Ultra™ Enterprise™ 10000 Server: SunTrust™ Reliability, Availability, and Serviceability

Technical White Paper

Please
Recycle

Executive Summary

Today, computers running the UNIX™ operating system are deployed in virtually every market segment. UNIX has achieved widespread use in the commercial marketplace and is the cornerstone of systems running mission critical applications. For users and their applications, downtime can have a dramatic effect on revenues and customer satisfaction.

Sun Microsystems, Inc. is addressing the needs of businesses operating in today's network-based, global computing world, by providing computing platforms that are reliable, available for use when needed, and easy to diagnose and service with minimal interruption.

Sun's newest Ultra™ Enterprise™ server, the Ultra Enterprise 10000, sets new standards in reliability, availability, and serviceability in open systems. This high performance system includes the most comprehensive set of RAS features, designated SunTrust™, that are available in an open system today.

The Ultra Enterprise 10000 provides businesses with outstanding levels of service, through SunTrust's features:

- Complete configurable redundancy
- Alternate Paths to disk and network
- Mainframe-style Partitioning
- Assured system recovery
- Improved data integrity with ECC
- Intelligent monitoring and reporting

- “No-outage” servicing
- Sophisticated diagnostic support

Beginning in concept and design, innovative technologies, such as the Gigaplane-XB™ interconnect and Dynamic Reconfiguration, are implemented in support of superior reliability, high availability, and convenient servicing.

Sun is committed to products and services that provide highly reliable solutions to its customers. Combined with a global 7x24 service organization, reliable Solaris™ operating environment, Solstice™ system and network management tools, Sun’s Ultra Enterprise 10000 family servers offer the most trusted open systems solutions today.

Contents

Introduction 1

The role of UNIX-based computer systems has changed dramatically over the past decade. UNIX systems originated in a pure research environment, and quickly became widespread in colleges and universities. In these environments, high availability was not critical, since the systems typically supported small numbers of students and researchers, rather than business-critical applications.

UNIX provided a rich development environment, ran on a variety of platforms, and often even shipped with source code. As a result, users employed UNIX to study operating systems design or make modifications to suit their specific needs. The very nature of UNIX users and their requirements placed relatively limited demands on hardware platforms, such that they did not need to be highly reliable, easily serviceable, or provide high availability.

Today, computers running the UNIX operating system are widespread in the commercial marketplace. More and more, these systems are running applications considered to be business-critical, where downtime can have a dramatic impact on revenue.

Today, systems vendors must meet the uptime demands of such business-critical applications by providing computing platforms that are reliable, available for use when needed, and easy to diagnose and service.

For these reasons, Reliability, Availability and Serviceability (RAS) have become critical components in today's computing platforms, and as such, must be designed into the system from the very beginning. RAS capabilities are as vital as competitive performance.

1

Sun Microsystems, Inc. continues to develop and evolve RAS technologies with its new products. Each step in this evolution takes the best from previous products and adds refinements and new technologies to the next product family. Most recently, Sun's new Ultra Enterprise 10000 server, the latest member of the Ultra Enterprise server family, expands upon RAS features found in the rest of the family. It shares environmental monitoring, ECC data protection, and modular subsystem design with the rest of the Ultra Enterprise family. Enterprise 10000 servers also have several additional features, such as full hardware redundancy, fault tolerant power and cooling subsystems, automatic recovery, mainframe-style partitioning, full software support for "no-outage" component swapping and advanced system monitoring tools.

This paper is intended to provide a detailed description of the RAS features available with Enterprise 10000 servers. For information on how Sun delivers mission-critical service and robust application suites as part of its overall RAS strategy, please see the executive brief entitled *Sun RAS Solutions for Mission-Critical Computing*.

Reliability, Availability, and Serviceability Overview 2

Introduction

Customers are seeking increased levels of reliability in the systems they deploy, requiring that their data and applications be available whenever they are needed. To meet customer's requirements, platforms providing data storage and supporting applications must have features that add to their overall reliability, as well as be serviceable such that problems are relatively easy to diagnose and repair.

In this section, we will discuss reliability, availability, and serviceability—collectively called RAS—in generic terms. The remainder of this document will explore specific RAS features and technologies of the Enterprise 10000 systems in more detail.

Reliability

Reliability is the starting point for building increasingly available systems. Reliability is a function of the design, component selection, and manufacturing processes. It is possible to define reliability as anything that serves to reduce the probability that a failure will occur (thereby increasing the systems availability), and which serves to maintain data integrity, thus reducing the probability that bad data will flow undetected through the system.

Mean Time Between Failure, or MTBF, is considered an important metric with respect to measuring a systems reliability. However, there is currently no industry adopted standard for measuring MTBF, which makes the MTBF number for a given system or component of questionable use for comparison against other vendors.

Availability

Availability may be described as the time a particular resource, such as a system, application, or data, is accessible and usable. Obtaining increasing levels of availability starts at the core system design and extends to the overall data processing or application architecture.

To some extent, reliability drives availability. The more failure reduction built into a system, the more available it is likely to be. Additional features and technologies can be built into a system to increase levels of availability by reducing failure recovery times and speeding up problem diagnosis and repair.

Availability is typically measured as a percentage of total uptime or services available over the course of a year. Customer uptime requirements are almost always stated this way. For example, a 99.99 percent availability requirement translates to 52.8 minutes of downtime per year, where a 99.9 percent availability requirement means about 8 hours of downtime per year.

What's unclear is whether or not the downtime factor includes planned maintenance or just unscheduled outages, and if the downtime represents a single time slot or a total of some number of time slots through the year (e.g., 8 x 1 hour outages or 1 x 8 hour outage). Of course, such questions can only be answered on a case-by-case basis, with the customer assessing their business requirements and the cost of downtime for a particular application.

Serviceability

Serviceability defines the time it takes to isolate and repair a fault, or, more succinctly, the time it takes to restore a system to service following a failure. Mean Time To Repair, or MTTR, is considered an important metric when discussing the serviceability of a system or some component of the system, keeping in mind that this is a unit of time and does not factor in cost of service (which can be significant).

Consider the cost of a 7x24 (seven days a week, 24 hours a day) service contract versus 5x8 coverage. Typically, the cost for the additional coverage increases by about 40 percent. A system that recovers automatically from most failures can obviate the need for 24 hour service coverage, which may also reduce support staff requirements. The net result is a highly available system with a lower cost of ownership.

Over the last few years, Sun has established a consistent trend in delivering increasingly modular and serviceable systems. Ranging from the reduction in numbers of jumpers and module slot dependencies, to tighter integration and greater environmental tolerances, Sun has made the time required for the replacement of a failed module much shorter. These enhancements, coupled with improved diagnostic capabilities has significantly reduced the service cycle on systems, and simultaneously increased their reliability and availability.

RAS

Reliability, availability and serviceability combine to provide for the continuous operation of the system. Designing and implementing technologies into systems that contribute to RAS is the first step in providing a platform that runs virtually continuously. Even fatal system errors that bring the machine down can be circumvented without manual intervention so the system can reboot and return to service in a relatively short amount of time.

System uptime, which is often the metric used to evaluate RAS, is typically discussed in terms of what percentage of uptime the system provides over the course of a year. The following two tables provide definitions of terms used when discussing availability, with corresponding uptime metrics. Table 2-1 is the work of Jim Gray, an acknowledged expert in computers and transaction oriented systems. Table 2-2, comes from a very large telecommunications company.

Looking at both tables, very different uptime percentages are used for the same terms. For example, Jim Gray defines Fault Tolerant as providing 99.99 percent uptime, while the telecommunications company suggests it refers to 99.999 percent uptime. Hence, when discussing availability (and therefore RAS), it is important that uptime requirements be clearly stated to ensure that goals can be met.

System Type	Unavailability Minutes/Year	Availability (Percent)	Availability Class
Unmanaged	50,000	90	1
Managed	5,000	99	2
Well-Managed	500	99.9	3
Fault Tolerant	50	99.99	4
Highly Available	5	99.999	5
Very H.A.	0.5	99.9999	6
Ultra H.A.	0.05	99.99999	7

Table 2-1 Availability terms and metrics (Source: Jim Gray)

Class	Downtime/Year	Percentage
Conventional	3.5 days	99.000
Highly Available	8.5 hours	99.9
Fault Resilient	1 hour	99.99
Fault Tolerant	5 minutes	99.999

Table 2-2 Availability terms and metrics (Source: major telecommunications company)

With the Enterprise 10000, meeting clearly defined availability requirements is principally a matter of configuring for higher RAS. For example, base systems can be configured for higher RAS levels by adding a redundant control board, additional I/O controllers, mirroring disks, etc. Alternate Pathing and no-interruption servicing strategies that are unique to the Enterprise 10000, may also be employed. Enterprise 10000 servers can be configured to the extent that no single point of failure exists in the hardware configuration.

Beyond that, building highly available platforms can be accomplished by increasing the amount of system redundancy (e.g., a hot spare backup system), at additional cost. And software can be added to detect failures and initiate failover to backup hardware resources. All of these approaches increase RAS, but substantially expand system cost and administrative complexity. As with

any other critical business issue, RAS decisions must be made based on the cost of downtime versus the investment required to adequately safeguard against it.

RAS and Sun

Reliability, Serviceability and Availability, or RAS, define the features, technologies and metrics used to assess and measure a given system's ability to operate continuously and reduce service times. The reliability of a system defines the technologies and manufacturing processes implemented to reduce system failures and insure data integrity. The availability of a system defines accessibility to the data and applications supported by the system. And a system's serviceability provides for short service cycles when failures do occur.

Combined, RAS provides for near continuous system operation, minimizing both planned and unplanned outages.

The remainder of this document discusses the RAS features of the new Enterprise 10000 product and details how they apply to the generic definitions presented here.

2

Enterprise 10000 RAS Implementation and Features 3

Introduction

Ultra Enterprise 10000 servers share the RAS capabilities now available in the Enterprise 3000, 4000, 5000, and 6000 servers. But the Enterprise 10000 goes beyond those features to incorporate RAS capabilities that demonstrate that Sun leads the UNIX marketplace in understanding and delivering quality systems for enterprise-wide mission-critical computing. This chapter describes the design philosophies employed, key underlying technologies, and then details the SunTrust RAS features themselves.

What is SunTrust?

SunTrust™ is a family of capabilities on the Ultra Enterprise 10000 server that deliver reliability, availability and serviceability that is unprecedented in the open systems world. Drawing on experience in data centers worldwide, Sun has designed the Enterprise 10000 server to incorporate a number of features formerly known only to mainframes. These features dramatically improve the time between failures, deliver highly available services, and simplify servicing.

As Sun develops additional products, the SunTrust features will be implemented in next generation products, constantly improving levels of reliability. Sun's product line, as it has done in the past, will evolve to deliver ever-increasing levels of dependability and trust.

3

Enterprise 10000 SunTrust Design Philosophies

RAS begins with a solid design for RAS; it cannot be added after the fact. Thus, it is important to reflect on the strategies that underlie the Enterprise 10000 SunTrust capabilities. These philosophies have resulted in a system capable of handling mission-critical applications for the most demanding customers.

In planning for reliability, the main objectives were to eliminate or reduce the probability of errors; to correct errors without impacting availability, and to detect uncorrectable errors such that data integrity is not compromised.

To maximize availability, the focus was to detect and isolate errors, develop tolerance and recovery mechanisms, and implement high levels of redundancy.

And to enhance serviceability, the goals were to minimize impact on production operations, facilitate the servicing process, and provide ease of access and parts handling.

Key Technologies and Design Features

The Enterprise 10000 server incorporates several unique design features that form the basis for its innovative SunTrust RAS features. Three key architectural elements are the Gigaplane-XB™ crossbar interconnect, Dynamic Reconfiguration, and the integrated System Service Processor (SSP).

Gigaplane-XB™ Interconnect

The Gigaplane-XB crossbar is the system interconnect, providing communications among processors, memory and I/O. It implements a very high speed point-to-point connection among all system boards. While using the same Ultra Port Architecture as the rest of the Ultra server line, the Enterprise 10000 server's implementation of UPA is a crossbar rather than a bus. The characteristics of this form of interconnect enable higher system availability.

Using a sophisticated and reliable ASIC design, the Gigaplane-XB utilizes separate paths for addresses and data. The Global Data Router (GDR) and Global Address Bus (GAB) each provide for redundancy and graceful recovery in case of failure. The GDR routes 16 bytes in parallel, via two 8 byte routers that have independent power and clock signals. Should one router fail, when the system recovers, work can continue on the other router. Similarly, the GAB,

comprised of four independent busses, can operate as 3, 2, or 1 if a failure or even multiple failures occur. Hence, even in the very unlikely event of a centerplane failure, the system can stay in operation, in degraded mode, until a convenient time for service. The Gigaplane-XB technology also enables the dynamic reconfiguration and system partitioning features, described in sections that follow.

Dynamic Reconfiguration

Dynamic Reconfiguration is a key technology that permits a physical or logical restructuring of an Ultra Enterprise 10000, while the system is in active use. This is sometimes confused with hot-plug or hot-swap terminology, which accommodates hardware changes in inactive parts. Dynamic Reconfiguration provides full software support for components to be added or removed during systems operation, without impacting work.

Dynamic reconfiguration is fundamental to Enterprise 10000 processes for boot-time testing, and “no-outage” upgrades and servicing.

Integrated System Service Processor (SSP)

The integrated System Service Processor (SSP) is the central console that monitors the Enterprise 10000 system and enables the operator or service provider to perform management and maintenance procedures. The SSP is a workstation with specially developed software which communicates with the Enterprise 10000, to obtain detailed information on all aspects of the system’s status. The SSP manages the boot up process, monitors the status of the system, including its power and cooling, and provides the interface for the operator to manage system partitioning and dynamic reconfiguration. An important responsibility of the SSP is to protect the system from environmental conditions which can cause errors or damage to the system. The SSP also provides for recovery from fatal errors, automatically reconfiguring the system after testing for component faults.

The SSP includes a suite of software accessible to the system administrator and service personnel. The primary interface is the Hostview program, a graphical view of all aspects of system activity. Hostview is described in detail in a later section.

Connected via Ethernet to the Enterprise 10000 server, the SSP offers flexibility in the location of the console. It can be located in the same room, or in a central monitoring center, or even in a different building.

Enterprise 10000 RAS features

This section discusses the following RAS features of Ultra Enterprise 10000 systems:

- Complete configurable redundancy
- Alternate Paths to disk and network
- Mainframe-style Partitioning via Dynamic System Domains
- Assured system recovery
- Improved data integrity with ECC
- Intelligent monitoring and reporting
- “No-outage” servicing
- Sophisticated diagnostic support

Complete Configurable Redundancy

All Components Optionally Redundant

Each Ultra Enterprise 10000 server may be configured to have 100% hardware redundancy. Each critical link may be made redundant. Options are available for redundant control boards, centerplane support boards, system boards, disk storage, bulk power subsystems, bulk power supplies, peripheral controllers and channels, and the System Service Processor. Even the system interconnect has built-in redundancy. A fully redundant system will always recover from a system crash, utilizing stand-by components or operating in degraded mode. This is described more fully in the section entitled Assured System Recovery.

By making components optional, the Enterprise 10000 server remains cost effective, and enables customers to add redundancy where they believe it will be needed.

Fault Tolerant Cooling and Power

Ultra Enterprise 10000 servers feature redundant power and cooling in order to survive failures of power supplies and fans without interrupting the operation of the systems. The Enterprise 10000 power system implements a current share design, which effectively ties the outputs of several smaller power supplies together, making them appear to system components as one large power supply.

The Enterprise 10000 is built to withstand failure of the power and cooling subsystems without adverse effect. AC power is supplied to the Enterprise 10000 system through up to four independent 30 Amp single-phase plugs. Each AC plug carries power to a pair of 2,000 watt bulk power supplies. When N+1 power supplies are installed in the Enterprise 10000 server, a power supply failure does not require any intervention, as the remaining power supplies deliver adequate power for ongoing operations. The faulty power supply may be replaced without any impact on the running system.

Power subsystem redundancy extends to an optional duplicate power cord, which should connect to separate in-house circuits, and could also be connected to an alternate power grid or Universal Power Supply (UPS) unit.

The Enterprise 10000 is cooled with 16 fan trays, located above and below the processor boards. Each tray includes two fans, providing protection against a single fan failure. The fan trays are powered by four separate circuit breakers, so that there is no single point of failure in the cooling subsystem.

Two types of sensors, measuring temperature and airflow, are mounted on each system board. The airflow sensor detects that the fans are functioning and the temperature sensor verifies that the ambient air is within the defined tolerance. Should either sensor detect a problem, a message is sent to the System Service Processor, where the condition is logged, made known to the operator, and corrective action taken. If a fan has failed, the SSP will instruct the paired fan to rotate faster as a means of minimizing the effect of the failed fan. Should multiple fan failures occur, the SSP will initiate a graceful shutdown of the system, in order to avert any potential damage to the system from excessive temperature.

The redundant power and cooling operations described above happen automatically and dynamically. For RAS purposes, it is also important to automatically notify the system administrator and/or operator that an event has occurred which requires attention. A failed component still needs to be

replaced, even if the system continues to run. The Enterprise 10000 offers multiple means of communicating its condition; via the Hostview program operating on the SSP, email messaging, and SNMP messages, which may be intercepted and acted upon by enterprise-wide system management tools.

System Service Processor Redundancy

While the System Service Processor is not strictly required for operations of the Enterprise 10000 server, if it is not functioning, it cannot monitor and respond to adverse environmental conditions or component failure. An undetected problem, such as overtemperature, might result in an unnecessary system outage. Thus, if the need for availability demands, a system may be configured with a stand-by SSP, which can be brought into action if the master SSP experiences a sustained hard failure.

The spare SSP is kept up-to-date with the active SSP's settings and status, so that it may be brought on-line if necessary.

Alternate Paths to Disk and Network

While mirroring or other RAID techniques can guard against failure of a disk subsystem, to assure continuous access to data, you must also protect against the failure of its controller. Similarly, a redundant path to the network assures ongoing connectivity. The Alternate Pathing function, supplied with the Enterprise 10000 server, permits each server also to be configured against the failure of SBus disk arrays and network controllers.

Alternate Pathing is software that manages redundant I/O and network controllers. Using Alternate Pathing to oversee two independent controllers, attached to separate system boards, assures end-to-end protection. The system administrator is responsible for defining the primary and alternate path for each device. When a failure is detected, the system administrator activates the alternate path and deactivates the primary path. These actions take place while the system is in operation. For supported disk arrays, this can be accomplished without losing any data being transmitted. Once the alternate path is activated, the failed component may be swapped out.

Alternate Pathing is also a foundation for dynamic reconfiguration. It enables I/O operations to be redirected to the alternate path if the system board serving the primary path must be removed from the configuration.

Mainframe-style Partitioning via Dynamic System Domains

Many system outages result from software failures. To protect a mission-critical application, one would like to minimize the chances for intrusions by unrelated software applications. In many data centers, this is accomplished by establishing a separate server for each application. In a mainframe environment, partitions may be established for the same purposes. Similar isolation can be implemented using a unique feature of the Ultra Enterprise 10000 server, called Dynamic System Domains. This capability enables partitioning of the system, creating essentially “systems within a system.”

Each domain is a fully functional, self-contained server. Domains are made up of one or more system boards; each domain has its own CPU, memory, I/O and network resources, and runs its own instance of Solaris. Because each domain has its own resources, it is physically separate from other domains to a large degree. Certain components, such as the control board, are shared system-wide, however.

Most hardware errors originating within a domain will not affect any other domain; any component specific to one domain cannot affect any other domain. For example, if a processor or memory SIMM fails in a domain, no other domain will be affected. The only type of failures that can affect more than one domain are those generated by hardware components shared by all domains, such as the control board or Gigaplane-XB interconnect, in which case the entire system, thus all domains, can be affected.

Data is also isolated. In the Enterprise 10000 server, data travels among CPU, memory and I/O via the Gigaplane-XB interconnect. The data traffic is managed by a technique called domain filtering, which permits data to travel only as defined by the domain configuration. When domains are activated, data to and from other domains is not permitted; only traffic within the domain is permitted.

A domain creates complete software isolation, because there is a completely independent instance of the Solaris operating system. A domain is completely shielded from any type of software error from other domains, including errors generated by a Solaris panic condition or a program crash.

Because of this software and hardware independence, all software and most hardware errors in a domain will be confined to that domain only and will not affect the rest of the system.

Configuring multiple domains allows system administrators to safely test new software or isolate mission-critical applications with the assurance that each domain is protected from errors in any other domain. This error isolation dramatically increases the Enterprise 10000 server's overall reliability, availability, and functionality.

Assured System Recovery

Automatic System Reconfiguration (ASR)

All Enterprise X000 systems incorporate the ability to detect failed hardware components and boot around the failed components. In the Enterprise 10000 server, this extends to all system components, when the system is configured for full redundancy. Therefore, a failure of any component: processor, memory bank, address bus, data router, ASIC, fan, power supply, control board or faulty I/O Board does not keep the entire system down. This approach should not be confused with fault tolerant systems, which are designed to continue running even if a hardware failure occurs. Such systems offer higher availability but very poor price/performance compared to highly available UNIX server solutions.

Specific types of hardware failures, such as a control board failure, will bring down an Enterprise 10000 system. Automatic System Recovery (ASR) enables the system to reboot immediately following a failure, automatically disabling the failed component. This approach also prevents a faulty hardware component from causing the system to crash again or keeping the entire system down.

ASR features involve testing various hardware components when the system is first powered on, or when an external reset is generated. The testing codes, POST and OBP, rely on extensive JTAG instrumentation.

In the Enterprise 10000, unlike the rest of the Ex000 line, the Power-On Self Test (POST) code and Open Boot PROM (OBP) code reside on the System Service Processor. When required, portions of these programs are downloaded into the Enterprise 10000 RAM.

POST

Power On Self Test, or POST, is a suite of hardware integrity tests. POST is executed at different points in time, such as when the system first powers up, and possesses different levels (tests) of execution. POST testing improves system reliability by helping protect against the system booting with faulty hardware that would compromise system availability or data integrity.

Once entered, POST is capable of various levels of testing. Typically, POST initializes and tests all ASICs, memory SIMMs, UltraSPARC Processor Modules, and the Control Board.

POST code is written such that every operation is checked for success or failure. POST testing is hierarchical in nature, starting with basic initialization and functionality testing and moving forward based on the success or failure of previously tested components. As it proceeds, POST disables failed components that could interfere with future testing.

JTAG Instrumentation

JTAG (Joint Test Action Group) is an IEEE standard for testing electronic circuits. JTAG is used in the Ultra Enterprise 10000 servers to communicate the status of components to the Control Board, which communicates it to the SSP, and to allow the SSP, via the Control Board, to modify the state of system components.

During Power On Self Test (POST), this information is used for initialization and diagnostics, and during system operation to verify the integrity of system boards, the power distribution system, and cooling system.

POST also functions as the end point for handling fatal system errors, whether they occur during POST testing or while system software (e.g., the Solaris operating system) is up and running. Following a fatal system error, POST will reconfigure a system around the failed component so the system can be restored to service. A fatal system error is an error that causes the system to be in an illegal hardware state, such as:

- UPA address parity error
- Gigaplane-XB address or control parity error
- DTag parity error
- External cache (E-cache) tag parity error

- UPA Master Port time-out
- Internal error

Improved Data Integrity with ECC

Error Correction Code (ECC) logic is used extensively throughout the Ultra Enterprise 10000 system data and address paths to provide high levels of data integrity.

Single-bit errors detected in the datapath between the UltraSPARC Data Buffer, local data routers, global data router, and the memory subsystem are corrected by the receiving UltraSPARC module and the error is logged.

An Enterprise 10000 system will also report and log correctable ECC errors. A correctable ECC error is any single bit error in a 72 bit field. Such error detection and correction is done on the fly. The ECC implementation can also detect double bit errors in the same 72 bit field and multiple bit errors in the same nibble (4 bits). An uncorrectable data error is not always a fatal system error.

A result of an uncorrectable data ECC error is determined by the mode of operation of the processor when the error occurs. If the processor is executing in user mode when the error occurs, the process will terminate and core dump with no other interruption to the system. If the processor is executing in kernel mode (operating system code), the processor will panic, and bring down the Enterprise 10000 domain (or system, if no domains are active).

In addition to providing ECC protection for data, the address and control lines on the Gigaplane-XB interconnect and the address lines on the UPA are parity protected. Parity protection is also used in the UltraSPARC processor's internal and external cache and in the Duplicate Cache Tags (Dtags).

The global address bus also includes ECC. Correctable errors are handled on the fly, while uncorrectable errors, such as a parity error, will cause an arbitration stop (arbstop.)

Intelligent Monitoring and Reporting

Hostview

Residing on the System Service Processor, the Hostview program is the primary interface to the Ultra Enterprise 10000 server for the operator or system administrator. Hostview is a graphical program that monitors the status of the Enterprise 10000 system and enables the operator to perform dynamic reconfiguration operations.

Using Hostview, the operator can view the system hardware and software configuration and status. Hostview displays the physical layout of the system, including control board(s), system boards and their processors. Hostview indicates with color which boards are grouped into partitions. The state of processors, e.g. booting or Solaris running, is also indicated.

Hostview monitors all system status functions:

- Power on each bulk power supply, control board and support board power supplies.
- Voltage on each power supply on each board
- Thermal conditions of power supplies, processors, ASICs,
- Fan speeds and status for each of the 32 fans installed

Hostview also enables an operator to:

- Power the Enterprise 10000 system on and off
- Dynamically reconfigure the Enterprise 10000 system; perform attach and detach operations, resetting domains, and running diagnostics.
- Create, modify and delete domains
- Access SSP message log for each domain
- Modify the blacklist file to enable or disable components in each domain
- Power the peripherals cabinets on and off

Using Netcon, the remote console function described in a later section, all of the Hostview functions may be accessed from a remote location.

Solstice SyMON

Solstice SyMON system monitor is a graphically based software tool designed to enable system administrators to maintain and monitor server systems. SyMON is an important and integral part of Sun's RAS strategy, and as such will evolve over time with new features and functionality, as well as support for a broader range of hardware.

Solstice SyMON, while not implemented with the first release of the Ultra Enterprise 10000 servers, will be available for Enterprise 10000 servers in future updates. The functions of Hostview will be merged with SyMON, presenting a single, consistent interface for all Enterprise X000 servers.

Solstice SyMON provides the following features:

- Simple, visual, system vital sign indicator
- A physical system view with per-device information, providing a graphical "picture" of the system, with information such as network interface types, disk types, processor module speeds, etc.
- A logical system view, showing the system components and their relationships
- A per-process display, showing process resource utilization and behavior
- An operating system performance statistics display
- Programmable event processing and notification, including scripting capability which enable events to result in sending emails, dialing pagers, etc.
- Event logfile and management (copy, size control)
- System logfile reader/browser
- On-line diagnostics via interface with SunVTS
- Performance monitoring capacity planning, including current
- load analysis
- Hardware failure prediction, based on historical events
- Solstice Site Manager™ and Solstice Domain Manager™ Agent support

SyMON makes the administration and management of Enterprise X000 systems easier and less time consuming. The system's reliability is improved through features that allow for the monitoring of system resource utilization and the setting of thresholds which trigger events when they are reached. For example, an administrator can be paged if processor utilization reaches 80 percent, or system paging activity becomes too high.

Administrators can react to such conditions and make corrections before end-users realize anything is wrong. Other features, such as hardware failure prediction and capacity tools also increase reliability by providing data that allows for a pro-active response to a potentially severe system condition. The fundamental effect is that the system is available when needed.

Solstice SyMON also improves system serviceability. The physical view presented by SyMON makes identification of a failed component a simple matter of viewing the graphical representation of the system. Clicking on a component in the physical system view produces specific information (e.g., I/O controller type, processor speed and cache size, memory SIMM size, etc.) so the correct replacement can be identified and made available.

SNMP Reporting

The System Service Processor (SSP) incorporates an SNMP proxy agent and detailed management information base (MIB). The MIB, Ultra-Enterprise-10000.mib, is aware of all of the instrumentation in Enterprise 10000 components. The MIB defines components as objects to be managed, and define what constitutes an event, such as overtemperature and over-voltage conditions. Such events are reported to the SSP, where the event detector daemon responds to the event by executing a response action script on the SSP. The process of reporting and responding to events is illustrated in the figures below.

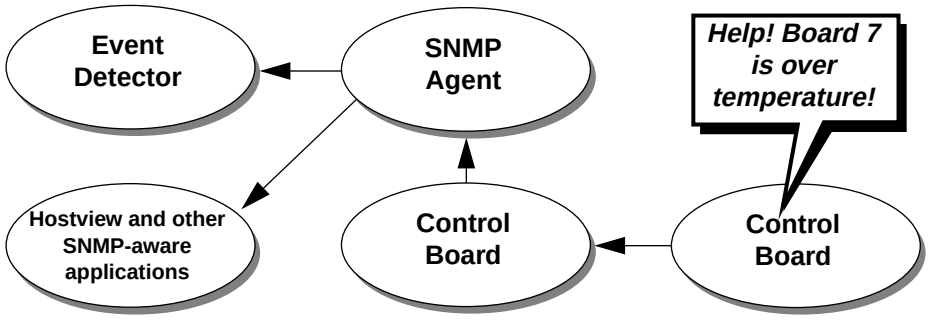

Figure 3-1 Event Recognition and Delivery

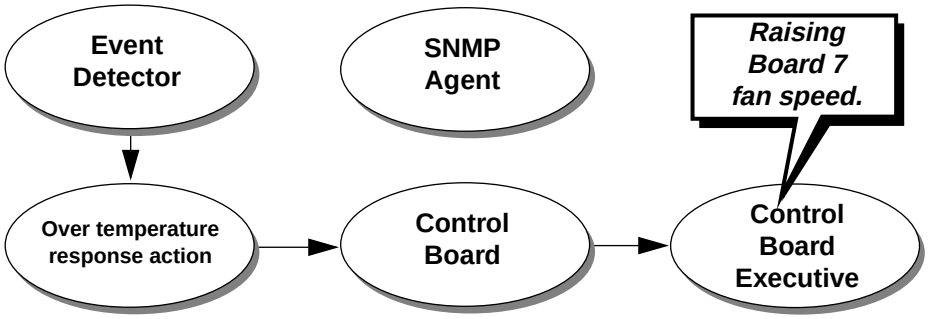

Figure 3-2 Response Action

Events may be optionally reported to other monitoring stations. By enabling other stations, the SNMP messages can be acted upon by SNMP-aware system or network management tools. Thus, the Enterprise 10000 can participate in enterprise-wide system management architectures.

"No Outage" Servicing

The Ultra Enterprise 10000 offers a number of features to reduce downtime and improve the convenience and reliability of servicing. By taking advantage of the Enterprise 10000 server's fully Dynamic Configuration / Reconfiguration capabilities, sites with little or no time window for servicing can perform most hardware upgrades or repairs while the system continues to operate.

Dynamic Reconfiguration Capabilities

Dynamic reconfiguration technology substantially increases a system's serviceability and availability, and offers advantages in several key areas:

- Adding hardware dynamically to enable the system to handle an increasingly larger workload and improve performance of a running system.
- Expanding I/O and network capacity by adding system boards.
- Removing/replacing faulty hardware with minimal system service disruption. The ability to queisce a running system to replace a system board. The time to resume the operating system is negligible in comparison with a full shutdown and cold reboot.

It is useful to define the terms that apply to hot swap technology, in order to better understand the functionality discussed.

- Dynamic Attach - Logically attaching new components to Solaris, making those components available for use
- Dynamic Detach - The process of logically detaching components that were available and on-line to the operating system, making them ready for removal
- Dynamic Reconfiguration - Essentially, support of Dynamic Attach and Dynamic Detach. Being able to add/remove components from a running system
- Hot Swap - The physical act of adding or removing a component on a powered-up system.

Using Dynamic Reconfiguration

Using dynamic attach operations, components can be added to running Enterprise 10000 systems and be activated without the necessity of a reboot. In other words, a system can grow dynamically. Using dynamic detach operations, components may be quiesced and removed, for repair or replacement. With the exception of the centerplane, all boards and power supplies may be removed and replaced without interrupting service.

The DR process manages the flow of processes and data to/from memory in order to preserve them when removing or adding components. The entire process is managed from the SSP console. When a system board has been

newly attached, the operator first applies power from the SSP or network console. The board may be assigned to a new domain, or added to a running system or domain. Then the alternate pathing is switched back, if appropriate. Finally, the Solaris scheduler allows new processes to be run on the board, and data begins to fill the board's memory. When a board is to be detached, the process is reversed; memory is drained of active data, new processes are not allowed onto the board's processors, running processes are allowed to complete. If alternate pathing is in use, the secondary path is activated, and finally, the board is powered off, ready to be physically removed.

Convenient and Secure Component Packaging

System boards are designed for convenient and error-free handling during servicing operations. This is provided by the following features:

- System boards are encased in a metal housing, yet remain under 30 pounds in weight
- System board connectors are keyed so that boards cannot be plugged in upside down
- Control and Centerplane Support boards are keyed such that they cannot be placed in each other's slots
- System boards have LEDs that indicate when the power to the board has been turned off, making it safe for removal from a running system

Sophisticated Diagnostics Support

SunVTS™ Diagnostic Tool

In order to provide better on-line diagnostic capabilities for the entire line of Enterprise X000 systems, the Sun Validation Test Suite (SunVTS) software package was developed. Shipping with all Enterprise X000 systems, SunVTS is a replacement for Sundiag, which previously provided online diagnostics for Sun systems.

The primary goal of the SunVTS™ software is to provide an environment with which Sun systems can be thoroughly stress tested. It is not intended to provide precise component-level fault isolation or POST-like functionality. Rather, it is intended to be used on a system running Solaris, providing a easy-

to-use GUI for initiating and logging the test results from various subsystems (e.g., processors, memory, I/O, etc.) or to test the system as a whole. Some of the key features of SunVTS are:

- UNIX level diagnostics, where system tests execute real UNIX code under the Solaris operating environment
- Automatic system probing, enabling the system configuration to be displayed via a user interface
- Two user interfaces, providing greater flexibility. A GUI-based interface as well as a character based interface will be available. The SunVTS kernel will be cleanly separable from the user interface, such that multiple user interfaces can communicate with the same SunVTS kernel. The character based interface permits the writing of shell scripts to control SunVTS
- An Application Programming Interface, providing a well-defined interface into the SunVTS kernel from other processes, as well as the user interfaces. A SunVTS execution could be initiated in a cron-like fashion, with no direct user interface at all
- Advanced configuration and execution control, allowing tests to be grouped together based on user requirements, with fine grained execution control for status and logging information

Remote Console and Console Dial-In Support

Many data centers encompass computers in multiple locations. It's desirable, then, to be able to communicate with and manage the computer systems, whatever their location. This enables a central operations center to monitor systems in other locations, allowing "lights out" operations. Alternatively, it enables a system administrator to monitor and interact with the Enterprise 10000 console remotely. The Enterprise 10000 supports a network console function, dubbed Netcon, permitting administration from remote sites. The network console is a set of functions that allow access to the SSP via network or dial in. All SSP functions are available, and may be accessed either as a graphical or command based interface.

3

Summary 4

Meeting the demands of business-critical operations requires core technology that is reliable, serviceable, and highly available. Optimum performance and value are no longer the only priorities for selection of a UNIX based server for commercial computing applications.

Sun has recognized this trend for some time, and has demonstrated a clear commitment to providing reliable computing solutions, including its Ultra Enterprise Servers, Enterprise Storage products, SPARCcluster HA and SPARCcluster PDB products, to name only a few.

Now Sun has developed new and comprehensive RAS features, leading the industry in highly available, serviceable and manageable UNIX servers. These services are complemented by highly regarded SunService programs designed to serve Sun's customers in planning, implementing, deploying and managing their key business systems.

The Enterprise 10000 system RAS features described in this paper further illustrate Sun's commitment to the technologies and innovation necessary to providing highly reliable solutions. Coupled with the Solaris operating environment, Enterprise storage products, Solstice system and network management tools, Sun's complete family of Ultra Enterprise 10000 server systems set new standards for affordable reliability, availability and serviceability.

References A

Sun Microsystems, Inc. posts product information in the form of data sheets, specifications, and white papers on its Internet World Wide Web Home page at *http://www.sun.com*.

Look for abstracts on these and other Sun technology white papers:

Token 53096 *Sun RAS Solutions for Mission-Critical Computing*, Sun Microsystems Computer Company, 1996.

Token 53092 *Ultra Enterprise X000 Server Family: Reliability, Availability, and Serviceability*, Executive Overview, Sun Microsystems Computer Company, 1996.

Token 53090 *Ultra Enterprise X000 Server Family: Reliability, Availability, and Serviceability*, Technical White Paper, Sun Microsystems Computer Company, 1996.

Token 53067 *Solstice SyMON System Monitor*, Executive Overview, Sun Microsystems Computer Company, 1996.

Token 53103 *Solstice SyMON System Monitor*, Technical White Paper, Sun Microsystems Computer Company, 1996.

Sun Microsystems Computer Company
A Sun Microsystems, Inc. Business
2550 Garcia Avenue
Mountain View, CA 94043 USA
415 960-1300
FAX 415 969-9131

Sales Offices

Argentina: +54-1-311-0700
Australia: +61-2-9844-5000
Belgium: +32-2-716-7911
Brazil: +55-11-524-8988
Canada: +905-477-6745
Chile: +56-2-638-6364
Colombia: +57-1-218-3933
Commonwealth of Independent States:
+7-095-956-5470
Czech / Slovak Republics:
+42-2-205-102-33
Denmark: +45-44-89-49-89
Estonia: +372-6-308-900
Finland: +358-0-525-561
France: +33-13-067-5000
Germany: +49-89-46008-0
Greece: +30-1-680-6676
Hong Kong: +852-2802-4188
Hungary: +36-1-202-4415
Iceland: +354-563-3010
India: +91-80-559-9595
Ireland: +353-1-671-4678
Israel: +972-3-695-6868
Italy: +39-39-60551
Japan: +81-3-5717-5000
Korea: +822-3469-0114
Latin America / Caribbean:
+1-415-688-9464
Latvia: +371-755-11-33
Lithuania: +370-7-298-586
Luxembourg: +352-491-1331
Malaysia: +603-264-9988
Mexico: +52-5-258-6100
Netherlands: +31-33-450-1234
New Zealand: +64-4-499-2344
Norway: +47-2218-5800
People's Republic of China:
Beijing: +86-10-6849-2828
Shanghai: +86-21-6247-4068
Guangzhou: +86-20-8777-9913
Chengdu: +86-28-678-0121
Poland: +48-2-658-4535
Portugal: +351-1-412-7710
Singapore: +65-224-3388
South Africa: +2711-805-4305
Spain: +34-1-596-9900
Sweden: +46-8-623-90-00
Switzerland: +41-1-825-7111
Taiwan: +886-2-514-0567
Thailand: +662-636-1555
United Arab Emirates:
+971-4-366-333
United Kingdom: +44-1-276-20444
United States: +1-800-821-4643
Venezuela: +58-2-286-1044
Worldwide Headquarters:
+1-415-960-1300

Printed in USA