



Sun Fire™ 6800/4810/4800/3800 System Auto Diagnosis and Recovery Enhancements

Updated for Firmware Release 5.15.3

*Peter Gonscherowski and Tricia Wittsack, PTS
Engineering Group*

Sun BluePrints™ OnLine—October 2003



<http://www.sun.com/blueprints>

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95045 U.S.A.
650 960-1300

Part No. 817-3910-10
Revision A, October 03
Edition: October 2003

Copyright 2003 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, U.S.A. All rights reserved.

Sun Microsystems, Inc. has intellectual property rights relating to technology that is described in this document. In particular, and without limitation, these intellectual property rights may include one or more of the U.S. patents listed at <http://www.sun.com/patents> and one or more additional patents or pending patent applications in the U.S. and in other countries.

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

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

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and in other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, Sun BluePrints, Sun Enterprise, Sun Fire, and Solaris, are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the US and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

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

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 2003 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, Etats-Unis. Tous droits réservés.

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

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

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

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

Sun, Sun Microsystems, le logo Sun, Sun BluePrints, Sun Enterprise, Sun Fire, et Solaris sont des marques de fabrique ou des marques déposées de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays.

Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciées de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.



Please



Adobe PostScript

Sun Fire™ 6800/4810/4800/3800 System Auto Diagnosis and Recovery Enhancements

This article describes the Sun Fire 6800/4810/4800/3800 system availability enhancements provided in the system controller (SC) firmware versions 5.15.0 and 5.15.3 releases and the Solaris OE kernel updates. This document is useful for support personnel who have a basic technical knowledge of the Sun Fire 6800/4810/4800/3800 systems.

This article covers the following topics:

- System Controller Firmware Enhancements
- Solaris OE Enhancements
- Conclusion

Enhancements have been added to both the Solaris™ Operating Environment (Solaris OE) and the Sun Fire firmware release 5.15.3. Improved auto diagnosis of hardware failures and system recovery are now available. These enhancements achieve increased availability and better serviceability of the Sun Fire™ 6800/4810/4800/3800 systems. Both firmware version 5.15.3 and either the Solaris 8 OE kernel update 24 or Solaris 9 OE kernel update 5 are required to benefit from these enhancements.

TABLE 1 lists the patches needed to benefit from the Sun Fire 6800/4810/4800/3800 auto diagnosis and recovery enhancements.

TABLE 1 Required Patches

112884-04	Sun Fire firmware release 5.15.3
108528-24	Solaris 8 OE kernel update 24
112233-09 and 116009-01	Solaris 9 OE kernel update 9 and patch 116009-01

Sun Fire firmware release 5.15.3 of the system controller (SC) introduces the ability to receive hardware failure messages from the Solaris OE and ensure that the identified hardware failure is not configured in future reboots or `setkeyswitch on` and `off` events. In Sun Fire firmware release 5.15.0, several enhancements were made to improve the availability, serviceability, diagnosability, and repair characteristics of Sun Fire 6800/4810/4800/3800 systems. This document discusses both the existing firmware release 5.15.0 enhancements and the recent firmware release 5.15.3 enhancements:

- Sun Fire firmware release 5.15.0 enhancements:
 - Auto diagnosis—Automated diagnosis of runtime hardware faults
 - Component health status—Persistent record of the information stored in the component
 - Auto restoration—Automatic restoration of a domain
 - Domain hang recovery—Detects and recovers from a domain hang
 - Repeated domain panic recovery—Runs power on-self test (POST) at increasing diagnosis levels to identify and isolate the faulty hardware (if any)
- Sun Fire firmware release 5.15.3 enhancement:
 - Persistent record of hardware failures identified by the Solaris OE—receives hardware failure messages from the Solaris OE and stores the component health status in the affected FRU

Additionally, some enhancements have been made in the Solaris OE to improve the availability of the domain:

- CPU off-lining—Off-line a CPU when an L2_SRAM module has an increased probability of experiencing fatal errors
- Communicate hardware failures to the system controller—Sends a message to the system controller when the Solaris OE identifies and isolates a faulty component

System Controller Firmware Enhancements

When the system encounters a fatal hardware error that causes a domain to be *error paused*, the hardware fault is automatically diagnosed. The auto diagnosis (AD) enhancement updates the component health status (CHS) on the affected FRU if the hardware failure can be isolated to a specific FRU or a set of specific FRUs. During the auto-restoration phase, POST consults the CHS and restores the domain with the fault isolated.

In addition to the preceding enhancements, if POST encounters a test failure, the CHS is stored in the appropriate FRU.

The SC firmware is enhanced to detect domain hangs and recover from such situations by resetting and rebooting the domain. Another SC firmware enhancement runs POST at increasing diagnostic levels when the domain panics repeatedly, so that the system can identify and isolate any persistent hardware faults.

Auto Diagnosis

The SC monitors the domains for hardware faults. AD is automatically invoked on hardware faults that cause a *domain pause* or *data parity errors*. On Sun Fire 6800/4810/4800/3800 systems the data path is protected by parity and ECC. Domain operation is not impacted if data parity errors occur. Domain pauses are fatal errors and stop domain operation. AD analyzes the following errors:

- Interconnect errors
- Data parity errors
- Internal ASIC errors

FIGURE 1 shows the AD phase, Steps 1 through 5. Depending on the fault, three types of diagnosis results are possible:

- Fault diagnosed to a single component
- Fault diagnosed to a set of components
- Unresolved fault diagnosis

Note that when a fault is diagnosed to a set of components, it does not mean that all the components are faulty, just that the fault is located in a subset of these components (usually one).

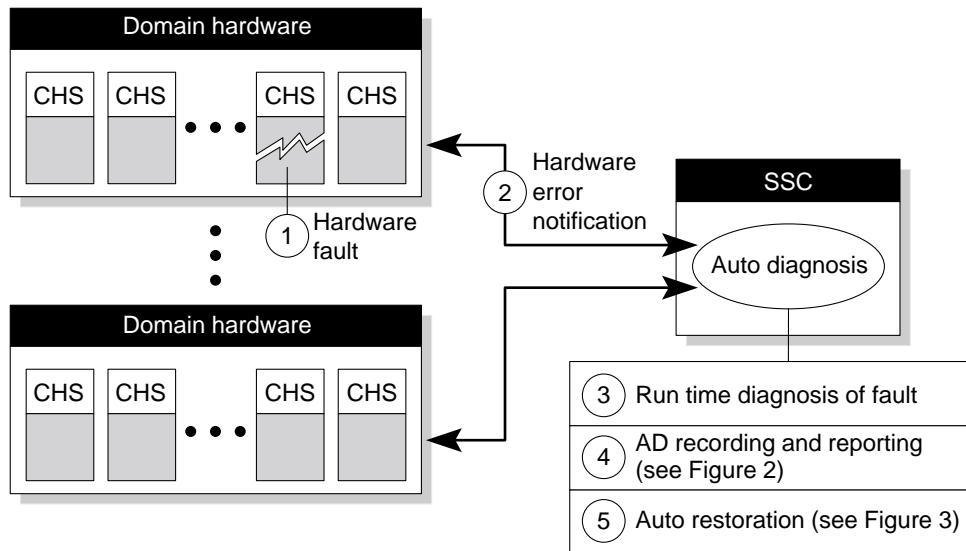


FIGURE 1 Auto Diagnosis Process

Auto Diagnosis Recording and Reporting

After the fault has been diagnosed, AD records its diagnosis persistently in the CHS and reports it to the domain console and loghost as shown in FIGURE 2.

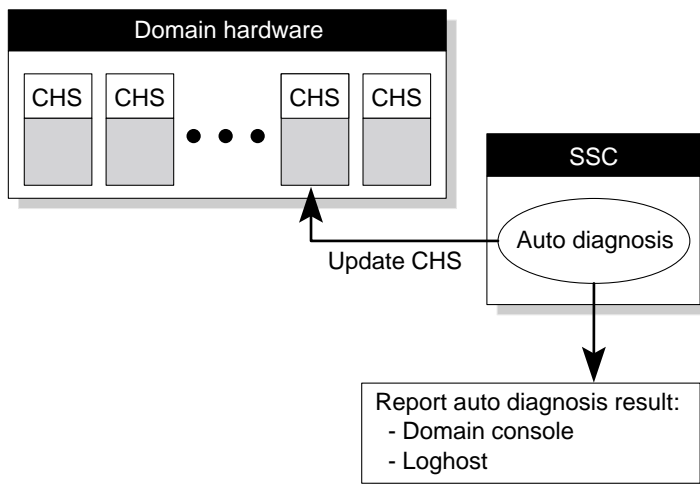


FIGURE 2 Auto Diagnosis Recording and Reporting

TABLE 2, “Example 1,” shows the AD result that is output to the domain console for a single FRU diagnosis.

TABLE 2 Example 1

```
[AD] Event: SF3800.ASIC.SDC.PAR_SGL_ERR.60111010
      CSN: 124H58EE DomainID: A ADInfo: 1.SCAPP.15.0
      Time: Thu Jan 23 20:47:11 PST 2003
      FRU-List-Count: 1;FRU-PN:5014362;FRU-SN: 011600; FRU-LOC:/N0/SB0
      Recommended-Action: Service action required
```

AD reports a unique event code for the failure type and the diagnostic time. A full description of the AD output format is in the *Sun Fire 6800/4810/4800/ 3800 Systems Platform Administration* manual. In this example AD determined that the error is within CPU/Memory board at FRU-LOC:N0/SB0.

The reported information enables your service provider to make a quick determination of the problem and initiate corrective service action.

CHS on a Sun Fire 6800/4810/4800/3800 is implemented for the following FRUs and components:

- CPU/Memory boards
- CPUs
- L2_SRAM modules
- DIMMs
- I/O assemblies
- Fireplane switches

Since the CHS and diagnostic information is persistently stored on a component, it moves with the component, which prevents the recurrence of a fault even if the component is moved to a different location. Preventing the recurrence of a fault improves the availability characteristic of Sun Fire 6800/4810/4800/3800 systems. The diagnosis information is contained inside the component. This makes service and repair of these systems easier.

Auto Restoration

POST performs the domain auto restoration function. POST runs automatically after AD or it is manually started by issuing the `setkey` command on the SC. POST consults the CHS of the domain hardware and reconfigures the domain to isolate the fault (FIGURE 3).

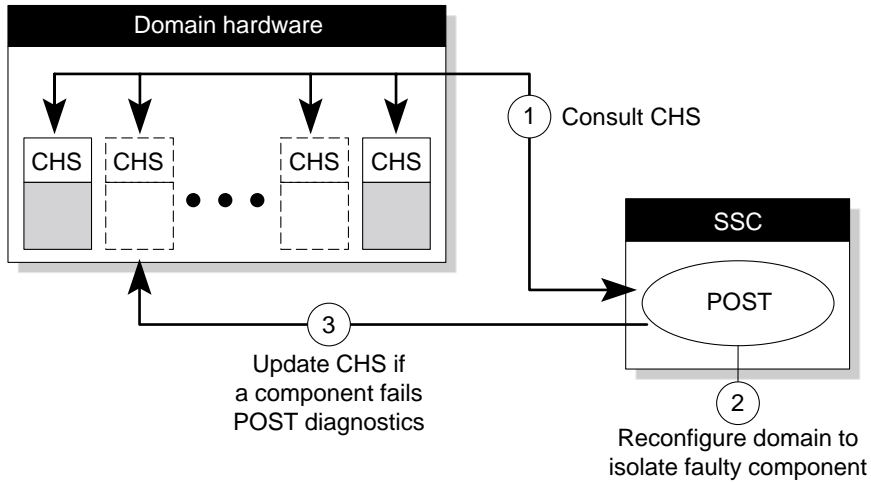


FIGURE 3 Auto Restoration

After the domain has been restored, you can run the `showcomponent` command to check which components have been disabled due to CHS.

If a FRU or component is disabled because of its CHS, immediate replacement is not necessary because the domain is restored with the fault isolated. Utilizing dynamic reconfiguration (DR), the FRU can be replaced at any time with minimal impact to the Solaris OE and user applications. For more information about DR, see [1] *Sun BluePrints OnLine* article “Sun Fire™ 3800-6800 Servers Dynamic Reconfiguration.”

Domain Hang Recovery

A situation in which a domain is not updating its heartbeats or is unreachable by using the console is categorized as a *domain hang*. A domain's heartbeat is a communication mechanism that informs the SC that it is alive. If a domain is not updating its heartbeat, this indicates a domain hang. A domain hang can occur due to hardware or software issues.

When using firmware version 5.15.0 or above on a Sun Fire 6800/4810/4800/3800 system, the SC acts as an external monitor for each domain. The SC monitors for a domain hang condition and initiates an XIR domain reset if the domain heartbeat register is not updated within the maximum time out limit. The domain heartbeat

monitoring is configurable for each domain using the `watchdog_timeout_seconds` parameter in the `/etc/systems` file of each domain.

The default time out value for a domain is three minutes. For additional details, refer to the `system(4)` man page. If `watchdog_timeout_seconds` is set to a value less than three minutes, the SC defaults to three minutes.

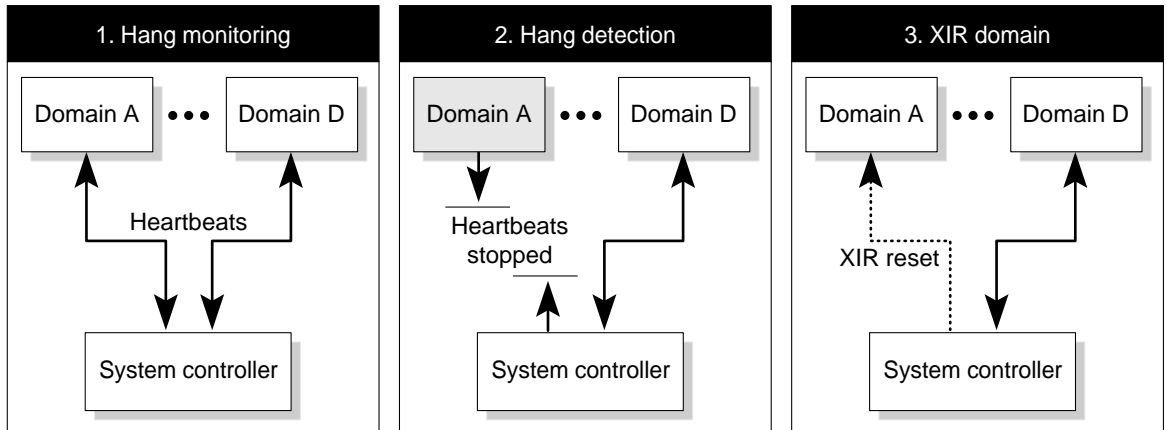


FIGURE 4 Domain Hang Restoration

TABLE 3, “Example 2,” shows the console output of a domain that was declared hung and was reset by the SC.

TABLE 3 Example 2

```
Jan 22 17:02:06 sc0 Domain-A.SC: Domain watchdog timer expired.
Jan 22 17:02:06 sc0 Domain-A.SC: Using default hang-policy (RESET).
```

In addition to the heartbeat monitoring, the SC also checks if the domain is picking up the interrupts sent to it by the SC. The SC sends interrupts to the domain when, for example, characters are entered on the domain console. If on a second interrupt, the previous one has not been picked up by the domain, the SC waits for one minute before declaring the domain hung. TABLE 4, “Example 3,” shows the console output of a domain that is hung because it has not been picking up its interrupts.

TABLE 4 Example 3

```
Jan 22 18:09:02 sc0 Domain-A.SC: Domain is not responding to interrupts.
Jan 22 18:09:02 sc0 Domain-A.SC: hang-policy is NOTIFY. Not resetting domain.
```

The *hang policy* is set by the `setupdomain` command to `notify` or `reset`. If set to `notify`, the SC reports the hang condition on the domain console and does not reset the domains (TABLE 4, “Example 3”). If set to `reset`, the SC reports the hang condition on the domain console and initiates a domain reset (TABLE 3, “example 2”). By default the hang policy is set up to dump core when it is reset. To identify the cause of the domain hang, consult your service provider while referring to the core file.

A system's hang-policy can be verified using the `showdomain` command in the domain shell. For more information about domain setup, refer to the *Sun Fire™ 6800/4810/4800/3800 Systems Platform Administration* manual.

Recovery From Repeated Domain Panics

Domain panics can be caused by software or hardware. To prevent hardware faults from causing panic-reboot loops, the SC firmware runs POST at increasing diagnostic levels on recurring panics.

On the first panic, the domain reboots and writes a core file, which can be used to analyze the problem. However, if further panics occur within a short time period, it is desirable to run POST automatically at a higher level as part of domain restoration. POST diagnostics verify the status of the hardware and could identify and isolate faulty components (if any). After identifying faulty components, POST updates the appropriate CHSs. With firmware release 5.14.0 and higher, the SC keeps track of the number of domain panics over time. A panic reboot of a domain has a unique register signature that differs from the normal reboot of a domain. If the domain is manually rebooted in the meantime, the panic-reboot counter is reset.

On recurring panics, the domain POST diagnostic level is increased to the next higher level from `diag-level quick`. In increasing order, POST levels are `init`, `quick`, `default`, `mem1`, and `mem2`. The domain is put into standby position if it continues to panic undetected by the user after the highest level of POST is run (FIGURE 5). For further analysis, consult your service provider while referring to the core file.

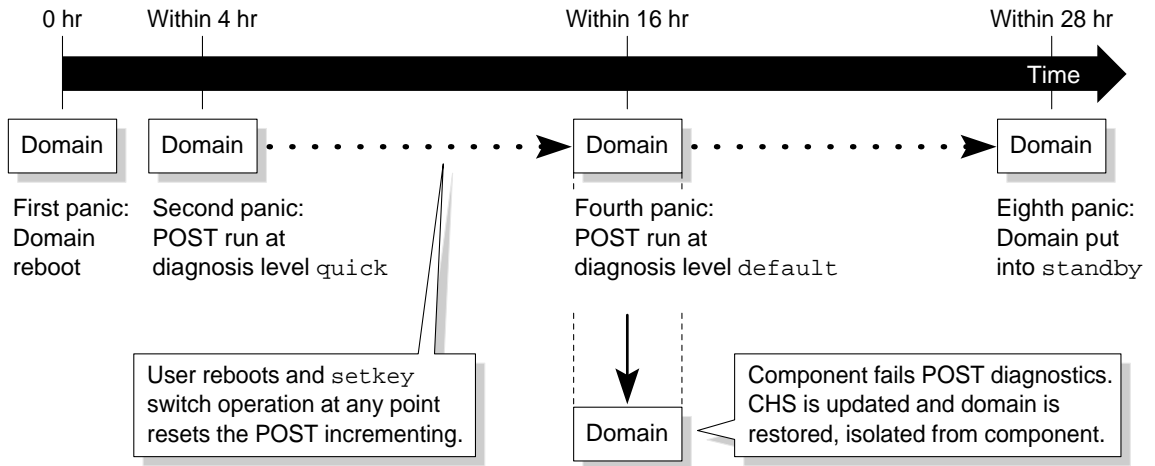


FIGURE 5 Domain Panic Restoration

This feature prevents a panic-reboot loop of domains. If recurring panics are caused by a software bug, the increased POST level minimizes hardware as a possible cause. Downtime for running further POST diagnostics is not required because the system automatically takes the necessary measures.

Persistent Record of Hardware Failures From the Solaris OE

As of Sun Fire firmware release 5.15.3, hardware failures diagnosed by the Solaris OE are persistent. The SC receives and stores hardware failure messages from the Solaris OE if the system is using the appropriate Solaris kernel update that applies to Solaris 8 (KU-108528-24) or Solaris 9 (KU-112233-09). Solaris 9 kernel update -09 also requires patch 116009-01. Domain reboots and `setkeyswitch off` and `on` events no longer configure components that the Solaris OE has previously diagnosed as failed.

When the system controller receives a hardware event message from the Solaris OE, the SC updates the CHS of the affected FRU. Any future POST consults the CHS and does not configure any listed faulty components. System availability is improved as future domain configurations deconfigure the failed L2_SRAM modules. The Solaris OE can currently identify and isolate the following types of faults:

- L2_SRAM ECC SERD
- L2_SRAM ECC UC

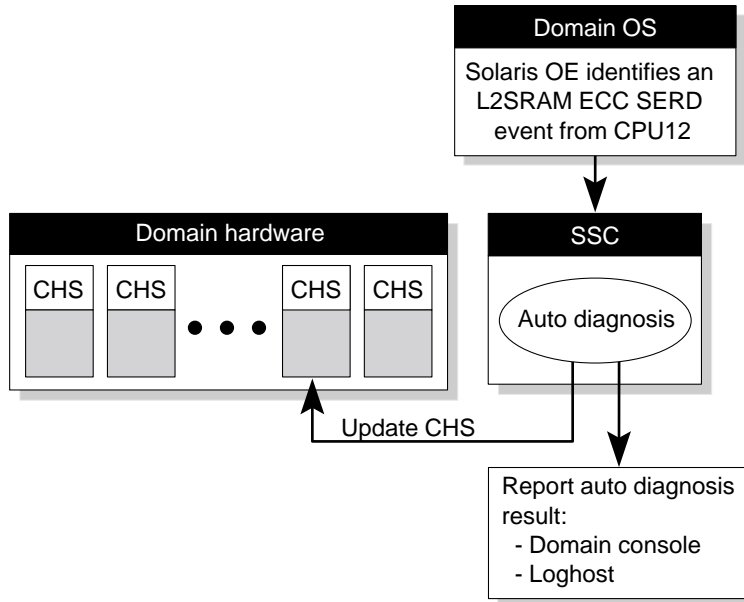


FIGURE 6 Persistent Failure Record

The soft error rate discrimination (SERD) algorithm detects when a specified number of distinct CPU events have occurred on the same processor in a 24-hour period. After the specified CPU SERD events, the CPU becomes a candidate for Solaris OE off-lining. The AD failure message (TABLE 5, “Example 4,”) specifies the phrase SF-SOLARIS-DE in the ADInfo Domain Log message to identify which hardware failures were received from the Solaris OE.

TABLE 5 Example 4

<pre> [DOM] Event: SF6800.L2SRAM.SERD.f.1b.1004000000091.f4470000 CSN: 044M347B DomainID: A ADInfo: 1.SF-SOLARIS DE.5_9_GENERIC_112233-09 Time: Mon Jun 02 23:34:59 PDT 2003 FRU-List-Count: 1; FRU-PN: 3704125; FRU-SN: 090K01; FRU-LOC: /N0/SB3/P3/E0 Recommended-Action: Service action required </pre>
--

The SC’s persistent failure record of hardware faults identified by the Solaris OE improves availability. System serviceability, diagnosis, and repair are also easier and quicker.

Solaris OE Enhancements

Kernel updates for Solaris 8 OE and Solaris 9 OE on UltraSPARC™ III systems enhance the correctable error (CE) L2_SRAM module handling. Multiple CEs on accessing an L2_SRAM module indicate a higher probability of experiencing an uncorrectable error (UE). To prevent a fatal UE, the Solaris OE attempts to off-line CPUs. The availability of domains increases because the Solaris OE does not access L2_SRAM modules that have an increased failure probability.

The enhanced Solaris OE kernels have the ability to communicate hardware failures to the SC. If the system is using the appropriate kernel update for the Solaris 8 OE (KU-108528-24) or the Solaris 9 OE (KU-112233-09) with patch 116009-01, a message is sent to the SC when the Solaris OE identifies and isolates a faulty L2_SRAM module. The failed L2_SRAM module is not reconfigured into a domain on future domain reboots or `setkeyswitch off` and `on` operations because the system controller has recorded the component as faulty in its CHS.

CPU Off-lining

Similar to memory page retirement, the Solaris OE keeps track of the number of ECCs over time on an L2_SRAM module (FIGURE 7). Two types of ECCs are considered here—nonfatal multibit errors (UCU, CPU, WDU, EDU) and nonfatal single-bit correctable errors (UCC, CPC, WDC, EDC). If an L2_SRAM module experiences one nonfatal multibit error or three single-bit correctable errors in a 24-hour window, the L2_SRAM module is diagnosed with an increased probability of suffering a fatal failure in future. In this scenario, the Solaris OE has been enhanced to automatically attempt to off-line the affected CPU module. It is possible that the CPU off-line may not succeed because there might be processes bound to that CPU.

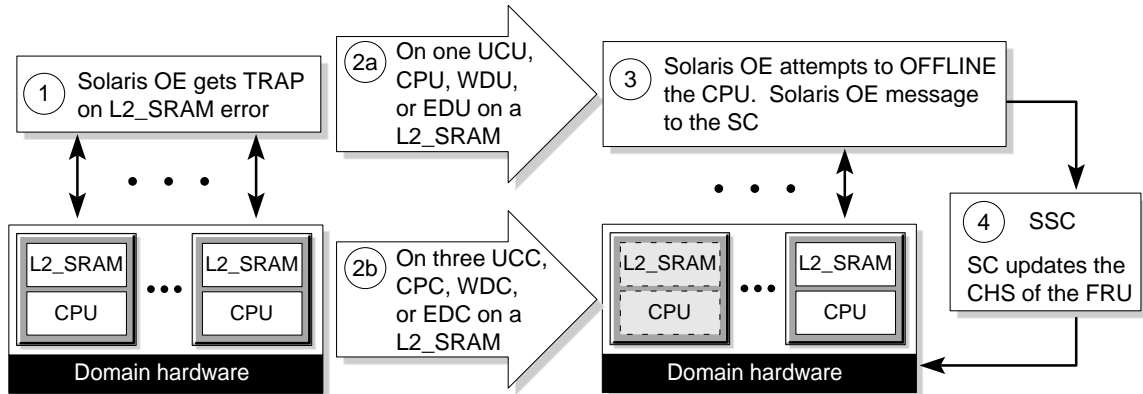


FIGURE 7 Solaris OE L2_SRAM Error Handling

TABLE 6, “Example 6,” shows the messages on successfully off-lining a CPU that experienced more than two CE events in a 24-hour window.

TABLE 6 Example 6

```
Feb  3 06:38:40 doma SUNW,UltraSPARC-III: NOTICE: [AFT1] CPU6
offlined due to more than 2 xxC Events in 24:00:00 (hh:mm:ss)
```

Once a CPU is off-lined the Solaris OE sends a message to the system controller. The system controller updates the CHS of the affected FRU so that the faulty CPU is not configured into a domain on future reboots or `setkeyswitch` off and on events.

Off-lining the CPU associated with L2_SRAM modules with a higher probability of experiencing a fatal error increases the availability of the Solaris OE. Communication between the Solaris OE and the SC to persistently store the CHS increases availability and provides easier diagnosis and serviceability of the system. Dynamically reconfigured CPU/Memory boards can be replaced with minimal impact to the Solaris OE and user applications.

Conclusion

To benefit from the enhancements, upgrading to firmware version 5.15.3 and KU 108528-24 for Solaris 8 OE and KU 112233-09 for Solaris 9 OE plus patch 116009-01 is required. In a mission critical environment, high availability is achieved by system resiliency, the appropriate configuration, serviceability, and efficient and automated

restoration processes. The enhancements made in version 5.15.3 address all of those elements and increase availability, serviceability, and diagnosability of Sun Fire 6800/4810/4800/3800 systems.

TABLE 7 lists the requirements for the various Auto Diagnosis and Recovery Features that are available for the Sun Fire 6800/4800/4810/3800 Systems. Note that SC firmware release 5.15.3 requires the domains use Solaris 8 KU24 or Solaris 9 KU6 with patch 116009-01. For additional patch IDs, refer to TABLE 1.

TABLE 7 Auto Diagnosis and Recovery Feature Kernel Update Requirements

SC firmware version	5.15.0 - 5.15.2	5.15.3
Auto diagnosis	√	√
Component health status	√	√
Domain hang recovery	√	√
Repeated domain panic recovery	√	√
Persistent updating of CHS by the Solaris OE for CPU off-line for L2_SRAM conditions		√ With either Solaris 8 KU 24 (108528-24) or Solaris 9 KU 09 (112233-09) and patch 116009-01

TABLE 8 lists the enhancements that are within the Solaris OE only.

TABLE 8 Enhancements Within the Solaris OE Only

CPU off-Lining for L2_SRAM conditions within Solaris OE	Solaris 8 OE KU20 108528-20) or Solaris 9 KU09 (112233-09)
---	--

Standardized messaging, CHS, and AD are powerful tools for users and service providers. AD in combination with DR on Sun Fire 6800/4810/4800/3800 systems greatly increases the availability and decreases scheduled downtime for maintenance.

The proactive actions of the Solaris OE on L2_SRAM events makes the Solaris OE more resilient and increases its availability.

References

[1] Gonscherowski, Peter, "Sun Fire™ 3800-6800 Servers Dynamic Reconfiguration," *Sun BluePrints Online*, April 2002

Related Resources

Sun Fire™ 6800/4810/4800/3800 Systems Firmware 5.14.X Release Notes, Sun Microsystems, 817-1305

Sun Fire™ 6800/4810/4800/3800 Systems Firmware 5.15.X Release Notes, Sun Microsystems, 817-1001

Sun Fire™ 6800/4810/4800/3800 Systems Platform Administration Manual, Sun Microsystems, 817-0999

Sun Fire™ 6800/4810/4800/3800 System Controller Command Reference Manual, Sun Microsystems, 817-1000

Acknowledgments

Thanks to Kris Bayuk who encouraged us to write this Sun Blueprint article and Domingo Goyena, James Hsieh, David Re, Sundeep Panicker, Frederic Vecoven, Jim Quigley, Mark A. Matthews, and the Serviceability Team under the lead of Marc Young for their constructive feedback and time reviewing the document.

About the Authors

Peter Gonscherowski is a member of the PTS Engineering group in Newark, California. He is a member of the team responsible for providing engineering support for Sun Fire 6800/4810/4800/3800 and Sun Enterprise servers. Prior to his current role, Peter worked as a Mission Critical engineer and back-line engineer for Sun Enterprise Services in Europe.

Tricia Wittsack is a member of the PTS Engineering group in Newark, California. She is a member of the team responsible for providing engineering support for Sun Fire 6800/4810/4800/3800 and Sun Enterprise servers. Prior to her current role, Tricia worked in the PTS Global Labs in Newark, California.

