logged to the domain log.
Domain Panic	All ASR recovery attempts are logged to the domain log.
Domain Panic Hang	Each occurrence of a domain hang and its accompanying information is logged to the domain log.
Domain Panic	All ASR recovery attempts after a domain panic and hang are logged to the domain log.
Repeated Domain Panic	All ASR recovery attempts after repeated domain panics are logged to the domain message log.
Solaris OS Hang Events	All operating system hang events are logged to the domain message log.
Solaris OS Hang Events	All OS hang events result in a domain panic in order to obtain a core image for analysis of the Solaris hang. This information and subsequent recovery action is logged to the domain message log.
Solaris OS Hang Events	SMS monitors for the inability of the domain software to satisfy the request to panic. Upon determining noncompliance with the panic request, SMS aborts the domain and initiates an ASR reboot. All subsequent recovery action is logged to the domain message file.
Hot-Plug Events	All HPU insertion events of system boards to a domain are logged in the domain message log.
Hot-Unplug Events	All HPU removals are logged to the platform message log.
Hot-Unplug Events	All HPU removals from a domain are logged to the domain message log.
POST-initiated Configuration Events	All POST-initiated hardware configuration changes are logged in `/var/opt/SUNWSMS/adm/`*`domain_id`*`/post`.
Environmental Events	All sensor measurements outside of acceptable operational limits are logged as environmental events to the platform log file.
Environmental Events	All environmental events that affect one or more domains are logged to the domain message log.
Environmental Events	Significant actions taken in response to environmental events are logged to the platform message log.
Environmental Events	Significant actions taken in response to environmental events within a domain are logged to the domain message log.
Hardware Error Events	Hardware error and related information is logged to the platform message log.
Hardware Error Events	Hardware error and related information within a domain is logged to the domain message file.

TABLE 9-1 SMS Log Type Information *(Continued)*

Type	Description
Hardware Error Events	Log entries about hardware error for which data was collected include the name of the data file(s).
Hardware Error Events	All significant actions taken in response to hardware error events are logged to the platform message log.
Hardware Error Events	All significant actions taken in response to hardware error events affecting a domain(s) are logged to the domain(s) message log.
SC Failure Events	All SC hardware failure and related information is logged to the platform message log.
SC Failure Events	The occurrence of an SC failover event is logged to the platform message log.

Log File Management

SMS manages the log files, as necessary, to keep the SC disk utilization within acceptable limits.

The message log daemon (`mld`) monitors message log size, file count *per directory*, and age every 10 minutes. `mld` executes the first limit to be reached.

TABLE 9-2 MLD Default Settings

	File Size (in Kb)	File Count	Days to Keep
platform messages	2500	10	0
domain messages	2500	10	0
domain console	2500	10	0
domain syslog	2500	10	0
domain post	20000*	1000	0
domain dump	20000*	1000	0
sms_core.*daemon*	50000	2	0

* total per directory not file

Assuming 20 directories, the defaults represent approximately 4Gbytes of stored logs.

Caution – The parameters show in TABLE 9-2 are stored in `/etc/opt/SUNWSMS/config/mld_tuning`. `mld` must be stopped and restarted for any changes to take effect. Only an administrator experienced with system disk utilization should edit this file. Improperly changing the parameters in this file could flood the disk and *hang* or *crash* the SC.

- When a log message file reaches the size limit `mld` does the following:

  Starting with the oldest message file x.X, it moves that file to x.X+1, except when the oldest message file is `message.9` or core file is `sms_core.daemon.1`, then it starts with x.X-1.

  For example, `messages` becomes `messages.0`, `message.0` becomes `messages.1` and so on up to `messages.9`. When `messages` reaches 2.5MB then `messages.9` is deleted and all files are bumped up by one and a new empty `messages` file is created.

- When a log file reaches the file count limit `mld` does the following:

  When `messages` or `sms_core.`*daemon* reaches its count limit, then the oldest message or core file is deleted.

- When a log file reaches the age limit `mld` does the following:
- When any message file reaches *x* days, it is deleted.

Note – By default, the age limit (`*_log_keep_days`) is set to zero and not used.

- When a `post`*date.time.sec*`.log` or a *dump_name*`.`*date.time.sec* file reaches the file size, count or age limit, `mld` deletes the oldest file in the directory.

Note – Post files are provided for service diagnostic purposes and not intended for display.

For more information, refer to the `mld` and `showlogs` man pages, and see “Message Logging Daemon” on page 44.

Domain Reboot Events

SMS monitors domain software status (see “Software Status” on page 119) to detect domain reboot events.

Domain Reboot Initiation

Since the domain software is incapable of rebooting itself, SMS software controls the initial sequence for all domain reboots. In consequence, SMS is always aware of domain reboot initiation events.

SMS software logs the initiation of each reboot and the passage through each significant stage of booting a domain to the domain-specific log file.

Domain Boot Failure

SMS software detects all domain reboot failures.

Upon detecting a domain reboot failure, SMS logs the reboot failure event to the domain-specific message log.

SC resident per-domain log files are available for failure analysis. In addition to the reboot failure logs, SMS can maintain duplicates of important domain-resident logs and transcripts of domain console output on the SC as described in “Log File Maintenance” on page 146.

Domain reboot failures are handled as follows:

- The first attempt to recover a domain from software failure uses a quick reboot procedure. The response to `reboot` or `reset` requests is always a fast bringup procedure.
- The first attempt to recover domain from hardware failure uses the `reboot` procedure. The POST default diagnostic level is used in the `reboot` procedure.
- If the domain recovery fails during the POST run, `dsmd` will retry POST at the default diagnostic level for up to four consecutive domain recovery failures after the first recovery attempt fails.
- If the domain recovery fails during the IOSRAM layout, OpenBoot PROM download and jump, OpenBoot PROM run, or Solaris software boot, `dsmd` reruns POST at the default diagnostic level. `dsmd` retries domain recovery domain at the default level for up to four attempts after the first recovery attempt fails. (All in all, `dsmd` will try domain recovery attempts at most five times).

- Once the system has been recovered and Solaris software has been booted, any domain failures within four hours is treated as repeated domain failure and is recovered by running POST at the default level.
- If there are no domain failures within four hours of Solaris software running, then the domain is considered successfully recovered and healthy.
- A subsequent domain hardware failure is handled by the `reboot` procedure.
- A subsequent domain software failure is handled by quick reboot procedure, and the `reboot` or `reset` request is handled by the fast bringup procedure.

SMS tries all ASR methods at its disposal to boot a domain that has failed booting. All recovery attempts are logged in the domain-specific message log.

Domain Panic Events

When a domain panics, it informs `dsmd` so that a recovery reboot can be initiated. The panic is reported as a domain software status change (see “Software Status” on page 119.

Domain Panic

`dsmd` is informed when the Solaris software on a domain panics.

Upon detecting a domain panic, `dsmd` logs the panic event including information, to the domain-specific message log.

SC resident per-domain log files are available to assist in domain panic analysis. In addition to the panic logs, SMS can maintain duplicates of important domain-resident logs and transcripts of domain console output on the SC as described in “Log File Maintenance” on page 146.

In general, after an initial panic where there has been no prior indication of hardware errors, SMS requests that a fast reboot be tried to bring up the domain. For more information, see “Fast Boot” on page 95.

After a panic event, `dsmd` tries the ASR reboot on the panicked domain. This recovery action is logged in the domain-specific message log.

Domain Panic Hang

The Solaris panic dump logic has been redesigned to minimize the possibility of hangs at panic time. In a panic situation, Solaris software may operate differently either because normal functions are shutdown or because it is disabled by the panic. An ASR reboot of a panicked Solaris domain is eventually started, even if the panicked domain hangs before it can request a reboot.

Since the normal heartbeat monitoring (see “Solaris Software Hang Events” on page 154) of a panicked domain may not be appropriate or sufficient to detect situations where a panicked Solaris domain will not proceed to request an ASR reboot, `dsmd` takes special measures as necessary to detect a domain panic hang event.

Upon detecting a panic hang event, `dsmd` logs each occurrence including information, to the domain-specific message log.

Upon detection of a domain panic hang (if any), SMS aborts the domain panic (see “Domain Abort / Reset” on page 96) and initiates an ASR reboot of the domain. `dsmd` logs these recovery actions in the domain-specific message log.

SC resident log files are available to assist in panic hang analysis. In addition to the panic hang event logs, `dsmd` maintains duplicates of important domain-resident logs and transcripts of domain console output on the SC as described in “Log File Maintenance” on page 146.

Repeated Domain Panic

If a second domain panic is detected shortly after recovering from a panic event, `dsmd` classifies the domain panic as a repeated domain panic event.

In addition to the standard logging actions that occur for any panic, the following action is taken when attempting to reboot after the repeated domain panic event.

With each successive repeated domain panic event, SMS attempts a full-test-level boot against the next untried administrator-specified degraded configuration (see “Degraded Configuration Preferences” on page 88).

After all degraded configurations have been tried, successive repeated domain panic events will continue full-test-level boots using the last specified degraded configuration.

Upon determining that a repeated domain panic event has occurred, `dsmd` tries the ASR method at its disposal to boot a stable domain software environment. `dsmd` logs all recovery attempts in the domain-specific message log.

Solaris Software Hang Events

dsmd monitors the Solaris heartbeat described in “Solaris Software Heartbeat” on page 123 in each domain while Solaris software is running (see Section “Software Status” on page 119). When the heartbeat indicator is not updated for a period of time, a Solaris software hang event occurs.

dsmd detects Solaris software hangs.

Upon detecting a Solaris hang, dsmd logs the hang event including information, to the domain-specific message log.

Upon detecting a Solaris hang, dsmd requests the domain software to panic in order to obtain a core image for analysis of the Solaris hang (“Domain Abort/Reset” on page 96). SMS logs this recovery action in the domain-specific message log.

dsmd monitors the inability of the domain software to satisfy the request to panic. Upon determining noncompliance with the panic request, dsmd aborts the domain (see “Domain Abort/Reset” on page 96) and initiates an ASR reboot. dsmd logs these recovery actions in the domain-specific message log.

Although the core image taken as a result of the panic will only be available for analysis from the domain, SC resident log files are available to assist in domain hang analysis. In addition to the Solaris hang event logs, dsmd can maintain duplicates of important domain-resident logs and transcripts of domain console output on the SC.

Hardware Configuration Events

Changes to the hardware configuration status are considered hardware configuration events. esmd detects the following hardware configuration events on the Sun Fire 15K system.

Hot-Plug Events

The insertion of a hot-pluggable unit (HPU) is a hot-plug event. The following actions take place:

- SMS detects HPU insertion events and logs each event and additional information to a platform message log file.

- If the inserted HPU is a system board in the logical configuration for a domain, SMS also logs its arrival in the domain's message log file.

Hot-Unplug Events

The removal of a hot-pluggable unit (HPU) is a hot-unplug event. The following actions take place:

- Upon occurrence of a hot-unplug event, SMS makes a log entry recording the removal of the HPU to the platform message log file.
- A hot unplug event that detects the removal of a system board from a logical domain configuration logs it to that domain's message log file.

Post-Initiated Configuration Events

POST can run against different server components at different times due to domain-related events such as reboots and dynamic reconfigurations. As described in "Hardware Configuration" on page 123, SMS includes status from POST and identifying failed-test components. Consequently, changes in POST status of a component are considered to be hardware configuration events. SMS logs POST-initiated hardware configuration changes to the platform message log.

Environmental Events

In general, environmental events are detected when hardware status measurements exceed normal operational limits. Acceptable operational limits depend upon the hardware and the server configuration.

esmd verifies that measurements returned by each sensor are within acceptable operational limits. esmd logs all sensor measurements outside of acceptable operational limits as environmental events to the platform log file.

esmd also logs significant actions taken in response to an environmental event (such as those beyond logging information or updating user displays) to the platform log file.

esmd logs significant environmental event response actions that affect one or more domain(s) to the log file(s) of the affected domain(s).

esmd handles environmental events by removing from operation the hardware that has experienced the event (and any other hardware dependent upon the disabled component). Hardware can be left in service, however, if continued operation of the hardware will not harm the hardware or cause hardware functional errors.

The options for handling environmental events are dependent upon the characteristics of the event. All events have a time frame during which the event must be handled. Some events kill the domain software; some do not. Event response actions are such that esmd responds within the event time frame.

There are a number of responses esmd can make to environmental events, such as increasing fan speeds. In response to a detected environmental event, which requires a power off, esmd undertakes one of the following corrective actions:

- esmd uses immediate power off if there is no other option that meets the time constraints.
- If the environment event does not require immediate power off and the component is a MaxCPU board, esmd will attempt to DR the endangered board out of the running domain and power it off.
- If the environment event does not require immediate power off and the component is a centerplane support board (CSB), esmd will attempt to reconfigure the bus traffic to use only the other CSB and power the component off.
- Where possible, if the environment event does not require immediate power off and the component is any type of board other than a MaxCPU or CSB, esmd notifies dsmd of the environment condition and dsmd sends an "orderly shutdown" request to the domain. The domain flushes uncommitted memory buffers to physical storage.

If the software is still running and a viable domain configuration remains after the affected hardware is removed, a remote DR operation to remove the hardware from the domain allows it to continue running in degraded mode.

If either of the last two options takes longer than the allotted time for the given environmental condition, `esmd` will immediately power off the component regardless of the state of the domain software.

SMS illuminates the fault indicator LED on any hot-pluggable unit that can be identified as the cause of an environmental event.

So long as the environmental event response actions do not include shutdown of the system controller(s), all domain(s) whose software operations were terminated by an environmental event or the ensuing response actions are subject to ASR reboot as soon as possible.

ASR reboot begins immediately if there is a bootable set of hardware that can be operated in accordance with constraints imposed by the Sun Fire 15K system to assure safe and correct operation.

Note – Loss of system controller operation (for example, by the requirement to power both SCs down) eliminates all possibility of Sun Fire 15K platform self-recovery actions being taken. In this situation, some recovery actions can require human intervention, so although an external monitoring agent may not be able to recover the Sun Fire 15K platform operation, that monitoring agent may serve an important role in notifying an administrator about the Sun Fire 15K platform shutdown.

The following provides a little more detail about each type of environmental event that can occur on the Sun Fire 15K system.

Over-Temperature Events

`esmd` monitors temperature measurements from Sun Fire 15K hardware for values that are too high. There is a critical temperature threshold that, if exceeded, is handled as quickly as possible by powering off the affected hardware. High, but not critical, temperatures are handled by attempting slower recovery actions.

Power Failure Events

There is very little opportunity to do anything when a full power failure occurs. The entire platform, domains as well as SCs, are shut off when the plug is pulled without the benefit of a graceful shut down. The ultimate recovery action occurs when power is restored (see “Power-On Self-Test (POST)” on page 97).

Out-of-Range Voltage Events

Sun Fire 15K power voltages are monitored to detect out-of-range events. The handling of out-of-range voltages follows the general principles outlined at the beginning of “Environmental Events” on page 156.

Under-Power Events

In addition to checking for adequate power before powering on any boards, as mentioned in “Power Control” on page 103, the failure of a power supply could leave the server inadequately powered. The system is equipped with power supply redundancy in the event of failure. `esmd` does not take any action (other than logging) in response to a bulk power supply hardware failure. The handling of under power events follow the general principles outlined at the beginning of “Environmental Events” on page 156

Fan Failure Events

`esmd` monitors fans for continuing operation. Should a fan fail, a fan failure event occurs. The handling of fan-failures will follow the general principles outlined at the beginning of “Environmental Events”.

Hardware Error Events

As described in “Hardware Error Status” on page 125, the occurrence of Sun Fire 15K hardware errors is recognized at the SC by more than one mechanism. Of the errors that are directly visible to the SC, some are reported directly by PCI interrupt to the UltraSPARC IIi processor on the SC, and others are detected only through monitoring of the Sun Fire 15K hardware registers.

There are other hardware errors that are detected by the processors running in a domain. Domain software running in the domain detects the occurrence of those errors in the domain, which then reports the error to the SC. Like the mechanism by which the SC becomes aware of the occurrence of a hardware error, the error state retained by the hardware after a hardware error is dependent upon the specific error.

`dsmd` implements the mechanisms necessary to detect all SC-visible hardware errors.

`dsmd` implements domain software interfaces to accept reports of domain-detected hardware errors.

`dsmd` collects hardware error data and clears the error state.

`dsmd` logs the hardware error and related information as required, to the platform message log.

`dsmd` logs the hardware error to the domain message log file for all affected domain(s).

Data collected in response to a hardware error that is not suitable for inclusion in a log file may be saved in uniquely named file(s) in `/var/opt/SUNWSMS/adm/`*`domain_id`*`/dump` on the SC.

SMS illuminates the fault indicator LED on any hot-pluggable unit that can be identified as the cause of a hardware error.

The actions taken in response to hardware errors (other than collecting and logging information as described above) are twofold. First, it may be possible to eliminate the further occurrence of certain types of hardware errors by eliminating from use the hardware identified to be at fault.

Second, all domains that crashed either as a result of a hardware error or were shut down as a consequence of the first type of action are subject to ASR reboot actions.

Note – Even in the absence of actions to remove from use hardware identified to be at fault, the ASR reboot actions are subject to full POST verification. POST will eliminate any hardware components that fail testing from the hardware configuration.

In response to each detected hardware error and each domain-software-reported hardware error, dsmd undertakes corrective actions.

ASR reboot with full POST verification will be initiated for each domain brought down by a hardware error or subsequent actions taken in response to that error.

Note – Problems with the ASR reboot of a domain after a hardware error are detected as domain boot failure events and subject to the recovery actions described in “Domain Boot Failure” on page 151.

---

dsmd logs all significant actions, such as those beyond logging information or updating user displays taken in response to a hardware error in the platform log file. When a hardware error affects one or more domains, dsmd logs the significant response actions in the message log files of the affected domain(s).

The following sections summarize the types of hardware errors expected to be detected / handled on the Sun Fire 15K system.

Domain Stop Events

Domain stops are uncorrectable hardware errors that immediately terminate the affected domain(s). Hardware state dumps are taken before dsmd initiates an ASR reboot of the affected domain(s). These files are located in: /var/opt/SUNWSMS/adm/*domain_id*/dump. dsmd logs the event in the domain log file.

CPU-Detected Events

A RED_state or Watchdog reset traps to low-level domain software (OpenBoot PROM or kadb), which reports the error and requests initiation of ASR reboot of the domain.

An XIR signal (reset -x) also traps to low-level domain software (OpenBoot PROM or kadb), which retains control of the software. The domain must be rebooted manually.

Record Stop Events

Correctable data transmission errors (for example, parity errors) can stop the normal transaction history recording feature of Sun Fire 15K ASICs. SMS reports a transmission error as a record stop. SMS dumps the transaction history buffers of the Sun Fire 15K ASICs and re-enables transaction history recording when a record stop is handled. `dsmd` records record stops in the domain log file.

Other ASIC Failure Events

ASIC-detected hardware failures other than domain stop or record stop include console bus errors which may or may not impact a domain. The hardware itself will not abort any domain but the domain software may not survive the impact of the hardware failure and may panic or hang. `dsmd` logs the event in the domain log file.

SC Failure Events

SMS monitors the main SC hardware and running software status as well as the hardware and running software of the spare SC, if present. In a high-availability SC configuration, SMS handles failures of the hardware or software on the main SC or failures detected in the hardware control paths (for example, console bus, or internal network connections) to the main SC by an automatic SC failover process. This cedes main responsibilities to the spare SC and leaves the former main SC as a (possibly crippled) spare.

SMS monitors the hardware of the main and spare SCs for failures.

SMS logs the hardware failure and related information to the platform message log.

SMS illuminates the fault indicator LED on a system controller with an identified hardware failure.

For more information, see “SC Failover” on page 127.

CHAPTER 10

SMS Utilities

This section discusses the SMS backup, configuration, restore and version utilities. For information and examples of these utilities, refer to the *System Management Services (SMS) 1.2 Reference Manual* and online man pages.

SMS Backup Utility

`smsbackup` creates a `cpio`(1) archive of files that maintain the operational environment of SMS.

Note – This utility runs on the SC and does not replace the need for routine and timely backups of SC and domain operating systems; and domain application data.

Whenever changes are made to the SMS environment, for example, by adding boards to or removing boards from a domain, you must run `smsbackup` again in order to maintain a current backup file for the system controller.

The name of the backup file is `smsbackup.X.X.cpio` - where X.X represents the active version from which the backup was taken.

`smsbackup` saves all configuration, platform configuration database, SMS, and log files. In other words, SMS saves everything needed to return SMS to the working state it was in at the time the backup was made.

Backups are *not* performed automatically. Whenever changes are made to the SMS environment, a backup should be performed. This process can be automated by making it part of `root cron` job run at periodic intervals depending on your site requirements.

The backup log file resides in `/var/sadm/system/logs/smsbackup`. You must specify the target location when running `smsbackup`.

Note – The target location must be a valid UFS file system directory. You cannot perform `smsbackup` to a tmp file system directory.

Whenever you run `smsbackup`, you will receive confirmation that it succeeded or be notified that it failed.

You must have superuser privileges to run `smsbackup`. For more information and examples, refer to the `smsbackup` man page.

Restore SMS backup files using the `smsrestore`(1M) command.

SMS Restore Utility

`smsrestore` restores the operational environment of the SMS from a backup file created by `smsbackup`(1M). You can use `smsrestore` to restore the SMS environment after the SMS software has been installed on a new disk or after hardware replacement or addition. Failover should be disabled and SMS stopped before `smsrestore` is performed. Refer to “Stopping and Starting SMS” on page 52 of the *System Management Services (SMS) 1.2 Installation Guide and Release Notes*.

If any errors occur, `smsrestore` writes error messages to `/var/sadm/system/logs/smsrestore`.

Note – This utility runs on the SC and does not restore SC operating system, domain operating system or domain application data.

`smsrestore` cannot restore what you have not backed up. Whenever changes are made to the SMS environment, for example, by shutting down a domain, you must run `smsbackup` in order to maintain a current backup file for the system controller.

You must have superuser privileges to run `smsrestore`. For more information and examples, refer to the `smsrestore` man page.

SMS Version Utility

`smsversion`(1M) administers adjacent, co-resident installations of SMS. Adjacent meaning you can use `smsversion` to switch between version SMS1.1 and SMS1.2 (assuming both are installed), but you cannot use `smsversion` to switch between SMS1.1 and a version of SMS following 1.2.

When you upgrade to the next sequential release of SMS (for example, 1.1 to 1.2), SMS must be stopped before running `smsversion`. Refer to “Stopping and Starting SMS” on page 52 of the *System Management Services (SMS) 1.2 Installation Guide and Release Notes*. `smsversion` backs up important system and domain information and switches to the target SMS version.

Note – This utility runs on the SC and does not replace the need for routine and timely backups of the SC and domain operating system; and domain application data.

You can switch back to the previous sequential SMS version (for example, 1.2 to 1.1) at a later time. `smsversion` permits unlimited 2-way SMS version-switching between sequential co-resident installations.

Switching between non-sequential releases is *not* supported.

Without options, `smsversion` displays the active version and exits when only one version of SMS is installed.

If any errors occur, `smsversion` writes error messages to `/var/sadm/system/logs/smsversion`.

You must have superuser privileges to run `smsversion`. For more information and examples, refer to the `smsversion` man page.

▼ Upgrade Flow

Switch between two adjacent, co-resident installations of SMS as follows:

On the main SC:

Make certain your configuration is stable and backed up using `smsbackup`.

Being stable means the following commands should *not* be running: `smsconfig`, `poweron`, `poweroff`, `setkeyswitch`, `cfgadm`, `rcfgadm`, `addtag`, `deletetag`, `addboard`, `moveboard`, `deleteboard`, `setbus`, `setdefaults`, `setobpparams`, `setupplatform`, `enablecomponent` or `disablecomponent`.

Deactivate failover using `setfailover off`.

On the spare SC:

Run `/etc/init.d/sms stop`

Run `smsversion`.

Run `smsrestore`.

If necessary, run `smsconfig -m` and reboot.

Only run `smsconfig -m` if you changed your network configuration using `smsconfig -m` *after* creating the `smsbackup` you just restored.

On the main SC:

Stop SMS using `/etc/init.d/sms stop`

On the spare SC:

If `smsconfig -m` was run, reboot otherwise run `/etc/init.d/sms start`

When the SC comes up it will become the main SC.

On the former main SC:

Repeat steps 4-6 and 8.

On the new main SC:

Activate failover using `setfailover on`.

For more information refer to the *System Management Services (SMS) 1.2 Installation Guide and Release Notes*.

SMS Configuration Utility

`smsconfig` configures the MAN networks, modifies the hostname and IP address settings used by the MAN daemon, `mand`(1M) and administers domain directories access control lists (ACLs).

UNIX Groups

`smsconfig` configures the UNIX groups used by SMS to describe user privileges. SMS uses a default set of UNIX groups installed locally on each SC. `smsconfig` allows you to customize those groups using the `-g` option. `smsconfig` allows you to add users to groups using the `-a` option and remove users from groups using the `-r` option.

For information and examples on adding, removing, and listing authorized users, refer to the *System Management Services (SMS) 1.2 Installation Guide and Release Notes* and `smsconfig`(1M) man page.

Access Control List (ACL)

Traditional UNIX file protection provides read, write and execute permissions for the three user classes: file owner, file group, and other. In order to provide protection and isolation of domain information, access to each domain's data is denied to all unauthorized users. SMS daemons, however, are considered authorized users and will have full access to the domain file systems. For example:

- `sms-esmd`—needs to read the blacklist files in each domain directory `$SMSETC/config/[A-R]`
- `sms-osd`—needs to read from and write to the `bootparamdata` file in each domain: `$SMSVAR/data/[A-R]`
- `sms-dsmd`—need to write to hpost logs for every domain `$SMSVAR/adm/[A-R]/post`

`smsconfig` sets the ACL entries associated with the domain directories so that the domain administrator has full access to the domain. A plus sign (+) to the right of the mode field indicates the directories that have ACL defined.

```
domain_id:sms-user:> ls -al
total 6
drwxrwxrwx   2 root     bin          512 May 10 12:29 .
drwxrwxr-x  23 root     bin         1024 May 10 12:29 ..
-rw-rw-r--+ 1 root     bin          312 May  4 16:15 blacklist
```

To add a user account to the ACL, the user must already belong to a valid SMS group as described in the *System Management Services (SMS) 1.1 Installation Guide and Release Notes.*

Note – UFS file system attributes, such as the ACL, are supported in UFS file systems only. If you restore or copy directories with ACL entries into the `/tmp` directory, all ACL entries will be lost. Use the `/var/tmp` directory for temporary storage of UFS files and directories.

Network Configuration

For each network, `smsconfig` can singularly set one or more *interface* designations within that network. By default, `smsconfig` steps through the configuration of all three internal, enterprise networks.

To configure an individual network, append the *net_id* to the command line. Management networks *net_id*s are designated `I1`, `I2`, and `C`.

Configure a single domain within an enterprise network by specifying both the desired domain and its *net_id*. A domain can be excluded from the I1 management network by using the word `NONE` as the *net_id*.

Note – Once you have configured or changed the configuration of the MAN network you *must* reboot the SC in order for the changes to take effect.

You must have superuser privileges to run `smsconfig`. For more information and examples, refer to the *System Management Services (SMS) 1.2 Installation Guide and Release Notes*, `smsconfig` man page and see “Management Network Services” on page 114.

MAN Configuration

`smsconfig -m` does the following:

1. Creates `/etc/hostname.scman[01]`
2. Creates `/etc/hostname.hme0` and / or `/etc/hostname.eri1` according to inputs to the External Network(s) prompts of `smsconfig`
3. Updates `/etc/netmasks` and `/etc/hosts`
4. Sets OBP variable `local-mac-address?`=true (default is false).

For more information on smsconfig refer to the smsconfig(1M) man page and see “Management Network Services” on page 114.

APPENDIX **A**

SMS man Pages

The SMS man pages are in the *SMS 1.2 Reference Manual* portion of your Sun FIre 15K system documentation set as well as online (after you have installed the SMS packages.)

The following is a list of SMS man pages:

- `addboard`(1M) - Assigns, connects, and configures a board to a domain
- `addtag`(1M) - Assigns a domain name (tag) to a domain
- `cancelcmdsync`(1M) - Removes a command synchronization descriptor from the command synchronization list
- `console`(1M) - Access the domain console
- `dca`(1M)- Domain configuration agent
- `deleteboard`(1M) - Unconfigures, disconnects and unassigns a system board from a domain
- `deletetag`(1M) - Removes the domain name (tag) associated with the domain
- `disablecomponent`(1M) - Adds the specified component from the ASR blacklist
- `dsmd`(1M) - Domain status monitoring daemon
- `dxs`(1M) - Domain *X* server
- `enablecomponent`(1M) - Removes the specified component from the ASR blacklist
- `esmd`(1M) - Environmental status monitoring daemon
- `flashupdate`(1M) - Update system board FROMs
- `fomd`(1M) - Failover management daemon
- `frad`(1M) - FRU access daemon
- `help`(1M) - Displays help information for SMS commands
- `hpost`(1M) - Sun Fire 15K power-on self test (POST) control application
- `hwad`(1M) - Hardware access daemon
- `initcmdsync`(1M) - Creates a command synchronization descriptor that identifies the script to be recovered
- `kmd`(1M) - Key management daemon
- `mand`(1M)- Management network daemon

- `mld`(1M) - Message logging daemon
- `moveboard`(1M) - Moves a system board from one domain to another
- `osd`(1M) - OBP server daemon
- `pcd`(1M) - Platform configuration database daemon
- `poweroff`(1M) - Controls power off
- `poweron`(1M) - Controls power on
- `rcfgadm`(1M) - Remote configuration administration
- `reset`(1M) - Sends reset to all ports (CPU or I/O) of a specified domain
- `resetsc`(1M) - Sends reset to the spare SC
- `runcmdsync`(1M) - Prepares a specified script for recovery after a failover
- `savecmdsync`(1M) - Adds a marker that identifies a location in the script from which processing can be resumed after a failover
- `setbus`(1M) - Performs dynamic bus reconfiguration on active expanders in a domain
- `setdatasync`(1M) - Modifies the data propagation list used in data synchronization
- `setdate`(1M) - Sets the date and time for the System Controller or a domain
- `setdefaults`(1M) - Removes all instances of a previously active domain
- `setfailover`(1M) - Modifies the state of the SC failover mechanism
- `setkeyswitch`(1M) - Changes the position of the virtual keyswitch
- `setobpparams`(1M) - Sets up OpenBoot PROM variables
- `setupplatform`(1M) - Sets up the available component list for domains
- `showboards`(1M) - Shows the assignment information and status of the system boards
- `showbus`(1M) - Displays the bus configuration of expanders in active domains
- `showcmdsync`(1M) - Displays the current command synchronization list
- `showcomponent`(1M) - Displays ASR blacklist status for a component
- `showdatasync`(1M) - Displays the status of SMS data synchronization for failover
- `showdate`(1M) - Displays the date and time for the System Controller or a domain
- `showdevices`(1M) - Displays system board devices and resource usage information
- `showenvironment`(1M) - Displays the environmental data
- `showfailover`(1M) - Manages or displays SC failover status
- `showkeyswitch`(1M) - Displays the position of the virtual key switch
- `showlogs`(1M) -Displays message log files
- `showobbparams`(1M) - Displays OBP bringup parameters
- `showplatform`(1M) - Displays the board available component list for domains
- `showxirstate`(1M) - Displays cpu dump information after sending a reset pulse to the processors

- `smsbackup`(1M) - Backs up the SMS environment
- `smsconfig`(1M) - Configures the SMS environment
- `smsconnectsc`(1M) - Access a remote SC console
- `smsrestore`(1M) - Restores the SMS environment
- `smsversion`(1M) - Displays the active version of SMS software
- `ssd`(1M) - SMS startup daemon
- `tmd`(1M) - Task management daemon

APPENDIX **B**

Error Messages

This section discusses user-visible error messages for SMS. The types of errors and the numerical ranges are listed. To view individual errors, you must install the SMSHelp software package (SUNWSMSjh). This section contains installation instructions for SUNWSMSjh, if it has not already been installed. Each error in SMSHelp contains the error ID, the text of the message, the meaning of the message, references for further information in the *System Management Services (SMS) 1.2 Administrator Guide* if applicable, and recovery action to take or suggested steps for further analysis.

Installing smshelp

This section explains how to manually install the SUNWSMSjh package using the standard installation utility, pkgadd.

▼ To Install theSUNWSMSjh Package

1. **Log in to the SC as superuser.**
2. **Load the SUNWSMSjh package on the server:**

```
# pkgadd -d . SUNWSMSjh
```

The software briefly displays copyright, trademark, and license information for each package. Then it displays messages about pkgadd(1M) actions taken to install the package, including a list of the files and directories being installed. Depending on your configuration, the following messages may be displayed:

```
This package contains scripts which will be executed
with superuser permission during the process of installing this
package.

Do you want to continue with the installation of this
package [y,n,?]
```

3. **Type y at each successive prompt to continue.**

   When this portion of the installation is complete, the SUNWSMSjh package has been installed and the superuser prompt is displayed.

4. **Remove the Sun Computer Systems Supplement CD from the CD-ROM drive, if applicable:**

```
# cd /
# eject cdrom
```

5. **Log out as superuser.**

▼ To Start `smshelp`

1. **Log in to the SC as a user with platform or domain group privileges.**
2. **In any terminal window, type:**

```
sc0:sms-user:> smshelp &
```

The smshelp browser will pop up. You can resize the panes if necessary, by placing your cursor to the right of the vertical scroll bar, pressing the left mouse button and dragging the cursor to the right.

3. Choose an error message.

Error messages are recorded in the platform and domain logs.

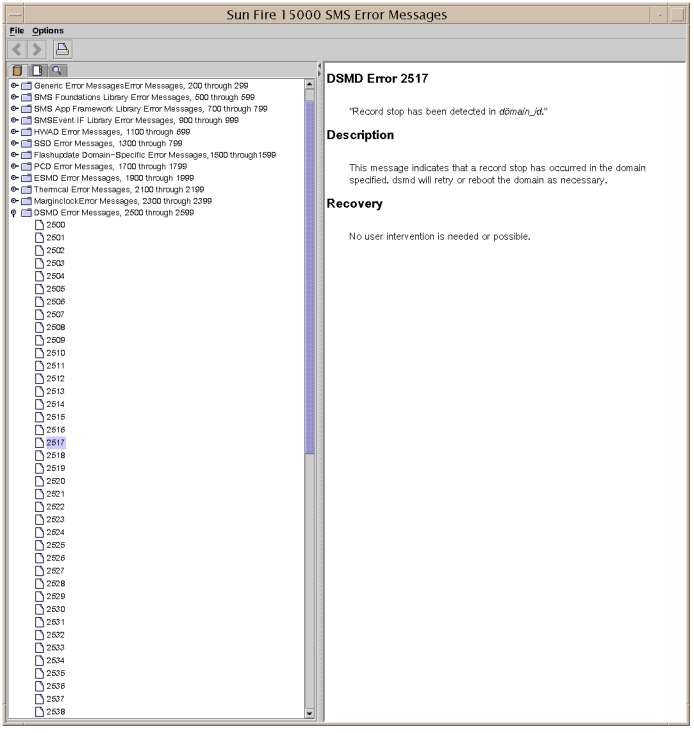
The message format follows the `syslog`(3) convention: (relevant parts of the message are in bold)

```
timestamp  host process_name  [pid]: [message_code
hight_res_timestamp level source_code_file_name
source_code_line_num] message_text
```

For example:

```
Oct  6 18:36:14 2001 xc17-sc0 dsmd[117469]-B(): [2517
16955334989087 WARNING EventHandler.cc 121] Record stop has been
detected in domain B.
```

Using the **message_code**, left click on the message folder containing your error message, in this case DSMD Error Messages, 2500 through 2599. Left click on error 2517.



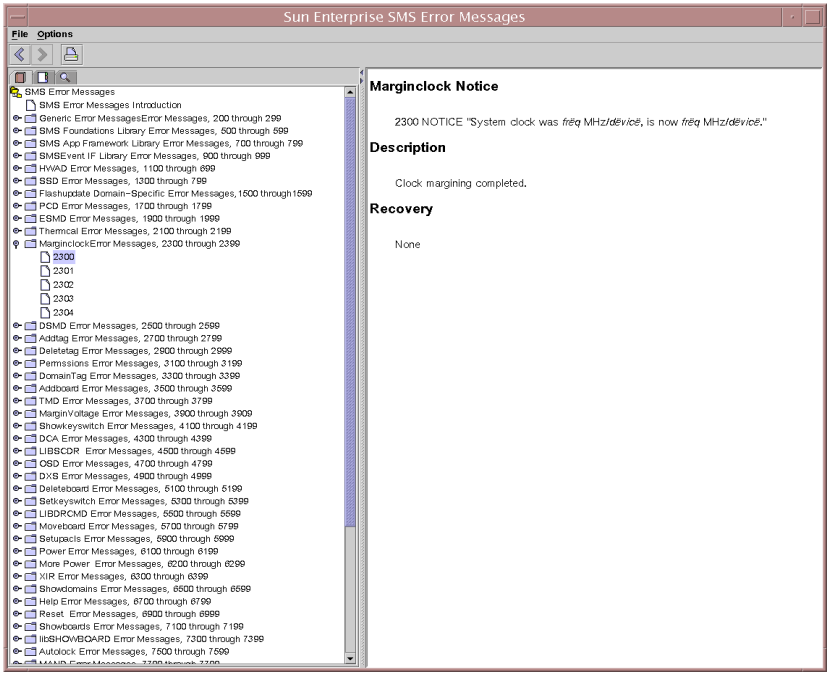
Types of Errors

This section describes the six types of errors reflected in the error messages in `smshelp`.

TABLE B-1 First Level Errors

Error	Description
EMERG	Panic conditions that would normally be broadcast to all users.
ALERT	Conditions that should be corrected immediately, such as a corrupted system database.
CRIT	Warnings about critical conditions, such as hard device failures.
ERROR	All other errors.
WARNING	Warning messages.
NOTICE	Conditions that are not error conditions but may require special handling.

The following is an example of a `smshelp` error message.


Sun Enterprise SMS Error Messages
File Options
SMS Error Messages
SMS Error Messages Introduction
Generic Error MessagesError Messages, 200 through 299
SMS Foundations Library Error Messages, 600 through 699
SMS App Framework Library Error Messages, 700 through 799
SMSEvent IF Library Error Messages, 900 through 999
HWAD Error Messages, 1100 through 699
SSD Error Messages, 1300 through 799
Flashupdate Domain-Specific Error Messages,1500 through1599
PCD Error Messages, 1700 through 1799
ESMD Error Messages, 1900 through 1999
Thermical Error Messages, 2100 through 2199
MarginclockError Messages, 2300 through 2399
2300
2301
2302
2303
2304
DSMD Error Messages, 2500 through 2599
Addtag Error Messages, 2700 through 2799
Deletetag Error Messages, 2900 through 2999
Permssions Error Messages, 3100 through 3199
DomainTag Error Messages, 3300 through 3399
Addboard Error Messages, 3500 through 3599
TMD Error Messages, 3700 through 3799
MarginVoltage Error Messages, 3900 through 3909
Showkeyswitch Error Messages, 4100 through 4199
DCA Error Messages, 4300 through 4399
LIBSCDR Error Messages, 4500 through 4599
OSD Error Messages, 4700 through 4799
DXS Error Messages, 4900 through 4999
Deleteboard Error Messages, 5100 through 5199
Setkeyswitch Error Messages, 5300 through 5399
LIBDRCMD Error Messages, 5500 through 5599
Moveboard Error Messages, 5700 through 5799
Setupacls Error Messages, 5900 through 5999
Power Error Messages, 6100 through 6199
More Power Error Messages, 6200 through 6299
XIR Error Messages, 6300 through 6399
Showdomains Error Messages, 6500 through 6599
Help Error Messages, 6700 through 6799
Reset Error Messages, 6900 through 6999
Showboards Error Messages, 7100 through 7199
libSHOWBOARD Error Messages, 7300 through 7399
Autolock Error Messages, 7500 through 7599
Marginclock Notice
2300 NOTICE "System clock was freq MHz/device, is now freq MHz/device."
Description
Clock margining completed.
Recovery
None

Error Categories

The following table shows the different error categories in SMS. Non-sequential numbering is due to error messages reserved for internal or service use.

TABLE B-2 First Level Errors

Error Numbers	Message Group
0-499	Reserved for DEBUG, INFO and POST messages
500-699	Reserved for SMS Foundation Library messages
700-899	Reserved for SMS Application Framework messages
900-1099	Reserved for SMSEvent IF Library messages
1100-1299	Reserved for HWAD daemon and library messages
1300-1499	Reserved for ssd messages
1500-1699	Reserved for flashupdate messages
1700-1899	Reserved for pcd messages
1900-2099	Reserved for esmd messages
2500-2699	Reserved for dsmd messages
2700-2899	Reserved for addtag messages
2900-3099	Reserved for deletetag messages
3100-3299	Reserved for Permissions messages
3300-3499	Reserved for *domain_tag* messages
3500-3699	Reserved for addboard messages
3700-3899	Reserved for tmd messages
4100-4299	Reserved for showkeyswitch messages
4300-4499	Reserved for dca messages
4500-4699	Reserved for libscdr plugin messages
4700-4899	Reserved for osd messages
4900-5099	Reserved for dxs messages
5100-5299	Reserved for deleteboard messages
5300-5499	Reserved for setkeyswitch messages
5500-5699	Reserved for libdrcmd messages

TABLE B-2 First Level Errors

Error Numbers	Message Group
5700-5899	Reserved for moveboard messages
5900-6099	Reserved for setupplatform messages
6100-6299	Reserved for power command messages
6300-6499	Reserved for xir library messages
6500-6699	Reserved for showplatform messages
6700-6899	Reserved for help messages
6900-7099	Reserved for reset messages
7100-7299	Reserved for showboards messages
7300-7499	Reserved for libshowboards messages
7500-7699	Reserved for autolock messages
7700-7899	Reserved for mand messages
7900-8099	Reserved for showenvironment messages
8100-8299	Reserved for resetsc messages
8300-8499	Reserved for dynamic bus reconfiguration messages
8500-8699	Reserved for fomd messages
8700-8899	Reserved for kmd messages
8900-9099	Reserved for setdefaults messages
9100-9299	Reserved for mld messages
9300-9499	Reserved for showdevices messages
9500-9699	Reserved for showxirstate messages
9900-10000	Reserved for frad messages
10100-10299	Reserved for fruevent messages
10300-10499	Reserved for smsconnectsc messages
10700-10899	Reserved for EFE messages
50000-50099	Reserved for SMS generic messages

Glossary

A

ACL See *access control list (ACL).*

access control list (ACL) Access control list (ACL) provides greater control over file and folder permissions. ACL enables you to define file or folder permissions for the owner, owner's group, others, and specific users and groups, and default permissions for each of these categories.

active board list List of components that are in use in a domain. The pcd(1M) keeps the state of this list.

ADR See *Automated dynamic reconfiguration (ADR)*

application specific integrated circuit (ASIC) In the Sun Fire systems, any of the large main chips in the design, including the UltraSPARC processor and data buffer chips.

arbitration stop A condition that occurs when one of the Sun Fire 15K system ASICs detects a parity error or equivalent fatal system error. Bus arbitration is frozen, so all bus activity stops.

ASIC See *application specific integrated circuit (ASIC)*

assigned board list List of components that have been assigned to a domain by a domain administrator/configurator privileged user. The pcd(1M) keeps the state of this list.

ASR Automatic System Recovery.

auto-failover The process by which the SMS daemon, `fomd`, automatically switches SC control from the main SC to the spare in the event of hardware or software failure on the main.

Automated dynamic reconfiguration (ADR) The dynamic reconfiguration of system boards accomplished through commands that can be used to automatically `assign/unassign`, `connect/disconnect` and `configure/unconfigure` boards, and obtain board status information. You can run these commands interactively or in shell scripts.

automatic system recovery (ASR) Automatic system recovery consists of those procedures that restore the system to running all properly configured domains after one or more domains have been rendered inactive due to software or hardware failures or due to unacceptable environmental conditions

available component list List of available components which can be assigned to a domain by a domain administrator/configurator privileged user. The `pcd`(1M) keeps the state of this list. `setupplatform`(1M) updates it.

B

BBC Bootbus controller. An ASIC used on the CPU & IO boards (also System Controller boards), that connects the Bootbus to the Prom bus and the Console bus

BBSRAM See *bootbus SRAM (BBSRAM)*.

blacklist A text file that `hpost`(1M) reads when it starts up. The blacklist file specifies the Sun Fire system components that are not to be used or configured into the system. Platform, domain blacklist files can be edited using the `enablecomponent` and `disablecomponent` commands. The ASR blacklist is created and edited by `esmd`.

bootbus A slow-speed, byte-wide bus controlled by the processor port controller ASICs, used for running diagnostics and boot code. UltraSPARC starts running code from bootbus when it exits reset. In the Sun Fire 15K system, the only component on the bootbus is the `BBSRAM`.

bootbus SRAM (BBSRAM) A 256-Kbyte static RAM attached to each processor PC ASIC. Through the PC, it can be accessed for reading and writing from JTAG or the processor. Bootbus SRAM is downloaded at various times with `hpost`(1M) and OpenBoot PROM startup code, and provides shared data between the downloaded code and the SC.

C

cacheable address slice map (CASM) A table in the AXQ that directs cacheable addresses to the correct expander.

CASM See *cacheable address slice map (CASM).*

checkpoint data A copy of the state an SC client is in at a specific execution point that is periodically saved to disk.

CLI Command line interface.

cluster A cooperative collection of interconnected computer systems, each running a separate OS image, utilized as a single, unified computing resource.

community An IP network at a customer site that is physically separate from any other networks.

community name A string identifier that names a particular community. In the context of External Network Monitoring for Sun Fire 15K system, it is used as the interface group name. See *interface group name.*

CMR Coherent Memory Replication.

cmdsync Command synchronization. Commands that work together to control recovery during SC failover. For example, `cancelcmdsync`, `initcmdsync` and `savecmdsync`.

CPU Central processing unit.

D

darb An ASIC on the Sun Fire 15K centerplane that handles data arbitration.

DCU See *domain configuration unit (DCU).*

DHCP Dynamic Host Configuration Protocol.

DIMM See *dual in-line memory module (DIMM).*

dstop See *domain stop.*

disk array A collection of disks within a hardware peripheral. The disk array provides access to each of its housed disks through one or two Fibre Channel modules.

disk array controller A controller that resides on the host system and has one or two Fibre Channel modules.

disk array port A Fibre Channel module that can be connected to a disk array controller that is serviced by a driver pair; for example; `soc/pln` for SSAs.

domain A set of one or more system boards that acts as a separate system capable of booting the OS and running independently of any other domains. A machine environment capable of running its own OS. There are up to eighteen domains available on the Sun Fire 15K system. Domains that share a system are characteristically independent of each other.

domain configuration unit (DCU) Refers to a unit of hardware that can be assigned to a single domain. Domains are configured from DCUs. CPU/Memory, PCI I/O, and hsPCI I/O are DCUs. csb, exb boards and the `SC` are not.

domain_id Domain ID of a domain.

domain_tag Domain name assigned using `addtag` (1M)

domain stop An uncorrectable hardware error that immediately terminates the affected domain.

DR See *dynamic reconfiguration (DR).*

DRAM See *dynamic RAM (DRAM).*

drift file The name of the file used to record the drift (or frequency error) value computed by `xntpd`. Most commonly `ntp.drift`.

DSD Dynamic System Domain. See `Domain`.

dual in-line memory module (DIMM) A small printed circuit card containing memory chips and some support logic.

dynamic reconfiguration (DR) Enables you to logically attach and detach system boards to and from the operating system without causing machine downtime. DR is used in conjunction with hot-swap, which is the process of physically removing or inserting a system board. You can use DR to add a new system board, reinstall a repaired system board, or modify the domain configuration on the Sun Fire system.

dynamic RAM (DRAM) Hardware memory chips that require periodic rewriting to retain their contents. This process is called “refresh”. In the Sun Fire 15K system, DRAM is used only on main memory SIMMs and on the control boards.

E

ECC Error Correction Code.

Ecache See *external cache (Ecache).*

EEPROM Electrically Erasable Programmable Read-Only Memory.

Environmental Monitoring Systems have a large number of sensors that monitor temperature, voltage and current. The SC daemons `esmd` and `dsmd` poll devices in a timely manner and makes the environmental data available. The SC shuts down various components to prevent damage.

Ethernet address A unique number assigned to each Ethernet network adapter. It is a 48-bit number maintained by the IEEE. Hardware vendors obtain blocks of numbers that they can build into their cards. See also, *MAC address*

external cache (Ecache) 8Mbyte synchronous static RAM second-level cache local to each processor module. Used for both code and data. This is a direct-mapped cache.

external network A network that requires a physical cable to connect a node to the network. In the context of the Sun Fire 15K system, it is the set of networks connected to the RJ45 jacks located on the front of each Sun Fire 15K system. See *external network.*

external network interface One of the RJ45 jacks located on the front of each Sun Fire 15K System Controller.

F

Fibre Channel module An optical link connection (OLC) module on a disk array controller that can be connected to a disk array port.

Fireplane Centerplane in the Sun Fire 15K system.

FPROM Flash programmable read-only memory

FRU Field replaceable unit.

G

GDCD See *global domain configuration descriptor (GDCD)*

global domain configuration descriptor (GDCD) The description of the single configuration that `hpost`(1M) chooses. It is part of the structure handed off to OpenBoot PROM.

GUI Graphical user interface.

H

HA High availability.

HASRAM High availability SRAM.

heartbeat interrupt Interruption of the normal Solaris operating environment indicator, readable from the SC. Absence of heartbeat updates for a running Solaris system usually indicates a Solaris hang.

hpost Host POST is the POST code that is executed by the SC. Typically this code is sourced from the SC local disk.

HPCI Hot-pluggable PCI I/O assembly

HPU Hot-Pluggable Unit. A hardware component that can be isolated from a running system such that it can be cleanly removed from the system or added to the system without damaging any hardware or software.

HsPCI See *HPCI.*

I

I1 Network There are 18 network interfaces (NICs) on each SC that are connected in a point-to-point fashion to NICs located on each of the 18 expander IO slots. All of these point-to-point links are collectively called the I1 network.

I2C Inter-IC Bus. This is a two-wire bus that is used throughout various systems to run LEDs, set system clock resources, read thermcal information, and so on.

I2 Network There is an internal network between the two system controllers consisting of two NICs per system controller. This network is called the I2 network. It is not a private network and is entirely separate from the I1 network

IDnet Inter-Domain Network.

IDPROM Identification PROM. Contains information specific to the Sun Fire 15K internal machine, such as machine type, manufacturing date, ethernet address, serial number and host ID.

interface group A group of network interfaces that attach to the same community.

interface group name A string identifier that names a particular interface group. In the context of External Network Monitoring for Sun Fire 15K system, it is the name associated with a particular community.

ioctl A control device. This function performs a variety of control functions on devices and STREAMS. For non-STREAMS, the functions performed by this call are device-specific control functions.

IP link A communication medium over which nodes communicate at the link layer. The link layer is the layer immediately below IPv4/IPv6. Examples include Ethernets (simple or bridged) or ATM networks.

IPv4 Internet Protocol version 4.

IPv6 Internet Protocol version 6. IPv6 increases the address space from 32 to 128 bits. It is backwards compatible with IPv4.

IOSRAM Input-Output Static Random-Access Memory.

IPMP IP Network Multipathing. Solaris software which provides load spreading and failover for multiple network interface cards connected to the same IP link, for example, Ethernet.

J

JTAG A serial scan interface specified by IEEE standard 1149.1. The name comes from Joint Test Action Group, which initially designed it.

JTAG+ An extension of JTAG, developed by Sun Microsystems Inc., which adds a control line to signal that board and ring addresses are being shifted on the serial data line. Often referred to simply as JTAG.

K

kadb kadb is an interactive kernel debugger with a user interface. For more information refer to the kadb(1M) Solaris man page.

L

LCD Liquid Crystal Display.

LED Light Emitting Diode.

M

MAC address Worldwide unique serial number assigned to a network interface. IEEE controls the distribution of MAC addresses. See also *Ethernet address*.

mailbox See *Mbox*.

MAN SMS Management Network.

MaxCPU Dual CPU board.

Mbox Message passing mechanism between SMS software on the SC and OpenBoot PROM and the Solaris operating environment on the domain.

MIB Management Information Base.

metadisk A disk abstraction that provides access to an underlying group of two physical paths to a disk.

metanetwork A network abstraction that provides access to an underlying group of two physical paths to a network.

N

network interface card (NIC) Network adapter which is either internal or a separate card that serves as an interface to an IP link.

network time protocol (NTP) Network Time Protocol. Supports synchronization of Solaris time with the time service provided by a remote host.

NIC See *network interface card (NIC)*.

NIS+ Network Information Service Plus. A secure, hierarchical network naming service.

no-domain Describes the state of a board (DCU) that is not assigned to any domain.

NTP See *network time protocol (NTP)*.

O

OBP See *OpenBoot PROM*.

OpenBoot PROM A layer of software that takes control of the configured Sun Fire 15K system from `hpost`(1M), builds some data structures in memory, and boots the operating system. IEEE 1275-compliant OpenBoot PROM.

OS Operating system.

OSR Operating system resource.

P

path group A set of two alternate paths that provide access to the same device or set of devices.

physical path The electrical path from the host to a disk or network.

Platform A single physical computer.

POR Power-on-reset.

POST See *power-on self-test (POST).*

power-on self-test (POST) A test performed by `hpost`(1M). This is the program that takes uninitialized Sun Fire 15K hardware and probes and tests its components, configures what seems worthwhile into a coherent initialized system, and hands it off to OpenBoot PROM. In the Sun Fire 15K POST is implemented in a hierarchical manner with the following components: lpost, spost, and hpost.

PROM Programmable Read Only Memory.

R

RAM Random access memory.

RARP Reverse Address Resolution Protocol.

rstop *Record Stop*

Record Stop A correctable data transmission error.

RPC Remote procedure call.

S

SBBC See *BBC.*

SC System Controller. The Nordica board that assists in monitoring or controlling the system.

SEEPROM Serial EEPROM.

SMP Symmetric multi-processor.

SMS System Management Services software. The software that runs on the Sun Fire 15K SC and provides control/monitoring functions for the Sun Fire 15K platform.

SNMP Simple Network Management Protocol.

split-brain condition When both SCs think they are the main SC.

SRAM See *static RAM (SRAM).*

static RAM (SRAM) Memory chips that retain their contents as long as power is maintained.

System Board For next-generation Sun FIre servers, there are five types of system boards, four of which can be found in the Sun Fire 15K system. The system boards are: the CPU/Memory board, the I/O board, the WCI board, the Sun Fire 15K PCI controller board, and the Sun Fire 15K compact PCI controller board.

T

TCP/IP Transmission Control Protocol/Internet Protocol.

TOD Time of day.

tunnel switch The process of moving the SC/Domain communications tunnel from one IO board to another in a domain. Typically occurring when a the IO board with the tunnel is being dynamically reconfigured out.

U

URL Uniform Resource Locator.

UltraSPARC The UltraSPARC processor is the processor module used in the Sun Fire 15K system.

V

virtual keyswitch The SC provides a virtual keyswitch for each domain which controls the bringup process for each domain. The `setkeyswitch`(1M) command controls the position of the virtual keyswitch for each domain. Possible positions are: `on`, `off`, `standby`, `diag` and `secure`.

W

wPCI Sun Fire Link I/O assembly.

X

XIR eXternally Initiated Reset. Sends a “soft” reset to the CPU in a domain. It does not reboot the domain. After receiving the reset, the CPU drop to the OpenBoot PROM prompt.

Index

E

F

H

I

K

L

M

N

O

P

R

S

T

X