Sun Microsystems, Inc.
2550 Garcia Avenue
Mountain View, CA 94043
415 960-1300
FAX 415 969-9131

For U.S. Sales Office locations, call:
800 821-4643
In California:
800 821-4642

Australia: (02) 413 2666
Belgium: 32-2-759 5925
Canada: 416 477-6745
Finland: 358-0-502 27 00
France: (1) 30 67 50 00
Germany: (0) 89-46 00 8-0
Hong Kong: 852 802 4188
Italy: 039 60551
Japan: (03) 221-7021
Korea: 822-563-8700
Latin America: 415 688-9464
The Netherlands: 033 501234
New Zealand: (04) 499 2344
Nordic Countries: +46 (0) 8 623 90 00
PRC: 861-831-5568
Singapore: 224 3388
Spain: (91) 5551648
Switzerland: (1) 825 71 11
Taiwan: 2-514-0567
UK: 0276 20444

Elsewhere in the world,
call Corporate Headquarters:
415 960-1300
Intercontinental Sales: 415 688-9000

Conclusion 3

The SPARCcenter 2000 and SPARCserver 1000 systems provide highly scalable platforms for use in intense compute environments, in high transaction rate environments, and in configurations supporting large numbers of users. The ability to do this results from a modular design that can be thought of as a set of systems communicating and sharing resources over a common bus.

Coupled with this design is a set of RAS features that ensure system integrity. These features begin with chip level specifications, are seen in the manufacturing testing and acceptance, and continue on in the field with diagnosis tools, reconfiguration and redundancy.

The RAS strategy built into the SPARCcenter 2000 and SPARCserver 1000 systems guard data, protect investment, and provide a consistent path into future data processing needs.

Should any non-functional unit be found, *prtdiag* identifies the appropriate field replaceable unit (FRU). The unique aspect of *prtdiag* is its ability to run while Solaris is operational -- yet still provide POST level information. See Figure 2-2 below.

```
Failed Field Replaceable Units (FRU) in System:
===============================================

cpu-unit unavailable on System Board #1
Failed Field Replaceable Unit is SuperSPARC Module B

Analysis of most recent System Watchdog:
========================================
Log Date: Mon Mar 15 13:45:33 1993

Analysis for Board 1
--------------------
BW0 (CPU B)
  Client Device Error, Internal Error(s) = XPE
```

Figure 2-2 Solaris *prtdiag* example showing an error.

```
prtdiag -v
System Configuration:  Sun Microsystems  sun4d SPARCserver 1000
System clock frequency: 40 MHz
Memory size: 128Mb
Number of XDBuses: 1
====== CPU Units [MHz] ======== Memory Groups [MB] =============
             A      B                                grp0    grp1    grp2    grp3
                 --      --                                      ----       -
---      ----       ----
Board0: 40     40                                  32      32        0         0
Board1: 40     40                                  32      32        0         0
======================SBus Cards============================
Board0:         0: dma/esp(scsi)
                       1: dma/esp(scsi)
                            lebuffer/le(network)
                       2: <empty>
                       3: cgsix
Board1:         0: dma/esp(scsi)
                           lebuffer/le(network)
                       1: <empty>
                       2: <empty>
                       3: <empty>


No failures found in System
```

Figure 2-1 Solaris *prtdiag* example. Note no failures were found in the system.

POST is initiated from three different starting points. In Power On Reset, a DIAG switch permits two levels of testing to be run. If DIAG is set, the more rigorous level is run. In reboot, speed of reboot is emphasized. To promote this consideration, stored information from previous boots is consulted. In an error case, POST tries to recover from an unexpected error encountered since the last POST entry point initiation.

Solaris has a utility, called *prtdiag*, to provide board level diagnostics from a UNIX shell. *Prtdiag* provides a listing of the entire configuration including:

- Number, type, and location of processors
- Location and amount of memory
- SBus configuration

See Figure 2-1 which follows.

Using disk mirroring in conjunction with dual-mastered SCSI buses, a system can be protected from disk data loss due to any single point failure in either the system of the disk subsystems.

SPARCcenter / SPARCserver System Diagnostics RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
Power On Self Test (POST)		■			■	■	■		■
Solaris *prtdiag* On-Line Diagnostics							■	■	■

*Table 2-6***SPARCcenter / SPARCserver System Diagnostics RAS features**

2.6 *Self-Test and Diagnostics RAS Features*

The Power On Self Test (POST) is the key mechanism for guaranteeing a correctly functioning system. See Table 2-6. POST is an orderly series of tests designed to locate failures at component, subsystem, and system level. Its objectives are to detect problems and to isolate them from other subsystems; to configure around failure points; and to maximize the number of usable I/O components.

Viewed at a high level, POST consists of three test phases: board level, system level and reconfiguration. In the first phase, each boards' components are tested, the functional groupings are tested and finally, the board is deemed operational or broken. Next, the system level testing tests the interconnection pathways of the system. The goal is to ensure that functioning board level elements can work together. Memory is tested during this phase.

Finally in the reconfiguration phase, the information from the first two phases is used to configure the working system. In order to do this, some components might be bypassed because of failures detected during the testing phases.

SPARCcenter / SPARCserver System Disk RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
(Optional) Online DiskSuite disk mirroring		■	■	■	■	■			
Standard disks with high MTBF	■								
Dual-master SCSIadapters		■		■		■			

Table 2-5

2.5 I/O and Disk Subsystem RAS Features

All I/O takes place through an I/O unit. An I/O unit consists of IOCs (I/O Cache), an SBI (SBus Interface), an external page table for DVMA and multiple SBus peripherals. For example, the SPARCcenter 2000 has two IOCs per I/O unit -- each of which can control any or all of the SBus slots. The SPARCserver has one IOC per I/O unit. RAS features for the I/O and disk subsystem are summarized in Table 2-5.

A SCSI bus can be dual-mastered. This means that the bus is connected to two different SCSI controllers -- each of which can take over the full burden of controlling the bus and its string of devices. These two adapters can be on the same system or on different systems. One adapter is the active master and the other is the hot standby. Either adapter can be the master, but only one can be the master at a time. This feature offers both flexibility in configuration and availability.

Disk mirroring is available with the optional Online DiskSuite software. With disk mirroring, disk data is protected from destructive disk crashes and from adapter and bus failures. By configuring a mirror set with each disk on a different adapter, the data will be available even if a single bus or a single adapter fails.

SPARCcenter / SPARCserver System Bus RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
Bus parity protection & bus time-outs		■	■						
SPARCcenter 2000 dual XDBuses				■		■			

Table 2-4 **SPARCcenter / SPARCserver System Bus RAS features**

- **XDBus** -- Data and control lines are parity protected. Eight bits of parity accompany the 64-bit data path organized as byte parity. One bit of parity accompanies the 5 arbitration request lines. One bit of parity accompanies the 6 arbitration grant lines.
- **XBus**[6] -- Four bits of parity accompany 64 bits of data, organized as half-word parity.
- **SBus** -- One bit of parity is provided for the data lines.
- **SCSI** is implemented with byte parity.

In addition to parity protection, a bus time-out mechanism is used to reduce the effects of a bus hang. SBus time-outs are maintained by the SBus interface. Having bus time-out constants available makes it possible to recover automatically in the event of a bus hang.

6. The XBus is a packet-switched bus between the BusWatchers on the XDBus and the SuperCache on the SuperSPARCmodule.

SPARCcenter / SPARCserver System Memory RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
ECC memory		■	■						
DRAM failure isolated		■	■			■			
Errors logged							■		■
Reconfiguration isolation					■	■			
Modular components								■	■
Reliable SIMMs, minimize # of pins	■								

*Table 2-3***SPARCcenter / SPARCserver System Memory RAS features**

2.4 Bus Complex RAS Features

Error correction on a bus would be expensive to manufacture. Information on a bus is transient and can be regenerated. As a result, the SPARCcenter 2000 and SPARCserver 1000 systems implement bus parity on all bus traffic -- detecting all single bit bus errors. Furthermore, the SPARCcenter 2000 has two XDBuses and can run with only one of them operational. See Table 2-4 for a summary of the SPARCcenter 2000 and SPARCserver 1000 bus complex RAS features.

SuperCache SRAM and Bus Watcher tag RAM are parity protected for reliability.

To reduce the likelihood of failure and to help with diagnosis, the SuperSPARC and SuperCache chips come with extensive self test logic. Their Built In Self Test (BIST), using a pseudo-random pattern applied at system clock speed, gives a greater that 80% detection rate for single-stuck-at faults in combinational logic. (A stuck-at fault is a bit that is stuck in either a one or zero position.) ASICs used in these systems conform to the JTAG standard (IEEE 1149.1).[5]

2.3 Memory Subsystem RAS Features

The SPARCcenter 2000 and SPARCserver 1000 systems have many RAS features. These are summarized in Table 2-3.

Dynamic Random Access Memory (DRAM) is protected by an Error Correcting Code (ECC) that operates in 64-bit segments. It corrects all single bit errors and detects all double bit errors. It also detects all triple-bit and quadruple-bit errors that occur within the same nibble of a 64-bit segment.

Even if an entire DRAM fails, the ECC protection and memory architecture maintain correct data due to the memory crossbar and built-in ECC. The failure of a DRAM will cause no more than a single bit error in any one 64-bit memory word.

ECC memory protection, previously presented as a reliability feature, can also be regarded as an availability feature. Correctable errors are repaired transparently; user services are not affected.

4. SPARCcenter 2000 only.

5. JTAG implements a boundary scan testing standard that makes it possible to locate point failure on ICs or at board level integration. This information is used in higher level test sequences that test integration and interaction of component functionality.

SPARCcenter / SPARCserver System Processor RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
Symmetric multiprocessing				■				■	
Cache SRAM & Bus Watcher tag RAM parity protected	■	■	■						
Built In Self Test (BIST)					■	■	■		■

*Table 2-2***SPARCcenter / SPARCserver System Processor RAS features**

2.2 Processor RAS Features

The SPARCcenter 2000 and SPARCserver 1000 processors are packaged on a daughter card called the SPARCmodule. The SPARCmodule contains highly integrated VLSI chips that include the SPARC processor, the SuperCache and the bus interfaces. This daughter card is a field replaceable unit. Each system supports two SPARCmodules per system board providing for redundant processors. The RAS features of the SPARCcenter 2000 and SPARCserver 1000 processor modules are summarized in Table 2-2.

If a failure should occur in a SuperSPARC or SuperCache, the failed module can be deactivated so that it will not interfere with system operation. The deactivation takes place during a reboot cycle initiated by an error interrupt. Since the SPARCcenter 2000 and SPARCserver 1000 features fully symmetrical multiprocessing, loss of a processor module may degrade performance but full functionality will still be maintained.

The SPARCcenter 2000 and SPARCserver 1000 implement a bus watcher mechanism on the XDBus. These bus watchers intercept bus commands bound for addresses on the boards which they service. If a Bus Watcher (BW) ASIC fails, the processor module serviced by the failed BW ASIC is de-configured[4], leaving all memory and I/O intact.

In some cases, environmental failures cannot be tolerated and will trigger a system shutdown. Out of range temperature or voltage, and persistent power are examples. These shutdowns initiate without warning and start immediatly so that component damage will be avoided. Other, less critical, enviromental irregularities are reported to the system console as warnings. The system administrator may then choose to act on the warning.

An Uninterruptable Power Supply (UPS) is also available as an optional feature. The SPARCcenter 2000 and SPARCserver 1000 systems can be configured with any of a variety of third party UPSs.[2]

2.1.2 LED Notifications

These systems have a variety of LEDs for at-a-glance checking of the system status. There are three LEDs on the front panel for monitoring DC power, single event error, and boot status.

LEDs on the control board monitor power supply levels and provide status information on the XDBus, whether a service processor is present and whether system reset is active.

Under special circumstances, the eight LEDs on every system board indicate any failed system boards after boot is completed.

2.1.3 Jumperless Configurations

Efforts have been taken to make the SPARCcenter 2000 and SPARCserver 1000 systems service cycle foolproof. This was done by eliminating meaningless or illegal configuration options. For instance, all connections are keyed so that boards and cables cannot be connected improperly. There are no slot dependencies on the buses so boards can be installed more easily. There are no jumper configurations to worry about.[3]

---

1. Single cycle dropouts are defined as the loss of a cycle in the 60 Hz AC power

---

2. Sun makes no recommendations regarding which to use.

---

3. The one exception to this is the SBus SCSI Ethernet controller (FSBE board) in which there is a jumper to enable or disable the tpe-link-test.

SPARCcenter / SPARCserver System Packaging RAS features	Reduce probability of error	Tolerate or recover from error	Ensure data integrity	Redundant components	Failure detection	Reconfiguration	Identify problem	Reduce illegal configurations	Reduce service cycle
Design for worst case environmental conditions: Temperature, voltage	■								
Modular design, reduce number of components	■							■	■
Reliable VLSI components	■								■
Jumperless configuration & no slot dependencies								■	■
Control panel and lights on system board							■		■
Field Replaceable Units (FRU)								■	■
Environmental sensors			■				■		
GTL signals	■								

*Table 2-1***SPARCcenter / SPARCserver System Packaging RAS features**

2.1.1 Environmental Safeguards

The SPARCcenter 2000 and SPARCserver 1000 systems are designed to tolerate environmental fluctuations. Their operating temperature range is 0 - 40C. The SPARCcenter 2000 and SPARCserver 1000 systems are implemented using low-swing GTL logic reducing noise in the signaling.

Power failure sensing features detect irregularities in both AC and DC power. AC power can tolerate brownouts of 160VAC for at least 15 minutes. It can also tolerate single cycle dropouts without failure.[1] If a persistent power failure is detected, DC power will continue for 5 ms.

SPARCcenter 2000 & SPARCserver 1000 RAS

2

The SPARCcenter 2000 and SPARCserver 1000 systems provide increased levels of reliability, availability and serviceability not typically seen in micro-processor based systems. Reliability is achieved through reducing probability of error, tolerating or recovering from errors and ensuring data integrity in spite of failures. Availability is provided through redundant components, on-line and boot-time failure detection, and failover or reconfiguration. Serviceability elements include easy identification of problems and the shortening of the service cycle.

The SPARCcenter 2000 and SPARCserver 1000 systems were designed to meet the goals of RAS: there are component level RAS features, module level RAS features and system level RAS features. For example, component RAS features include the use of high reliability SIMMs and VLSI chips. The module level RAS features use replication, such as multiple memory controllers, I/O controllers and processors. The system level features include environmental sensors, diagnostic registers and boot-time reconfiguration. The high end of the family, the SPARCcenter 2000, adds additional RAS features with the inclusion of a second XDBus backplane bus. This additional bus not only provides important availability features, but contributes to the outstanding performance of this machine. Additionally, there are many optional features, such as UPS and disk mirroring that add even more to the RAS feature set.

2.1 System Packaging RAS Features

The RAS features of the SPARCcenter 2000 and SPARCserver 1000 systems are summarized in Table 2-1.

1.4 Reliability, Availability and Serviceability

Reliability, availability and serviceability each deal with separate facets of assuring continuous operation. Each faces a distinct set of problems and provides solutions geared to those problems. Yet each supports the other and cannot be used separately. RAS, in reality, is a single set of features that provide more benefit than could be achieved by the simple sum of its individual components.

- There needs to be failure detection mechanisms and a procedure for orderly shutdown in emergency circumstances.
- There needs to be a comprehensive set of self-test procedures capable of locating failures at various levels of system integration.
- At each level, there must be backup resources available and a way to map it in to replace the failed unit.
- Finally, the system must boot itself to a level from which the users can easily return to the work under way at the time of the failure.

The key to availability is the systems' failure detection and failure correction mechanisms. Timely failure detection can prevent damage being done to the system itself or to data maintained on the system. Also, failure detection can be used to trigger failure recovery handlers -- so that it may be possible to recover from or correct failures under some circumstances.

If a problem should occur that is outside the scope of the reliability and availability features, the next step is to schedule service -- or the serviceability aspect of RAS.

1.3 Serviceability

All systems require service -- either scheduled or emergency. The goal of serviceability is to minimize the Mean Time To Repair (MTTR) -- the average time between the point of failure recognition and the return of the system to full functional operation.

There are four recognizable approaches to serviceability:

- Make it easy to identify a problem and to locate its source. This can be done by monitoring error logs and periodically running diagnostic programs.
- Make it difficult to improperly configure replacement components -- to reduce or eliminate illegal configurations.
- Minimize the number of serviceable components.
- Simplify the service cycle.

Failure detection and isolation together with a rapid service cycle answer the immediate question of serviceability. Use of field replaceable units (FRU) is a powerful strategy toward this end. With FRUs, even if only a single component on a module has failed, the entire module is replaced. The module with the faulty component is then returned to a service center where component-level servicing is done. There, test equipment can determine if the failure was due to a malfunctioning component or whether other conditions on the module caused the component to fail. Technicians with special equipment can make and verify repairs. FRUs make precise repairs easier while minimizing the need to stock a large spares inventory.

1

1.1 Reliability

Reliability is defined as reducing the likelihood of problems occurring. Reliability is achieved by a combination of engineering design decisions and manufacturing precautions. A common metric used to quantify reliability is Mean Time Between Failure (MTBF).

A large body of information and considerable expertise has grown up around failure anticipation and prevention. The first step in this process is to understand how a system will be used and to identify circumstances that may lead to failure -- including isolated events and systems interactions. The engineering design principle is to anticipate problems and design with them in mind; to consider effects at discrete levels and at systems levels.

No matter how thorough a reliability design, a failure potential always exists. This raises the question of availability.

1.2 Availability

Where reliability aims to prevent failures, availability protects users in the event of a failure. Availability means guaranteeing data integrity across system failures and assuring a high rate of access to system services-- that is, minimize the time that the system is unavailable.

At its extreme, availability consists of a fully fault-tolerant system that transparently re-configures itself in the event of any component failure -- to achieve constant availability of the system and its resources. It is a very costly solution that is justified in specialized circumstances.

More applicable -- to today's right-sizing efforts -- is an economical approach to availability that still offers a high degree of access. This approach uses redundancy, error correction, on-line diagnostics and auto re-configuration. When failure occurs, the system falls back to a reboot sequence where redundant components can assume the responsibility of their failed counterparts. Combining fast rebooting with auto re-configuration improves availability.

To offer this economical approach to availability:

- A system must built from redundant, modular, isolatable components.
- Data paths must be protected.
- Data stores must employ error correction.

Introduction to Reliability, Availability and Serviceability

1

As computers have evolved into essential elements of work environments, so has the requirement that computer-based systems provide regular service on demand. This is true of all types of computer equipment -- from single user PCs to enterprise systems.

Regardless of the type of equipment, confidence in continuing operation is required. However, confidence does not come without a price and the level provided needs to be gauged against a degree of acceptable risk. To PC users, a periodic backup of disk data may be a satisfactory compromise of risk and cost. For large server systems serving many clients responsible for volumes of key information, the degree of acceptable risk is much lower.

As organizations look to right-size, they are looking for reasonable trade-offs of cost and availability. To date, low cost servers based on microprocessor technology have not provided high availability. They have been difficult to maintain, with no provision for self-checking, self-configuring or data protection features.

The SPARCserver 1000 and SPARCcenter 2000 systems provide a comprehensive system solution for departments and enterprises. They are capable of serving a large number of users that require intensive compute power and high I/O bandwidth. With these systems, Sun provides an upgrade oriented, microprocessor based solution along with affordable reliability, availability and serviceability features. Referred to by the acronym RAS, these features combine to provide a base level of confidence.

Contents

Please
Recycle

Reliability, Availability, and Serviceability in the SPARCcenter 2000 and the SPARCserver 1000

Sun Microsystems Computer Corporation
2550 Garcia Avenue
Mountain View, CA 94043
U.S.A.

Part No: 8xx-xxxx-xx
Revision X, Month 1992