



Sun Fire™ E6900/E4900 Systems Overview

Sun Microsystems, Inc.
www.sun.com

Part No. 817-4119-10
February 2004, Revision 01

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

Copyright 2004 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, AnswerBook2, docs.sun.com, Sun Fire, and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc. in the U.S. and in other countries.

All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and in 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 2004 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, Californie 95054, Etats-Unis. Tous droits réservés.

Sun Microsystems, Inc. a les droits de propriété intellectuelle relatant à la technologie qui est décrit dans ce document. En particulier, et sans la limitation, ces droits de propriété intellectuelle 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, AnswerBook2, docs.sun.com, 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.

LA DOCUMENTATION EST FOURNIE "EN L'ÉTAT" ET TOUTES AUTRES CONDITIONS, DECLARATIONS ET GARANTIES EXPRESSES OU TACITES SONT FORMELLEMENT EXCLUES, DANS LA MESURE AUTORISEE PAR LA LOI APPLICABLE, Y COMPRIS NOTAMMENT TOUTE GARANTIE IMPLICITE RELATIVE A LA QUALITE MARCHANDE, A L'APTITUDE A UNE UTILISATION PARTICULIERE OU A L'ABSENCE DE CONTREFAÇON.



Please
Recycle



Adobe PostScript

Contents

Preface vii

1. Sun Fire Product Overview 1-1

1.1 Standard Features 1-1

1.2 Machine Configurations 1-4

1.2.1 Sun Fire E6900 System 1-4

1.2.2 Sun Fire E4900 System 1-6

2. System Features and Capabilities 2-1

2.1 Partitions and Domains 2-2

2.1.1 Partitions 2-2

2.1.2 Domains 2-3

2.2 Reliability 2-4

2.2.1 Reducing the Probability of Error 2-4

2.2.2 Correcting Errors Using Error-Correcting Code 2-4

2.2.2.1 Error-Correcting Code Protection of the Data
Interconnect 2-5

2.2.3 Detecting Uncorrectable Errors 2-5

2.2.3.1 Multiple-Bit Data Errors 2-5

2.2.3.2 Address Errors 2-5

2.2.3.3 System Timeout Errors 2-6

- 2.2.3.4 Power Corrected Failures 2-6
 - 2.2.4 Environmental Sensing 2-6
 - 2.2.4.1 Temperature 2-6
 - 2.2.4.2 Power Subsystem 2-6
 - 2.3 Availability 2-7
 - 2.3.1 Availability Goals for the Sun Fire System 2-7
 - 2.3.2 High Availability Capabilities of the Sun Fire System 2-7
 - 2.3.2.1 Cooling 2-8
 - 2.3.2.2 AC Power Switching 2-8
 - 2.3.2.3 ECC 2-8
 - 2.3.3 Resiliency Capabilities 2-8
 - 2.3.3.1 DC Power 2-9
 - 2.3.3.2 Logic Boards 2-9
 - 2.3.3.3 Processor 2-9
 - 2.3.3.4 Memory 2-9
 - 2.3.4 Redundant Components 2-10
 - 2.4 Serviceability Capabilities 2-10
 - 2.4.1 Mechanical Serviceability 2-11
 - 2.4.1.1 DC Power Supply Removal and Replacement 2-11
 - 2.4.1.2 Fan Tray Removal and Replacement 2-11
 - 2.4.1.3 Domain Isolation 2-11
 - 2.4.2 Nonconcurrent Service 2-11
 - 2.4.3 Remote Service 2-12
- 3. Hardware Overview 3-1**
 - 3.1 Standard Operation 3-1
 - 3.2 Data Interconnect 3-3
 - 3.3 Console Bus Interconnect 3-4

- 4. Sun Fire System Components 4-1**
 - 4.1 CPU/Memory Board 4-1
 - 4.2 I/O Assemblies 4-1
 - 4.2.1 PCI I/O 4-1
 - 4.3 Repeater Board 4-2
 - 4.4 System Controller Board, Version 2 With Enhanced Memory 4-2
 - 4.4.1 Redundant System Controllers 4-4
 - 4.4.2 Virtual Domain Key Switches 4-5
 - 4.4.3 Solaris Console 4-5
 - 4.4.4 Virtual Time Of Day 4-5
 - 4.4.5 Environmental Monitoring 4-5

Preface

This document provides the following information about the Sun Fire™ E6900/E4900 systems:

- Machine configurations of the Sun Fire E6900/E4900 systems
- Hardware overview
- System components
- Reliability, availability, and serviceability features

Typographic Conventions

Typeface ¹	Meaning	Examples
AaBbCc123	The names of commands, files, and directories; on-screen computer output	Edit your <code>.login</code> file. Use <code>ls -a</code> to list all files. % You have mail.
AaBbCc123	What you type, when contrasted with on-screen computer output	% su Password:
<i>AaBbCc123</i>	Book titles, new words or terms, words to be emphasized. Replace command-line variables with real names or values.	Read Chapter 6 in the <i>User's Guide</i> . These are called <i>class</i> options. You <i>must</i> be superuser to do this. To delete a file, type <code>rm filename</code> .

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

Related Documentation

Application	Title
Installation	<i>Sun Fire E6900/E4900 Systems Installation Guide</i>
Operation	<i>Sun Fire Cabinet Installation and Reference Guide</i> <i>Sun Fire E6900/E4900 System Getting Started</i> <i>Sun Fire E6900/E4900 Systems Service Manual</i>

Accessing Sun Documentation

You can view, print, or purchase a broad selection of Sun documentation, including localized versions, at:

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

Contacting Sun Technical Support

If you have technical questions about this product that are not answered in this document, go to:

<http://www.sun.com/service/contacting>

Sun Welcomes Your Comments

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

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

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

Sun Fire E6900/E4900 Systems Overview, part number 817-4119-10

Sun Fire Product Overview

This chapter discusses the features and capacity of the Sun Fire E6900 and Sun Fire E4900 systems.

This family of servers provides entry-level to high-end server functionality. The Sun Fire E6900 system has space for internal peripherals mounted within the 19-inch cabinet. You have the flexibility with the remaining system to install it in industry-standard 19-inch cabinets or have it preinstalled in a Sun Fire system cabinet. The Sun Fire system cabinet can hold one Sun Fire E4900 system.

1.1 Standard Features

The standard features of these systems include:

- Rackmountable in industry standard 19-inch rack (Sun Fire E4900)
- Support for up to 24 CPUs
- Support for up to 32 PCI I/O slots
- Extensive redundancy
- System controllers
- Support for multiple domains
- Concurrent hardware maintenance
- Common components
- Redundant power and cooling
- 9.6-Gbyte bus bandwidth

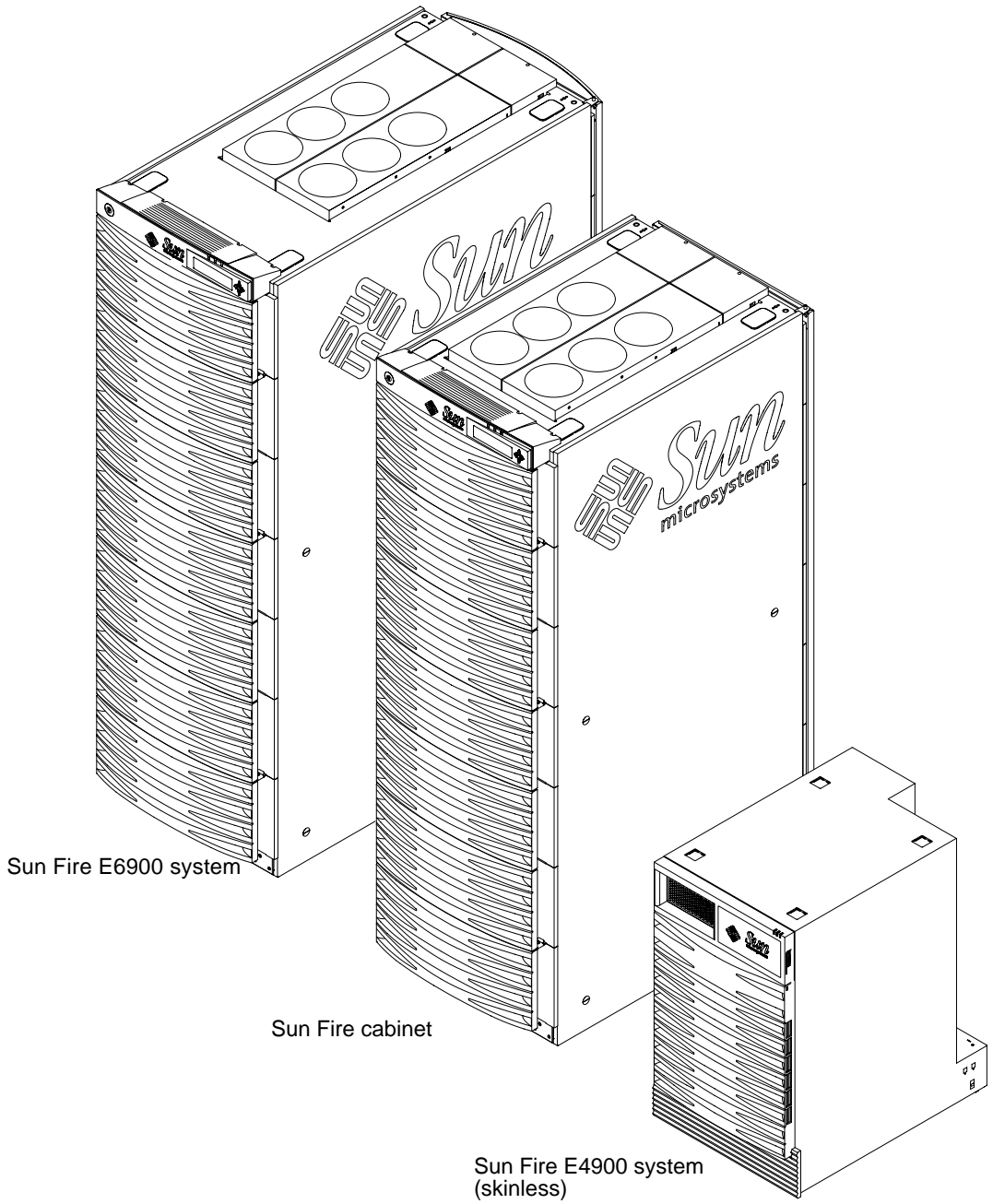


FIGURE 1-1 Sun Fire Systems and Sun Fire Cabinet

The Sun Fire E6900 and E4900 systems share the following components:

- CPU/Memory board
- CPU processors
- Memory DIMMs
- PCI I/O assembly
- PCI I/O cards
- System Controller board (version 2 with enhanced memory)
- Repeater board

1.2 Machine Configurations

Two machine configurations are available:

- Sun Fire E6900 system
- Sun Fire E4900 rackmountable system

1.2.1 Sun Fire E6900 System

The Sun Fire E6900 system has support for six CPU/Memory boards, four I/O assemblies, four Repeater boards, and two System Controller boards. Although there are four Repeater boards, they are logically two redundant repeaters (two boards together make up one logical repeater). FIGURE 1-2 shows front and rear views of the Sun Fire E6900 system cabinet. TABLE 1-1 lists the features of the Sun Fire E6900 system.

TABLE 1-1 Sun Fire E6900 System Features

Features	Quantity or Description
CPU/Memory boards	6
CPUs	24
Maximum memory	192 DIMM sockets
I/O assemblies	4 (PCI)
System Controller boards (version 2)	2
Repeater boards	4
Domains	4 maximum
Power supplies	6
Power requirements	200–240 VAC
Redundant cooling	Yes
Redundant AC input	Yes
Internal peripherals	None (However, space is available in the cabinet for peripherals options.)
Packaging	Sun Fire E6900 cabinet

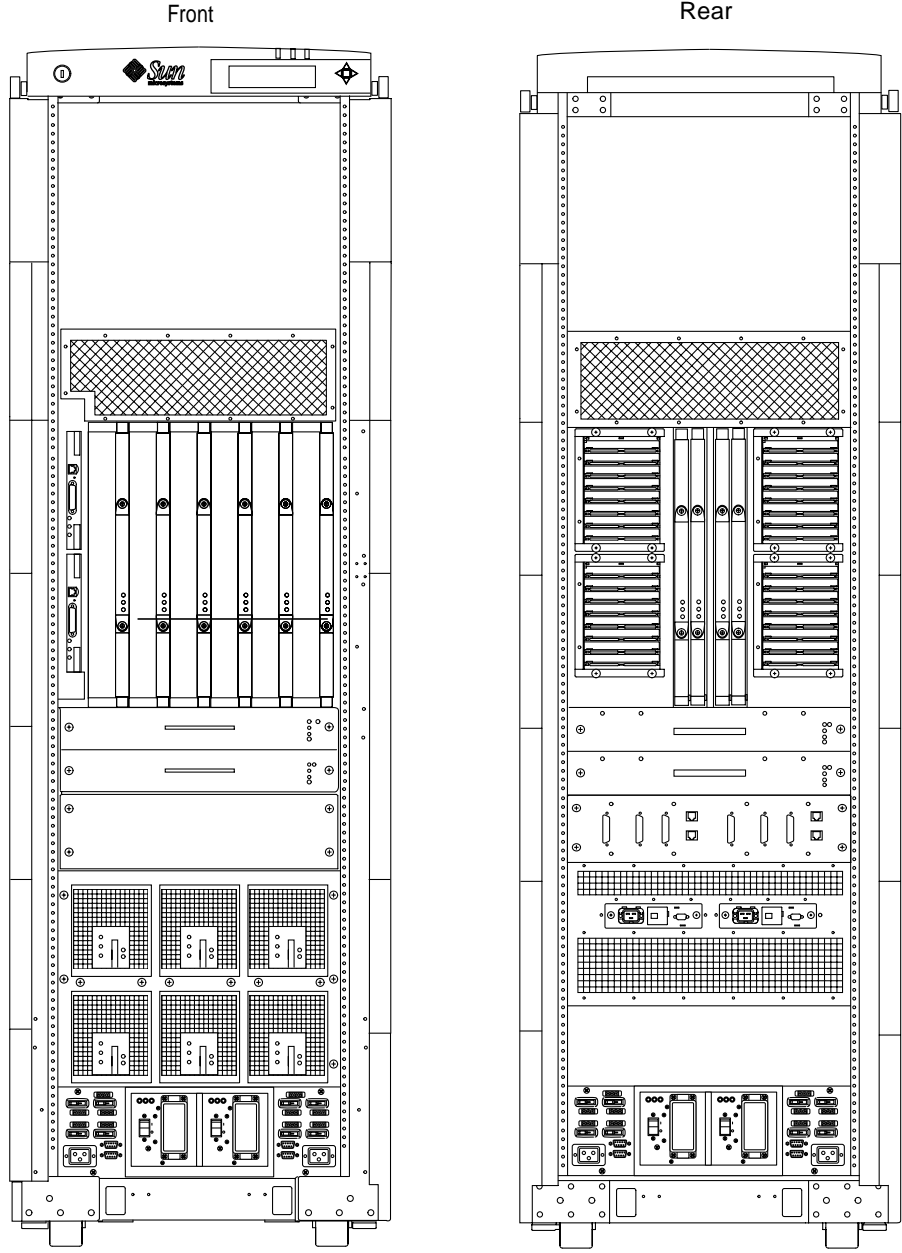


FIGURE 1-2 Sun Fire E6900 System Cabinet—Front and Rear Views

1.2.2 Sun Fire E4900 System

The Sun Fire E4900 system has support for three CPU/Memory boards, two I/O assemblies, two Repeater boards, and two System Controller boards. FIGURE 1-3 shows front and rear views of the Sun Fire E4900 system. TABLE 1-2 lists the features of the Sun Fire E4900 system.

TABLE 1-2 Sun Fire E4900 System Features

Features	Quantity or Description
CPU/Memory boards	3
CPUs	12
Maximum memory	96 DIMM sockets
I/O assemblies	2 (PCI)
System Controller boards (version 2)	2
Repeater boards	2
Domains	2 maximum
Power supplies	3
Power requirements	200–240 VAC
Redundant cooling	Yes
Redundant AC input	No
Internal peripherals	None
Packaging	Rackmountable, or mounted in a Sun Fire cabinet

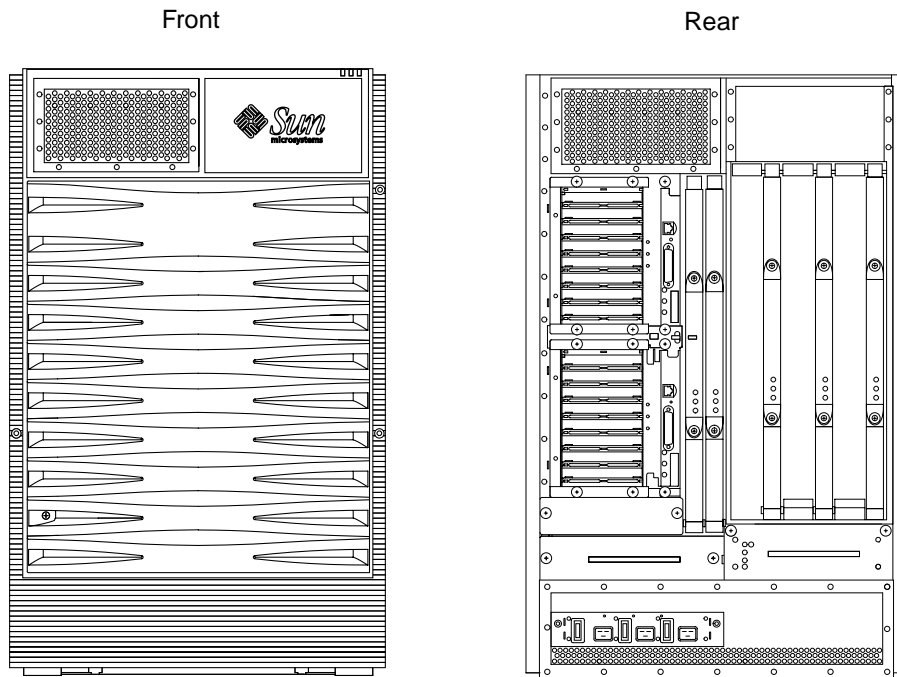


FIGURE 1-3 Sun Fire E4900 System—Front and Rear Views

System Features and Capabilities

The Sun Fire E6900/E4900 system's primary features include the ability to partition your system and create domains. These features provide greater reliability, availability, and serviceability, which means uptime. These features and capabilities are as follows:

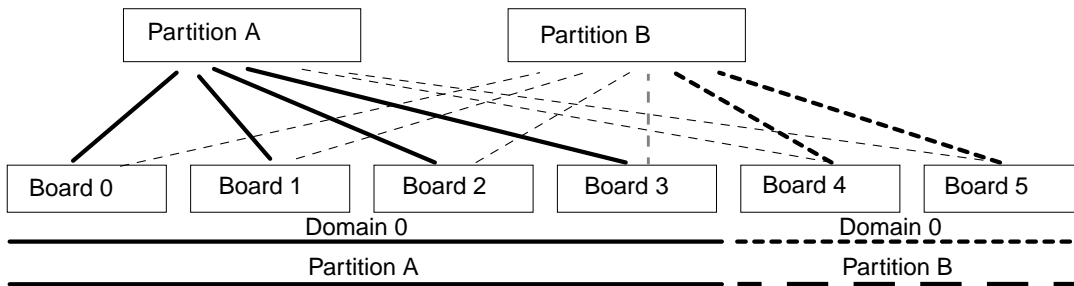
- *Partition*—The ability for the system to logically behave as two separate systems
- *Domain*—The ability to create logically independent multiple sections within a partition, with each domain running its own operating system
- *Reliability*—A function of the care with which the hardware and software design was executed, the quality of the components selected, and the quality of the manufacturing process (for example, ESD protection, clean rooms, and so forth)
- *Availability*—The percentage of time the customer's system is able to do productive work
- *Serviceability*—The system ensures that repair time (downtime) is kept to a minimum

2.1 Partitions and Domains

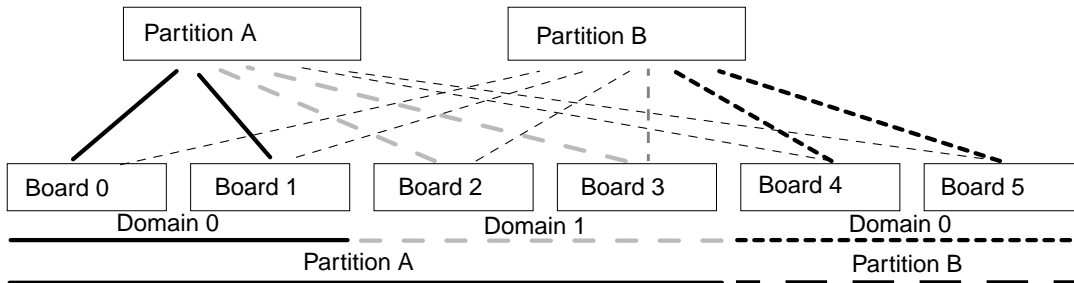
The Sun Fire system can be divided into partitions and domains. A single physical system can have multiple independent logical systems, each running its own operating system, by using partitions and domains. Partitions and domains differ only in terms of their flexibility and isolation.

FIGURE 2-1 Partitions and Domains on a Sun Fire E6900 System

Partitions A and B have separate Repeater boards.



A partition can have up to two domains.



- Active domain connection
- Active domain connection
- Active domain connection
- Inactive logical connection

2.1.1 Partitions

A single physical Sun Fire E6900 system can be divided into two partitions. All connections between boards of one partition and boards of the other partition are disabled. The system logically behaves as two separate systems.

If the partitions are assigned to the physical half of the Sun Fire E6900 system, then the power planes associated with each partition are also isolated. A Sun Fire E6900 system can be divided into two partitions by logically isolating one set of Repeater boards for each partition. Sun Fire E4900 systems also support two partitions.

Each partition on the Sun Fire E6900 system can have up to two domains, allowing for up to four domains total. For the Sun Fire E4900 system, if a single partition is established, it can support two domains; if two partitions are established, however, each partition will support only one domain.

2.1.2 Domains

The Sun Fire system can be logically divided into multiple domains. Since each domain is comprised of one or more system boards, a domain can have between one and 24 processors. Each domain runs its own instance of the operating system and has its own peripherals and network connections. You can configure domains without interrupting the operation of other domains on the same system.

Domains can be used for:

- Testing new applications
- Updating the operating system
- Configuring several domains to support separate departments

While production work continues on the remaining (and usually larger) domain, there will not be any adverse interaction between any of the domains. You can gain confidence in the correctness of applications without disturbing production work. When the testing work is complete, the system can be rejoined logically without rebooting (there are no physical changes when you use domains). Thus, if problems occur, the rest of your system is not affected.

The Sun Fire E6900 system can have up to four domains. The Sun Fire E4900 system can have up to two domains. Each instance of the Solaris™ Operating System runs in its own domain. Domains do not depend on each other and do not interact with each other.

A single partition on a Sun Fire E6900 system can be divided into two domains. Unlike partitions, domains share the Repeater boards. Each domain gets half the address bandwidth of a full system bus.

2.2 Reliability

The reliability capabilities of the Sun Fire system fall into four categories:

- Reducing the probability of errors
- Correcting errors using error-correcting code (ECC)
- Detecting uncorrectable errors
- Sensing environmental factors

2.2.1 Reducing the Probability of Error

All the ASICs are designed for worst-case temperature, voltage, frequency, and airflow combinations. The high level of logic integration in the ASICs reduces component and interconnect count.

A distributed power system improves power supply performance and reliability.

Extensive self-test upon power-on reboot after a hardware failure screens all of the key logic blocks in the Sun Fire system:

- Built-in self-test logic in all the ASICs.
- The power-on self-test (POST)—controlled from the System Controller board—tests each logic block first in isolation, and then with progressively more of the system. Failing components are electrically isolated from the centerplane. The result is that the system is booted only with logic blocks that have passed this self-test and that must operate without error.

All I/O cables have a positive lock mechanism and a strain-relief support to prevent accidental disconnections.

2.2.2 Correcting Errors Using Error-Correcting Code

The Sun Fire system contains a number of subsystems that are capable of recovering from errors without failing. Subsystems that have a large number of connections have greater odds of failure. The subsystems that have the highest probability of errors are protected from transient errors through the use of single-bit error correction that uses an error-correcting code.

2.2.2.1 Error-Correcting Code Protection of the Data Interconnect

The entire data path from the local data crossbars and the memory subsystem is protected by error-correcting code. Single-bit-data errors detected in these subsystems are corrected by the receiving UltraSPARC™ s400 module, and the system is notified for logging purposes that an error has occurred.

The memory subsystem does not check or correct errors but provides the extra storage bits. The Sun Fire data buffer chips use the error-correcting codes to assist in fault isolation.

If a correctable error is detected by the interconnect, the system controller is notified and enough information is saved to isolate the failure to a single net within the interconnect system. The data containing the error is sent through the interconnect unchanged, and the error is reported.

Memory errors are logged by software so that defective DIMMs can be identified and replaced during scheduled maintenance.

2.2.3 Detecting Uncorrectable Errors

Almost all internal system paths are protected by some form of redundant check mechanism. Transmission of bad data is thus detected, preventing propagation of bad data without notification. All uncorrectable errors result in an error condition. Recovery requires an operating system automatic reboot.

2.2.3.1 Multiple-Bit Data Errors

Multiple-bit ECC errors are detected by the receiving port, which notifies the operating system, so that depending upon what process is affected, the system as a whole can avoid failure.

Parity errors on external cache reads to the interconnect become multibit ECC data errors and are handled as other multibit errors.

2.2.3.2 Address Errors

Any single-bit or multiple-bit errors detected in the address interconnect are unrecoverable and are fatal to the operating system.

2.2.3.3 System Timeout Errors

Timeout errors detected by the port controller or memory controller are an indication of lost transactions. Timeouts are therefore always unrecoverable.

2.2.3.4 Power Corrected Failures

The Sun Fire system uses a highly reliable distributed power system. Each I/O subsystem, CPU/Memory board, System Controller board, or Repeater board within the system has DC-to-DC converters for that board only, with multiple converters for each voltage. When a DC-to-DC converter fails, the system controller is notified. The system board reporting the failure will then be deconfigured from the system. No guarantee is made regarding continued system operation at the time of the failure.

2.2.4 Environmental Sensing

The system chassis environment is monitored for key measures of system stability, such as temperature, airflow, and power supply performance. The system controller is constantly monitoring the system environmental sensors in order to have enough advance warning of a potential condition that the machine can be brought gracefully to a halt, avoiding physical damage to the system and possible corruption of data.

2.2.4.1 Temperature

The internal temperature of the system is monitored at key locations as a fail-safe mechanism. Based on temperature readings, the system can notify the administrator of a potential problem, begin an orderly shutdown, or power off the system immediately.

2.2.4.2 Power Subsystem

The Sun Fire system performs additional sensing to enhance the reliability by enabling constant health checks. DC voltages are monitored at key points within the system. DC current from each power supply is monitored and reported to the system controller. The CPU power control will shut down any overheating CPU without shutting down the system.

2.3 Availability

For organizations whose goal is to make information instantly available to users across the enterprise, high levels of availability are essential. This is especially true for a large shared resource system such as the Sun Fire system.

2.3.1 Availability Goals for the Sun Fire System

The Reliability, Availability, and Serviceability (RAS) goals for the Sun Fire system are to protect the integrity of the customers data and to maximize availability. The focus is on three areas:

- Problem detection and isolation—knowing what went wrong and ensuring that the problem is not propagated
- Tolerance and recovery—absorbing abnormal system behavior and fixing it, or dynamically circumventing it
- Redundancy—replicating critical components

To ensure data integrity at the hardware level, all data is error correction code (ECC) protected, and control buses are protected by parity checks out to the data on the disks. These checks ensure that errors are contained.

For tolerance to errors, resilience capabilities are designed into the Sun Fire system to ensure that the system continues to operate, even in a degraded mode. Because it is a symmetrical multiprocessing system, the Sun Fire system can function with one or more processors disabled. In recovering from a problem, the system is checked quickly to determine the fault and to ensure minimum downtime. The system can be configured with redundant hardware to reduce downtime.

2.3.2 High Availability Capabilities of the Sun Fire System

The Sun Fire system capabilities raise its availability from the normal commercial category to the *high availability* category. These capabilities are grouped as follows:

- *Fault-tolerant capabilities*—Any single point of failure is entirely transparent to users. Users see no loss of performance or capability in the specific areas of the system that are fault tolerant.

- *Resiliency capabilities*—These capabilities enable processing and data access to continue in spite of a failure, possibly with reduced resources. These capabilities usually require that you reboot your system.
- *Serviceability capabilities*—These capabilities lower or eliminate the repair time when a failure occurs.

2.3.2.1 Cooling

The Sun Fire system has redundant cooling. If one fan fails, the remaining fans automatically increase their speed, thereby enabling the system to continue to operate, even at the maximum specified ambient. Therefore, operation need not be suspended when a fan fails. You can replace a fan while the system is operating, again without any adverse impact on the availability metric. Of course, the Sun Fire system has comprehensive and fail-safe temperature monitoring to ensure that there is no over-temperature stressing of components in the event of a cooling failure.

2.3.2.2 AC Power Switching

AC power is supplied to the Sun Fire system through up to four independent, 30-ampere, single-phase Redundant Transfer Switches (RTS). Each RTS module carries power to two or three 2,200-watt bulk DC power supplies.

The AC connections must be controlled by separate customer circuit breakers, and can be on isolated power grids if a high level of availability is required. Optionally, third-party battery backup power can be used to provide AC power in the event of utility failure.

2.3.2.3 ECC

On the Sun Fire system, data errors are detected, corrected, and/or reported by the data buffer on behalf of its associated processor. Additionally, data errors passing through the interconnection will be detected and will cause a record stop condition for the ASICs. The ASICs detect and initiate this condition. These history buffers and record stop-condition bits can then be read and used by offline diagnostics.

2.3.3 Resiliency Capabilities

Resiliency capabilities enable processing and data access to continue in spite of a failure, possibly with reduced resources. These capabilities usually require that you reboot the system, and this is counted as repair time in the availability equation.

2.3.3.1 DC Power

The Sun Fire logic DC power system is modular at the system board level. Bulk 56-VDC is supplied through a circuit protector to each system board. This 56 volts is converted through several small DC-to-DC converters to the specific low voltages needed on the board. Failure of a DC-to-DC converter affects only that particular system board. You need to configure only as many bulk DC power supplies as are needed for the particular system configuration. The standard redundant configurations are three DC power supplies for up to three system boards and six DC power supplies for up to six system boards on the Sun Fire E6900 system.

2.3.3.2 Logic Boards

The System Controller board contains the system controller interface as well as the clock source and the emergency shutdown logic. Optionally, you can configure two System Controller boards in the system for redundancy.

The Repeater, CPU/Memory boards, and the I/O subsystems hold the DC-to-DC converters that power the address repeater, the system data controller, the system data crossbar, and all other ASICs. If one Repeater board fails, the system will continue to operate in a degraded mode, which includes two of the four address buses and data buses.

2.3.3.3 Processor

If you have a failure of an UltraSPARC-s400 processor, the dual data switch, the external cache SRAMs, or the associated support ASICs, the failed processor can be isolated from the remainder of the system by a power-on self-test (POST) configuration step. As long as there is at least one functioning processor available in the configuration, the system can operate.

2.3.3.4 Memory

When POST completes testing the memory subsystem, any faulty banks of memory will be identified. POST can then reconfigure the memory configuration using only reliable memory banks, taking advantage of the highly configurable nature of the address-match logic in the memory controller.

2.3.4 Redundant Components

Both the customer mean time between failure and the customer availability measures of the system are enhanced by the Sun Fire system's capability to configure redundant components. There are no components in the system that cannot be configured redundantly if the customer desires. Each system board is capable of independent operation. The Sun Fire system is built with multiple system boards and is inherently capable of operating when only a subset of the configured boards is functional.

In addition to the basic system boards, redundant configurable components include:

- System Controller boards (version 2)
- Repeater boards
- AC power switching
- DC power supplies
- Peripheral controllers and channels

You can configure systems with multiple connections to the peripheral devices, enabling redundant controllers and channels. Software maintains the multiple paths and can switch to an alternate path on the failure of the primary path.

The system controller is controlled through a console interface workstation. Redundant system controllers and interfaces can be configured if the customer desires.

2.4 Serviceability Capabilities

To reduce repair time, the Sun Fire system has been designed with a number of maintenance capabilities and aids. These are used by the Sun Fire system administrator and by the service provider.

Several capabilities enable service to be performed without forcing scheduled downtime. Failing components are identified in the failure logs in such a way that the field-replaceable unit (FRU) is clearly identified. All boards and power supplies in a properly configured system can be removed and replaced during system operation without scheduled downtime.

2.4.1 Mechanical Serviceability

Connectors are keyed so that boards cannot be installed upside down. No special tools are required to access the inside of the system. This is because all voltages within the cabinet are considered extra-low voltages (ELVs) as defined by applicable safety agencies.

No jumpers are required for configuration of the Sun Fire system. This makes for a much easier installation of new and/or upgraded system components. There are no slot dependencies other than the special slots required for the System Controller and Repeater boards.

The Sun Fire system cooling-system design includes capabilities that provide strength in the area of RAS. Standard proven parts and components are used wherever possible. FRUs and subassemblies are designed for quick and easy replacement with minimal use of tools required.

2.4.1.1 DC Power Supply Removal and Replacement

56-VDC power supplies can be hot-swapped with no interruption to the system. This assumes that the system is configured from the factory for power supply redundancy.

2.4.1.2 Fan Tray Removal and Replacement

If a fan fails, the remaining working fans are set to high-speed operation by the system controller to compensate for the reduced airflow. The system is designed to operate normally under these conditions until the failed fan assembly is serviced. The fan trays can be hot-swapped with no interruption to the system.

2.4.1.3 Domain Isolation

The Sun Fire system has an interconnect domain facility that enables the system boards to be assigned to separate domains. For example, one domain can do production while a second domain experimentally runs the next revision of the operating system or exercises a suspected bad board with production-type work.

2.4.2 Nonconcurrent Service

Nonconcurrent service requires the entire system to be powered off.

2.4.3 Remote Service

Every System Controller board has remote access capability that enables remote login to the system controller. Through this remote connection, all system controller diagnostics are accessible. You can run diagnostics remotely or locally on deconfigured system boards while the operating system is running on the other system boards.

Hardware Overview

The Sun Fire systems are a family of symmetrical shared-memory multiprocessors (SMPs).

You can view the Sun Fire systems at several functions of detail:

- Standard operation (simple SMP and OS functions)
- Interconnect (details of interconnect for OS boot and RAS features)
- Console bus interconnect (details of how the System Controller board controls the system)

3.1 Standard Operation

The standard operation is simply that of an SMP running an operating system with standard functions. It consists of CPU/Memory devices and I/O devices connected through an interconnect bus. Although the data interconnect is actually a crossbar switch, it is logically a bus. This is illustrated in FIGURE 3-1.

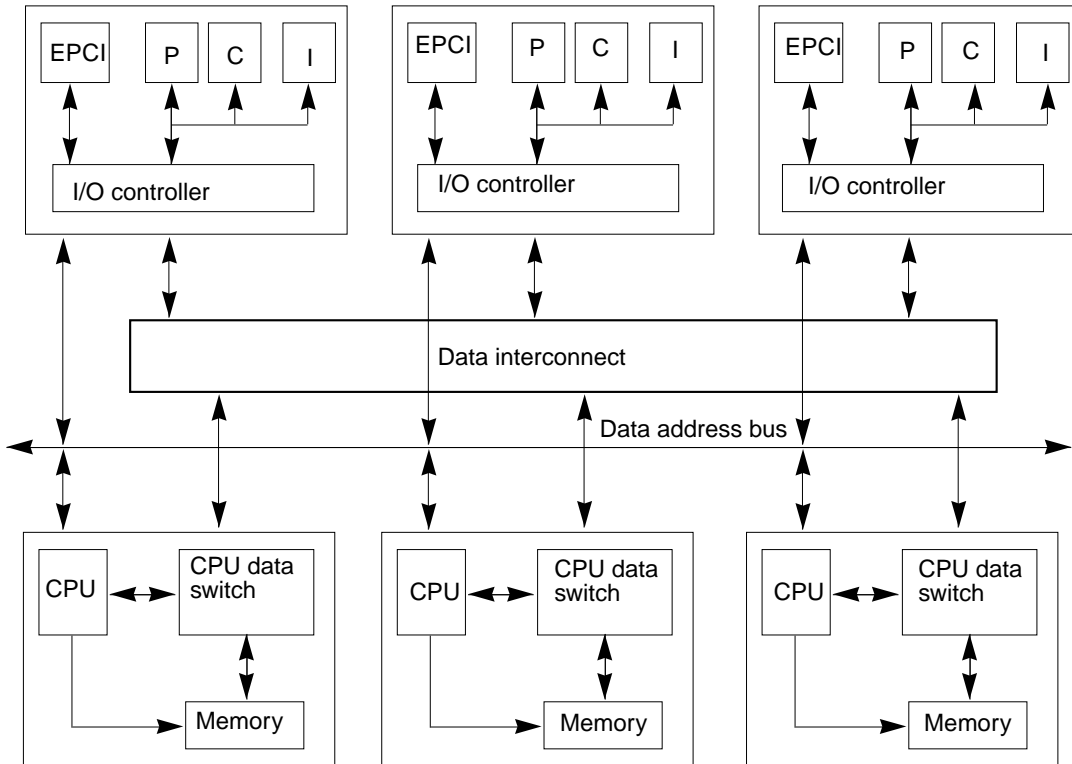


FIGURE 3-1 Standard Operation of the Sun Fire E6900/E4900 Systems

3.2 Data Interconnect

Although the standard operation of the Sun Fire system is that of a simple “bus-like” interconnect, it is actually a point-to-point switched interconnect, with two levels of repeaters or switches.

The switch is capable of complex functions such as:

- Dividing the system into completely isolated partitions
- Dividing the partition into logically isolated domains

To boot the operating system and to exercise the functions listed above, the system controller must be aware of the logical structure of the switch interconnect.

The Sun Fire E6900 system has six slots for CPU/Memory boards. The Sun Fire E4900 system has three slots for CPU/Memory boards. Each CPU/Memory board has up to four UltraSPARC s400 CPUs. The CPU also includes a memory controller, and each CPU can support one memory bank with up to eight DIMMs.

The Sun Fire E6900 system has four bays for the I/O assemblies. Two bays are included in the Sun Fire E4900 system for I/O assemblies. The Sun Fire E6900/E4900 systems support PCI. Each PCI I/O assembly has two I/O controllers, each with one 66-MHz PCI bus and one 33-MHz PCI bus.

The Sun Fire E6900 system is designed to greatly improve reliability, serviceability, and availability (RAS) over previous generations of systems. The Sun Fire system is designed to be able to recover from any hardware failure. Some failure recovery will not impact users (for example, a power supply failure) if the system is configured for redundant power supplies. Some failure recovery (for example, a CPU failure) will require a reboot, and will impact users, but a properly configured system will always be able to recover from any hardware failure.

The address path and data path are treated in slightly different ways. The address path has two completely redundant repeaters. A complete address repeater path requires two Repeater boards, as the Address Repeater (AR) function is bit-sliced across two ARs. On the Sun Fire E6900 system, the data path is bit-sliced across all four Repeater boards for standard operation. Optionally, a single pair of Repeater boards can be used in “double-pumped” mode, so that full functionality, although with lowered data bandwidth, is retained.

The Repeater boards have active devices. Because centerplanes are relatively hard to service, the Sun Fire E6900/E4900 systems were designed so that no active devices are present on the centerplane.

3.3 Console Bus Interconnect

The console bus enables the system controllers (SC) to read and write registers throughout the rest of the system. Only one of the two SCs can be master on the console bus at a time. Each system controller is connected to a console bus hub (CBH), and the two CBHs arbitrate for the use of the console bus.

Sun Fire System Components

The Sun Fire E6900/E4900 systems share many of the same components. These components are the CPU/Memory boards, the I/O assemblies, the Repeater boards, and the System Controller boards.

4.1 CPU/Memory Board

The CPU/Memory board is the same across the Sun Fire E6900/E4900 systems. This board supports up to four UltraSPARC s400 CPU modules, and eight banks of memory, two banks per CPU with four DIMM sockets per bank, for a total of 32 DIMMs.

All DIMMs must be the same capacity and size within a bank, and must not be intermixed on a board.

4.2 I/O Assemblies

The Sun Fire E6900/E4900 systems support PCI I/O devices.

4.2.1 PCI I/O

The I/O assemblies are logically and physically the same for the Sun Fire E6900/E4900 systems. The basic PCI I/O assembly will have six slots for standard PCI (33-MHz) device boards, plus two slots for PCI-66 (66-MHz) device boards.

4.3 Repeater Board

The Sun Fire E6900/E4900 systems are designed to be repaired and upgraded easier and faster than previous systems. This is due to the placement of active ASICs mounted on the Repeater boards. With two Repeater boards installed in the system, an alternate path is available through the second board, if one board fails.

The Repeater boards provide two functions: redundancy for reliability, and a higher bandwidth. The system can operate with only one Repeater board. The Repeater board acts as a switch and connects multiple CPU/Memory and I/O boards together. The three components are the Address Repeater (AR), the Sun Fire system data controller (SDC), and the data crossbar (DX).

In standard operation, the Sun Fire E6900 system has four Repeater boards which are used to route ten buses (six CPU and four I/O). If one of the Repeater boards fails, the system can continue to operate in a degraded mode with one pair of adjacent Repeater boards. The data width is cut in half and the two Repeater boards will route the ten buses.

Because the Sun Fire E4900 system supports only two Repeater boards, the two Repeater boards operate together to route five buses (three CPU and two I/O). If one of the Repeater boards fails, the data width is cut in half and one Repeater board can route the five buses.

4.4 System Controller Board, Version 2 With Enhanced Memory

The system controller board contains the system clock and a service processor. The processor on the board is a microSPARC IIep with its own POST/OBP flash PROM and memory. The processor also has a 33-MHz PCI bus with two devices on it. The system controller board also has a 10/100BASE-T Ethernet connection and an Ebus interface for a variety of devices. These include a TOD/NVRAM device, flash PROM for extra NVRAM space, a large flash PROM to hold the OS code, and one 16552 dual serial port device.

One System Controller board is required per system. The Sun Fire E6900 comes with two system controllers for redundancy. You can install a second System Controller board in the Sun Fire E4900 System if desired. With two System Controller boards, if one System Controller board fails, the other System Controller board can take control of the system without causing a disruption in the main system operation.

The System Controller board performs the following main functions:

- Sets up the system and coordinates the boot process
- Generates system clocks
- Monitors the environmental sensors throughout the system
- Analyzes and logs error messages and takes corrective action
- Sets up the system partitions and domains
- Provides the system console functionality
- Provides virtual domain key switches
- Provides network Solaris software console for each domain
- Provides virtual time of day for each domain

The system controller provides five ports: domain A console, domain B console, domain C console, domain D console, and the system controller shell. The system controller shell provides the following:

- Configuration control
- Environmental status
- Ability to reconfigure domains
- Ability to power on and off power grids
- Ability to change the system controller password
- Other generic system controller functions

The system controller software sequences the booting of the system by:

- Configuring hardware
- Setting up domains
- Powering on and off components (such as system boards, power supplies, and fans)
- Testing components
- Building the domains

The system controller software provides tools for changing the configuration of the system, and it also logs errors.

4.4.1 Redundant System Controllers

With two System Controller boards installed in Sun Fire E6900/E4900 systems, the second board is a redundant System Controller board. Each System Controller board can check the health and status of the other System Controller board. If the main System Controller board (SC0) fails, the redundant System Controller board (SC1) takes over operations without causing a disruption of the system operation.

4.4.2 Virtual Domain Key Switches

The system controller provides a virtual key switch for each domain. The `key switch` command controls the position of the virtual key switch for each domain.

4.4.3 Solaris Console

The system controller provides a Solaris software console for each domain. The Ethernet or serial port of the System Controller board is the physical connection for the Solaris software console. The serial port can support only one console at a time. However, the Ethernet port can support many consoles simultaneously. The system controller multiplexes these physical connections to provide console services for each domain and for the system controller.

4.4.4 Virtual Time Of Day

The Sun Fire system TOD/NVRAM chip is located on the System Controller board. The system controller multiplexes the physical TOD chip to provide TOD services for each domain and for the System Controller board. The system controller also provides for synchronizing the TOD between the main System Controller board and the redundant System Controller board.

4.4.5 Environmental Monitoring

The Sun Fire system has a large number of sensors that monitor temperature, voltage, and current. The system controller polls these devices periodically. If thresholds are exceeded, the system controller shuts down various components to prevent damage.

