SUN™ SPARC® ENTERPRISE™ SERVER FAMILY ARCHITECTURE

FLEXIBLE, MAINFRAME-CLASS COMPUTE POWER

White Paper
April 2007

Table of Contents

Chapter 1

Flexible, Mainframe-Class Compute Power for the Datacenter

Reliance upon technology within enterprises is greater than ever. Today, compute systems play a critical role in every function from product design to customer order fulfillment. In many cases, business success is dependent on continuous availability of IT services. Once only required in pockets of the datacenter, mainframe-class reliability and serviceability are now essential for systems throughout the enterprise. In addition, powering datacenter servers and keeping services running through a power outage are significant concerns.

While availability is a top priority, costs must also remain in budget and operational familiarity maintained. To deliver networked services as efficiently and economically as possible, enterprises look to maximize use of every IT asset through consolidation and virtualization strategies. As a result, modern IT system requirements reach far beyond simple measures of compute capacity. Organizations need highly flexible servers with built-in virtualization capabilities and associated tools, technologies, and processes that work to optimize sever utilization. With budgets still in mind, new computing infrastructures must also help protect current investments in technology and training.

Introducing the Sun™ SPARC® Enterprise™ Server Family

Sun™ SPARC® Enterprise™ servers are highly reliable, easy to manage, vertically-scalable systems with all of the benefits of traditional mainframes and none of the associated cost, complexity, or vendor lock-in (Figure 1-1). In fact, Sun SPARC Enterprise servers deliver a mainframe-class system architecture at open systems prices. With symmetric multiprocessing (SMP) scalability from one to 64 processors, memory subsystems as large as 2 TB, and high-throughput I/O architectures, Sun SPARC Enterprise servers easily perform the heavy lifting required by consolidated workloads. Furthermore, Sun SPARC Enterprise servers run the powerful Solaris™ 10 Operating System (OS) and include leading virtualization technologies. By offering Dynamic System Domains, eXtended system boards, Dynamic Reconfiguration, and Solaris Containers technology, Sun SPARC Enterprise servers bring mainframe-class, sophisticated resource control to an open systems compute platform.

Massive compute power, a resilient system architecture, flexible resource control features, and the advanced capabilities of the Solaris 10 OS combine in Sun SPARC Enterprise servers to provide organizations a best-in-class enterprise platform. As an added benefit, Sun SPARC Enterprise servers also offer improved performance over previous generations of Sun servers, with a clear upgrade path that protects existing investments in software, training, and datacenter practices. By taking advantage of Sun SPARC Enterprise servers, IT organizations can create a more powerful

infrastructure, optimize hardware utilization, and increase application availability — resulting in lower operational cost and risk.

Figure 1-1. The Sun SPARC Enterprise server family provides enterprises with scalable power, reliability, and flexibility.

Sun SPARC Enterprise Server Family Overview

The members of the Sun SPARC Enterprise server family share many of the same characteristics which provide power, reliability, and flexibility to enterprises. Sun SPARC Enterprise servers all feature a balanced, highly scalable SMP design that utilizes the latest generation of SPARC64 processors connected to memory and I/O by a new high-speed, low latency system interconnect, delivering exceptional throughput to software applications. Also architected to reduce planned and unplanned downtime, Sun SPARC Enterprise servers include stellar reliability, availability, and serviceability capabilities to avoid outages and reduce recovery time. Design features of Sun SPARC Enterprise servers, such as advanced CPU integration and data path integrity, memory

chipkill and memory mirroring, end-to-end data protection, hot-swappable components, fault resilient power options, and hardware redundancy boost the reliability of these servers.

Sun SPARC Enterprise servers also provide unmatched configuration flexibility. As in other Sun high-end servers, administrators can use Dynamic System Domains to physically divide a single Sun SPARC Enterprise server into multiple electrically isolated partitions, each running independent instances of the Solaris Operating System (Solaris OS). Hardware or software failures in one Dynamic System Domain do not affect applications running in other domains, unless the failed resource is shared across both domains.

Dynamic Reconfiguration can then reallocate hardware resources among Dynamic System Domains — without interrupting critical systems. Sun SPARC Enterprise Servers advance resource control one-step further with eXtended System Board technology, enabling allocation of sub-system board resources such as CPUs, memory, and I/O components to Dynamic System Domains. The fine-grain resource control provided by eXtended System Board technology helps enterprises to further optimize resource utilization.

Adding even more value, the range of compute power offered by the Sun SPARC Enterprise server family provides the levels of vertical scalability required for many deployment classes, enabling enterprises to match the right system to the job at hand. Rackmount Sun SPARC Enterprise M4000 and Sun SPARC Enterprise M5000 servers are economical, powerful, and reliable servers well-suited for mid-range system requirements, while Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers deliver the massive processing power needed for high-end computing (Table 1-1).

Table 1-1. The Sun SPARC Enterprise server family supports midrange and high-end compute requirements.

	Sun SPARC Enterprise M4000 Server	**Sun SPARC Enterprise M5000 Server**	**Sun SPARC Enterprise M8000 Server**	**Sun SPARC Enterprise M9000-32 Server**	**Sun SPARC Enterprise M9000-64 Server**
Enclosure	• 6 rack units	• 10 rack units	• One cabinet	• One cabinet	• Two cabinets
Processors	• SPARC64 VI • 2.15 GHz • 5 MB L2 cache • Up to four dual-core chips	• SPARC64 VI • 2.15 GHz • 5 MB L2 cache • Up to eight dual-core chips	• SPARC64 VI • 2.28 - 2.4 GHz • 5 - 6 MB L2 cache • Up to 16 dual-core chips	• SPARC64 VI • 2.28 - 2.4 GHz • 5 - 6 MB L2 cache • Up to 32 dual-core chips	• SPARC64 VI • 2.28 - 2.4 GHz • 5 - 6 MB L2 cache • Up to 64 dual-core chips
Memory	• Up to 128 GB • 32 DIMM slots	• Up to 256 GB • 64 DIMM slots	• Up to 512 GB • 128 DIMM slots	• Up to 1 TB • 256 DIMM slots	• Up to 2 TB • 512 DIMM slots
Internal I/O Slots	• Four PCIe • One PCI-X	• Eight PCIe • Two PCI-X	• 32 PCIe	• 64 PCIe	• 128 PCIe

	Sun SPARC Enterprise M4000 Server	Sun SPARC Enterprise M5000 Server	Sun SPARC Enterprise M8000 Server	Sun SPARC Enterprise M9000-32 Server	Sun SPARC Enterprise M9000-64 Server
External I/O Chassis	• Up to two units	• Up to four units	• Up to 8 units	• Up to 16 units	• Up to 16 units
Internal Storage	• Serial Attached SCSI • Up to two drives	• Serial Attached SCSI • Up to four drives	• Serial Attached SCSI • Up to 16 drives	• Serial Attached SCSI • Up to 32 drives	• Serial Attached SCSI • Up to 64 drives
Dynamic System Domains	• Up to two	• Up to four	• Up to 16	• Up to 24	• Up to 24

Meeting the Needs of Commercial and Scientific Computing

Suiting a wide range of computing environments, Sun SPARC Enterprise servers provide the availability features needed to support commercial computing workloads along with the raw performance demanded by the high performance community (Table 1-2).

Table 1-2. The power and flexibility of Sun SPARC Enterprise servers provide benefit to a broad range of enterprise applications.

Sun SPARC Enterprise M4000 and Sun SPARC Enterprise M5000 Servers	Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 Servers
• Server consolidation • Business processing (ERP, CRM, OLTP, Batch) • Database • Decision support • Datamart • Web services • System and network management • Application development • Scientific engineering	• Server consolidation • Business processing (ERP, CRM, OLTP, Batch) • Database • Decision support • Data warehouses • IT infrastructure • Application serving • Compute-intensive scientific engineering

Chapter 2
System Architecture

Continually challenged by growing workloads and demands to do more with less, IT organizations realize that meeting processing requirements with fewer, more powerful systems holds economic advantages. The Sun SPARC Enterprise server system interconnect, processors, memory subsystem, and I/O subsystem work together to create a scalable, high-performance platform ready to address server consolidation needs. By taking advantage of Sun SPARC Enterprise servers, organizations can load multiple projects onto a single platform and accelerate application execution at lower costs.

System Component Overview

The design of Sun SPARC Enterprise servers specifically focuses on delivering high reliability, outstanding performance, and true SMP scalability. The characteristics and capabilities of every subsystem within these servers work toward this goal. The Sun SPARC Enterprise server high-bandwidth system bus, powerful SPARC64 VI processor chips, dense memory options, and fast PCI-Express (PCIe) and PCI-eXtended (PCI-X) expansion slots deliver high levels of uptime and throughput, as well as dependable scaling for enterprise applications.

System Interconnect

Based on mainframe technology, the Jupiter system interconnect enables performance scalability and reliability for Sun SPARC Enterprise servers. Multiple system controllers and crossbar units provide point-to-point connections between CPU, memory, and I/O subsystems. Providing more than one bus route between components enhances performance and enables system operation to continue in the event of a faulty switch. Indeed, the system interconnect used in these servers delivers as much as 737 GB/second of peak bandwidth, offering substantially more system throughput than Sun's previous generation of high-end servers. Additional technical details for the system interconnect on each Sun SPARC Enterprise server are found in *Chapter 3 – System Bus Architecture*.

The SPARC64 VI Processor

All Sun SPARC Enterprise servers utilize the Fujitsu SPARC64 VI processor. Based on decades of mainframe experience and designed for outstanding reliability and speed, the Fujitsu SPARC64 VI dual-core, multithreaded processor takes advantage of 90 nm technology and consumes less than 120 watts of power. Additional technical details about the SPARC64 VI processor are found in *Chapter 4 – SPARC64 VI Processor*.

Memory

The memory subsystem of Sun SPARC Enterprise servers increases the scalability and throughput of these systems. In fact, the Sun SPARC Enterprise M9000 server accommodates up to 2 TB of memory. Sun SPARC Enterprise servers use DDR-II DIMMs with 8-way memory interleave to enhance system performance. While multiple DIMM sizes are not supported within a single bank, DIMM capacities can vary across system boards. Available DIMM sizes include 1GB, 2GB, and 4GB densities. Further details about the memory subsystem of each Sun SPARC Enterprise server are described in Table 2-1.

Table 2-1. Sun SPARC Enterprise server memory subsystem specifications.

	Sun SPARC Enterprise M4000 Server	**Sun SPARC Enterprise M5000 Server**	**Sun SPARC Enterprise M8000 Server**	**Sun SPARC Enterprise M9000-32 Server**	**Sun SPARC Enterprise M9000-64 Server**
Maximum Memory Capacity	• 128 GB	• 256 GB	• 512 GB	• 1 TB	• 2 TB
DIMM slots	• 32	• 64	• Up to 128	• Up to 256	• Up to 512
Bank Size	• 4 DIMMs	• 4 DIMMs	• 8 DIMMs	• 8 DIMMs	• 8 DIMMs
Number of Banks	• 8	• 16	• Up to 16	• Up to 32	• Up to 64

Beyond performance, the memory subsystem of Sun SPARC Enterprise servers are built with reliability in mind. ECC protection is implemented for all data stored in main memory, and the following advanced features foster early diagnosis and fault isolation that preserve system integrity and raise application availability.

- Memory patrol — Memory patrol periodically scans memory for errors. This proactive function prevents the use of faulty areas of memory before they can cause system or application errors, improving system reliability.
- Memory chipkill —The memory chipkill function of these servers enables single-bit error correction, enabling processing to continue despite events such as burst read errors that are sometimes caused by memory device failures.
- Memory mirroring — Memory mirroring is an optional, high-availability feature appropriate for execution of applications with the most stringent availability requirements. When memory mirroring mode is enabled on Sun SPARC Enterprise servers, the memory subsystem duplicates the data on write and compares the data on read to each side of the memory mirror. In the event that errors occur at the bus or DIMM level, normal data processing continues through the other memory bus and alternate DIMM set. In Sun SPARC Enterprise M4000 and Sun SPARC Enterprise M5000 servers, memory is mirrored within the same memory module, using the common memory address controller (MAC) Application Specific Integrated Circuit (ASIC) (Figure 2-1 and Figure 2-2).

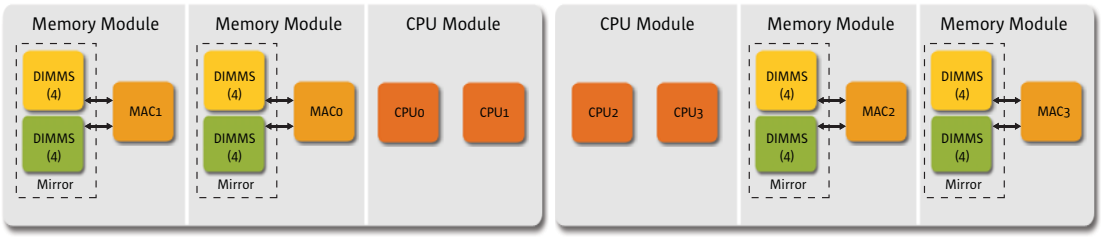

Figure 2-1. Sun SPARC Enterprise M4000 server memory mirroring architecture.

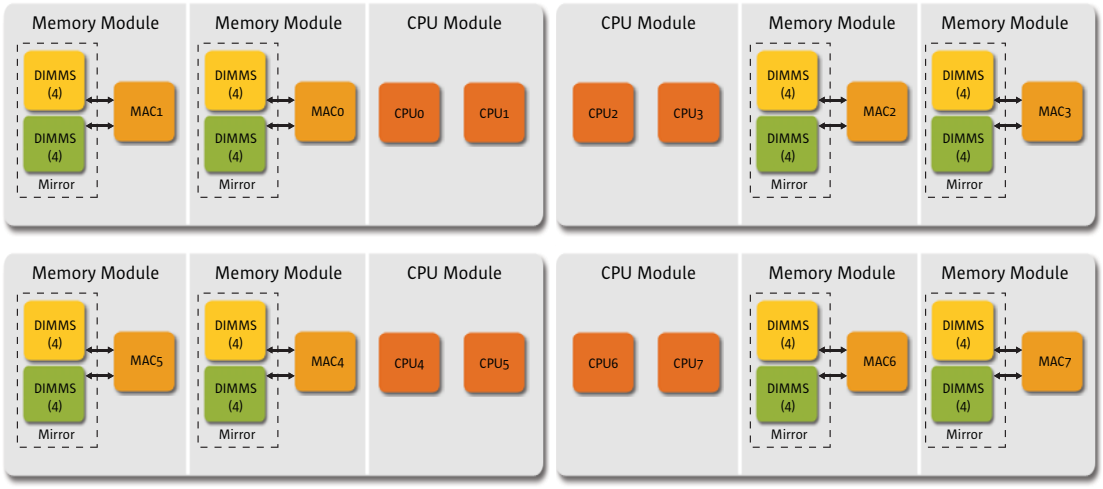

Figure 2-2. Sun SPARC Enterprise M5000 server memory mirroring architecture.

On Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers, memory is mirrored across adjacent MAC ASICs to increase reliability (Figure 2-7). However, in a Quad-XSB configuration, paired DIMMs are split across different Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 Quad-XSBs. As such, memory mirroring is incompatible with the optional configuration of Quad-XSBs on Sun SPARC Enterprise high-end server system boards.

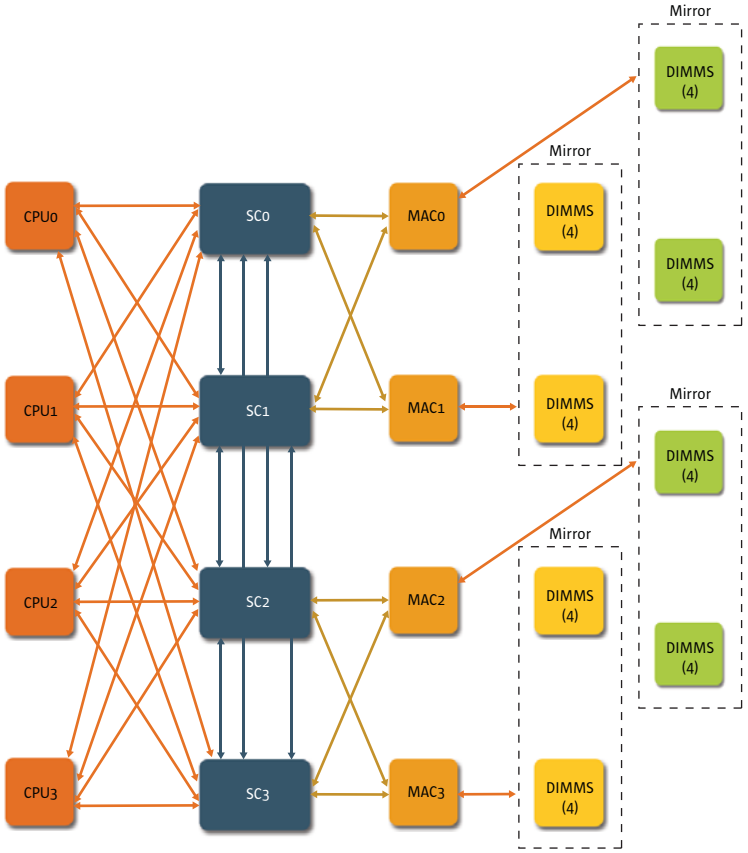

Figure 2-3. Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 server memory mirroring architecture.

System Clock

While the implementation of the system clock varies within each member of the Sun SPARC Enterprise server family, all are engineered with reliability in mind. In particular, Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers use a clock chip with redundant internal components. These high-end servers also implement two sources of clock signal and a dual signal source synchronous line exists between the clock chip and the system boards, enabling the system to restart in the event one route fails. Further enhancing availability and easing maintenance, the Sun SPARC Enterprise M9000 server provides for configuration of a redundant clock control unit.

PCI-Express and PCI-eXtended Technology

Sun SPARC Enterprise servers use a PCI bus to provide high-speed data transfer within the I/O subsystem. In order to support PCIe expansion cards, all Sun SPARC Enterprise servers use a PCIe physical layer (PCIe PHY) ASIC to manage the implementation of the PCIe protocol. PCIe technology doubles the peak data transfer rates of original PCI technology and reaches 2.5 Gb/second of throughput. In fact, PCIe was developed to

accommodate high-speed interconnects such as Fibre Channel, Infiniband, and Gigabit Ethernet. Sun SPARC Enterprise servers also support PCI-X expansion cards for fast access to external devices. PCI-X is backward compatible with existing PCI cards, but increases bandwidth enabling data transfer of up to 1 GB/second for 64-bit devices. Additional technical details for Sun SPARC Enterprise server I/O subsystems can be found in *Chapter 5 – I/O Subsystem*.

Service Processor – Extended System Control Facility

Simplifying management of compute systems leads to higher availability levels for hosted applications. With this in mind, mid-range and high-end models of Sun SPARC Enterprise servers include an eXtended System Control Facility (XSCF). The XSCF consists of a dedicated processor that is independent of the server and runs the XSCF Control Package (XCP) to provide remote monitoring and management capabilities. This service processor regularly monitors environmental sensors, provides advanced warning of potential error conditions, and executes proactive system maintenance procedures as necessary. Indeed, while power is supplied to the server, the XSCF constantly monitors the platform even if the system is inactive. XCP enables Dynamic System Domain configuration, audit administration, hardware control capabilities, hardware status monitoring, reporting, and handling, automatic diagnosis and domain recovery, capacity on demand operations, and XSCF failover services. Additional technical details about the XSCF and XCP are found in *Chapter 7 – System Management*.

Power and Cooling

Sun SPARC Enterprise servers use separate modules for power and cooling. Sensors placed throughout the system measure temperatures on processors and key ASICS as well as the ambient temperature at several location. Hardware redundancy in the power and cooling subsystems combined with environmental monitoring keep servers operating even under power or fan fault conditions.

Fan Unit

Sun SPARC Enterprise server family members use fully redundant, hot-swap fans as the primary cooling system (Table 2-2 and Table 2-3). If a single fans fails, the XSCF detects the failure and switches the remaining fans to high-speed operation to compensate for the reduced airflow. Sun SPARC Enterprise servers operate normally under these conditions, enabling ample time to service the failed unit. Replacement of fans units can occur without interrupting application processing.

Power Supply

The use of redundant power supplies and power cords adds to the fault resilience of Sun SPARC Enterprise servers (Table 2-2 and Table 2-3). Power is supplied to Sun SPARC Enterprise servers by redundant hot-swap power supplies, enabling continued server

operation even if a power supply fails. Since the power units are hot-swappable, removal and replacement can occur while the system continues to operate.

As an option, Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers can be ordered with a three-phase power supply unit and corresponding server cabinet. Models with a three-phase power supply permit two configurations, a star connection that connects a neutral line and each phase, and a delta connection that connects each phase.

Table 2-2. Sun SPARC Enterprise midrange servers power and cooling specifications.

	Sun SPARC Enterprise M4000 Server	**Sun SPARC Enterprise M5000 Server**
Fan Units	• Four fan units • Two 172 mm fans • Two 60 mm fans • One of each type is redundant	• Four fan units • Four 172 mm fans • Two fan groups, each containing two fan units • One redundant fan per fan group
Power Supplies	• 2350 watts • Two units • 1+1 redundant • Single-phase	• 4590 watts • Four units • 2+2 redundant • Single-phase
Power Cords	• Two power cables • 1+1 redundant power cables	• Four power cables • 2+2 redundant power cables

Table 2-3. Sun SPARC Enterprise high-end servers power and cooling specifications.

	Sun SPARC Enterprise M8000 Server	**Sun SPARC Enterprise M9000-32 Server**	**Sun SPARC Enterprise M9000-64 Server**
Fan Units	• 12 fan units • Four 172 mm fans • Eight 60 mm fans • N+1 redundant	• 16 fan units • 16 172 mm fans • N+1 redundant	• 32 fan units • 32 172 mm fans • N+1 redundant
Power Supplies	• 10,500 watts • 9 units • N+1 redundant	• 21,300 watts • 15 units • N+1 redundant	• 42,600 watts • 30 units • N+1 redundant
Options	• Single-phase • Three-phase • Dual-grid	• Single-phase • Three-phase • Dual-grid	• Single-phase • Three-phase • Dual-grid
Power Cords	• 3 power cables (single feed) • 6 power cables (dual feed) • 2 power cables (three-phase)	• 5 power cables (single feed) • 10 power cables (dual feed) • 2 power cables (three-phase)	• 10 power cables (single feed) • 20 power cables (dual feed) • 4 power cables (three-phase)

Optional Dual Power Feed

While enterprises can control most factors within the datacenter, utility outages are often unexpected. The consequences of loss of electrical power can be devastating to IT operations. In order to enable organizations to reduce the impact of such incidents, mid-range and high-end models of Sun SPARC Enterprise servers are dual power feed capable. The AC power subsystem in these servers is completely duplicated, enabling optional reception of power from two external, independent AC power sources. The use of a dual power feed and redundant power supplies increases system availability, as server operations can remain unaffected even after a single power grid failure.

Operator Panel

Mid-range and high-end models of Sun SPARC Enterprise servers feature an operator panel to display server status, store server identification and user setting information, change between operational and maintenance modes, and turn on power supplies for all domains. During server startup, the front panel LED status indicators verify XSCF and server operation.

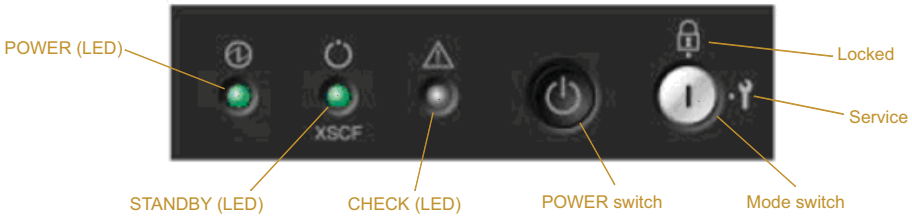

Figure 2-4. The Sun SPARC Enterprise server operator panel.

Midrange Systems — Sun SPARC Enterprise M4000 and Sun SPARC Enterprise M5000 Servers

Sun SPARC Enterprise M4000 and Sun SPARC Enterprise M5000 servers are economical, high-power compute platforms with enterprise-class features. These midrange servers are designed to reliably carry datacenter workloads that support core business operations.

Sun SPARC Enterprise M4000 Server

The Sun SPARC Enterprise M4000 server enclosure measures six rack-units (RU) and supports up to four processor chips, 128 GB of memory, and up to two Dynamic System Domains. In addition, the Sun SPARC Enterprise M4000 server features four short internal PCIe slots and one short internal PCI-X slot, as well as two disk drives, one DVD drive, and an optional DAT tape drive. Two power supplies and four fan units power and cool the Sun SPARC Enterprise M4000 server.

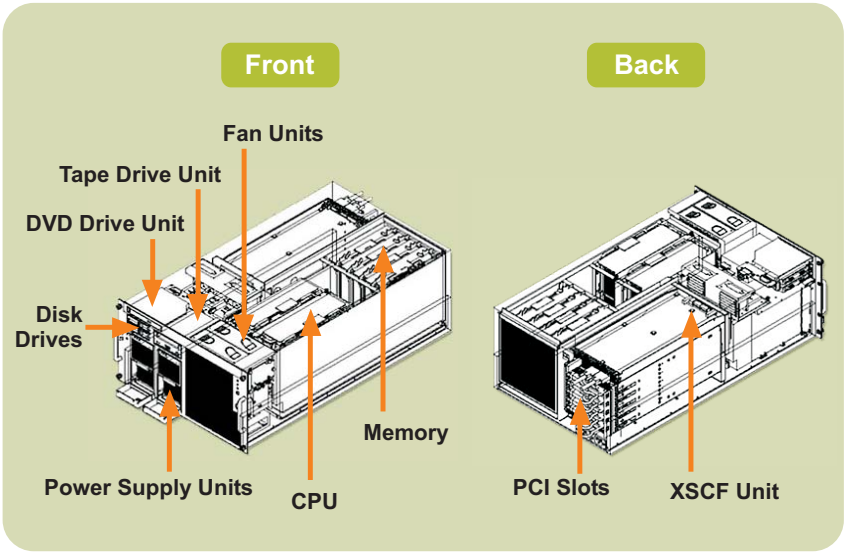

Figure 2-5. Sun SPARC Enterprise M4000 server enclosure diagram.

Sun SPARC Enterprise M5000 Server

The Sun SPARC Enterprise M5000 server enclosure measures 10 RU and supports up to eight processor chips, 256 GB of memory, and up to four Dynamic System Domains. In addition, the Sun SPARC Enterprise M5000 server features eight short internal PCIe and two short internal PCI-X slots, as well as four disk drives, one DVD drive, and an optional DAT tape drive. Four power supplies and four fan units power and cool the Sun SPARC Enterprise M5000 server.

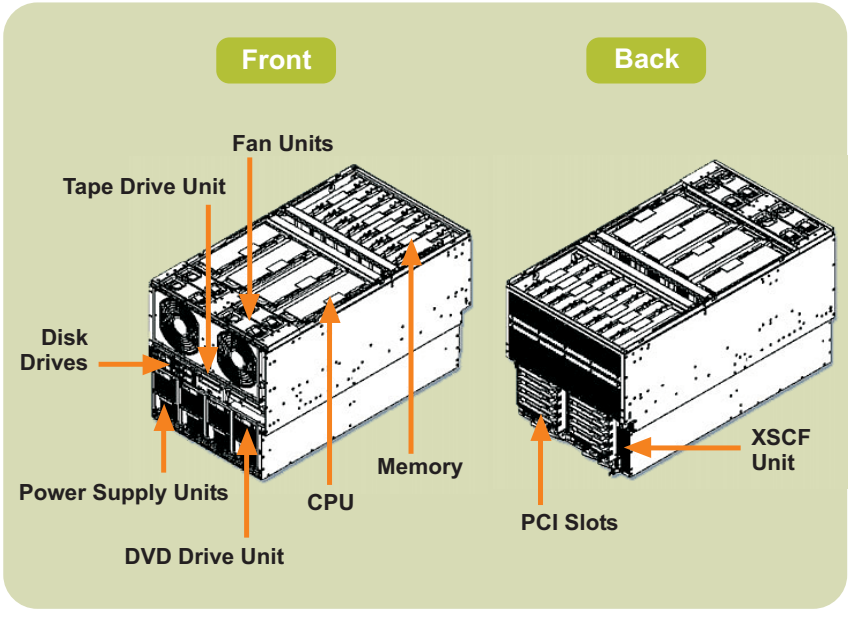

Figure 2-6. Sun SPARC Enterprise M5000 server enclosure diagram.

High-End Systems — Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 Servers

Designed to deliver outstanding performance for even the most challenging workloads, high-end Sun SPARC Enterprise servers merge mainframe reliability, advanced performance technology often used in supercomputers, and an open systems environment to create reliable, high-throughput, flexible systems.

Sun SPARC Enterprise M8000 Server

The Sun SPARC Enterprise M8000 server is mounted in an enterprise system cabinet and supports up to four CPU Memory Units (CMU) and four I/O Units (IOU). Fully configured, the Sun SPARC Enterprise M8000 server houses 16 processor chips, 512 GB of memory, 32 short internal PCIe slots, and can be divided into 16 Dynamic System Domains. In addition, the Sun SPARC Enterprise M8000 server supports up to 16 disk drives, one DVD drive, and an optional DAT tape drive. Nine power supplies and 12 fan units power and cool the Sun SPARC Enterprise M8000 server.

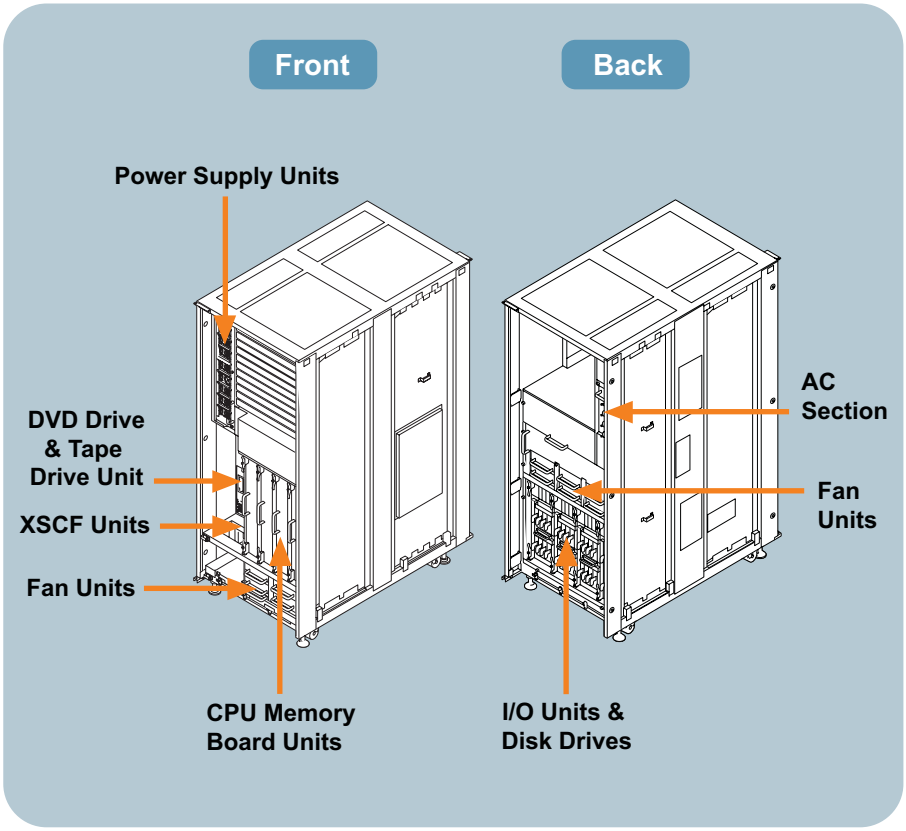

Figure 2-7. Sun SPARC Enterprise M8000 server enclosure diagram.

Sun SPARC Enterprise M9000-32 Server

The Sun SPARC Enterprise M9000-32 server mounts in an enterprise system cabinet and supports up to eight CMUs and eight IOUs. Fully configured, the Sun SPARC Enterprise M9000-32 server houses 32 processor chips, 1 TB of memory, 64 short internal PCIe slots, and can be divided into 24 Dynamic System Domains. In addition, the Sun SPARC Enterprise M9000-32 server supports up to 32 disk drives, one DVD drive, and an optional DAT tape drive. Power and cooling for the Sun SPARC Enterprise M9000-32 server is provided by 15 power supplies and 16 fan units.

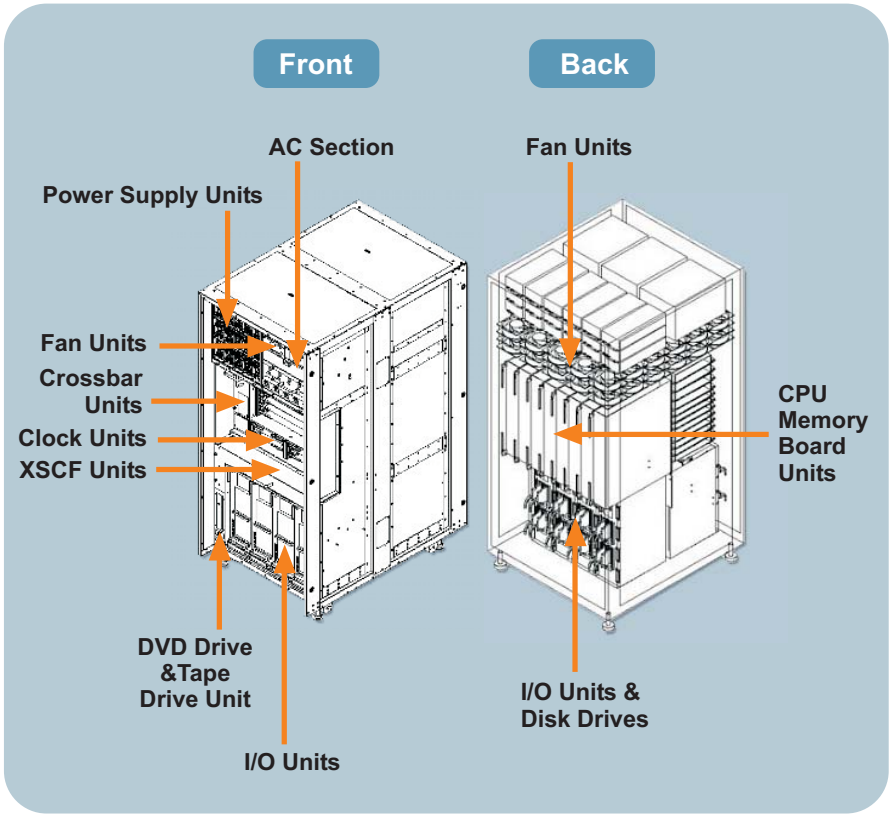

Figure 2-8. Sun SPARC Enterprise M9000-32 server enclosure diagram.

Sun SPARC Enterprise M9000-64 Server

An expansion cabinet can be added to an existing base cabinet to create the Sun SPARC Enterprise M9000-64 server which supports up to 16 CMUs and 16 IOUs. Fully configured, the Sun SPARC Enterprise M9000-64 server houses 64 processor chips, 2 TB of memory, 128 short internal PCIe slots, and can be divided into 24 Dynamic System Domains. In addition, the Sun SPARC Enterprise M9000-64 server supports up to 64 disk drives, two DVD drives, and two optional DAT tape drives. The Sun SPARC Enterprise M9000-64 server utilizes 30 power supplies and 32 fan units for power and cooling.

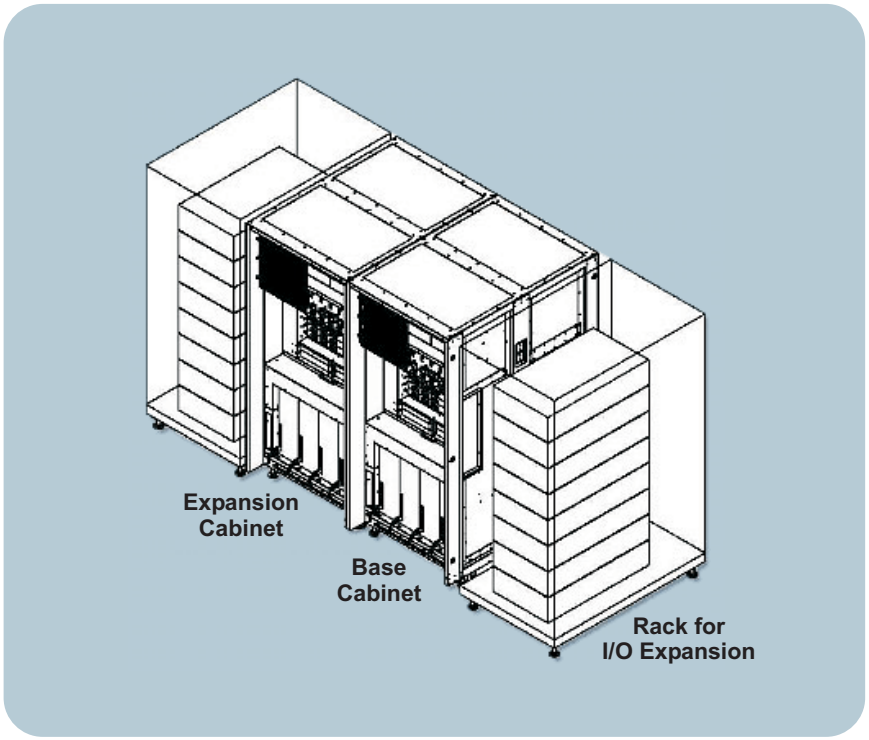

Figure 2-9. Sun SPARC Enterprise M9000-64 server enclosure diagram.

Chapter 3

System Bus Architecture — Jupiter Interconnect

High end systems containing dozens of CPUs only provide scalability if all processors actually contribute to the performance of the application. The ability to deliver near-linear scalability and fast, predictable performance for a broad set of applications rests largely on the capabilities of the system bus. Sun SPARC Enterprise servers utilize a system interconnect designed to deliver massive bandwidth and consistent, low latency between components. The Jupiter system bus benefits IT operations by delivering balanced and predictable performance to application workloads.

Sun SPARC Enterprise Server Interconnect Architecture

The Jupiter interconnect design maximizes the overall performance of Sun SPARC Enterprise servers. Implemented as point-to-point connections which utilize packet-switched technology, this system bus provides fast response times by transmitting multiple data streams. Packet-switching enables the interconnect to operate at much higher system-wide throughput by eliminating “dead” cycles on the bus. All routes are uni-directional, non-contentious paths with multiplexed address, data, and control plus ECC in each direction.

System controllers within the interconnect architecture on all Sun SPARC Enterprise servers direct traffic between local CPUs, memory, I/O subsystems, and interconnect paths. On high-end systems, the system bus is implemented as a crossbar switch between system boards to support high-throughput data transfer with consistent latency times between all components. In addition, the physical addressing of memory on a motherboard of a Sun SPARC Enterprise midrange server or CMU of a Sun SPARC Enterprise high-end server is evenly spread out across all system controllers on that same board, improving performance.

Sun SPARC Enterprise M4000 System Interconnect Architecture

The Sun SPARC Enterprise M4000 system is implemented within a single motherboard. This server design features one logical system board with two system controllers. Both system controllers connect to each other, as well as CPU modules, memory address controllers, and the IOU (Figure 3-1).

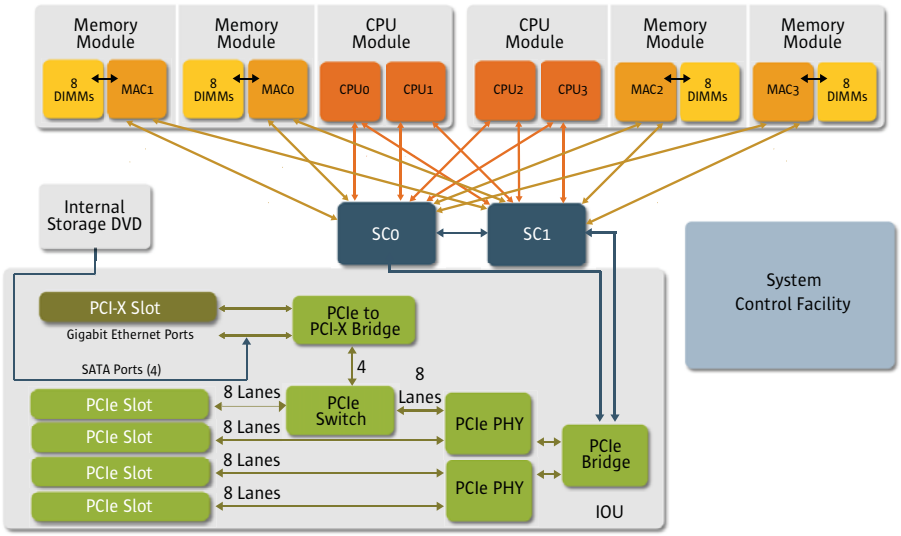

Figure 3-1. Sun SPARC Enterprise M4000 system interconnect diagram.

Sun SPARC Enterprise M5000 System Interconnect Architecture

The Sun SPARC Enterprise M5000 system is implemented within a single motherboard but features two logical system boards. Similar to the Sun SPARC Enterprise M4000 server design, each logical system board contains two system controllers which connect to each other, as well as CPU modules, memory access controllers, and an IOU. In addition, each system controller connects to a corresponding system controller on the other logical system board (Figure 3-2).

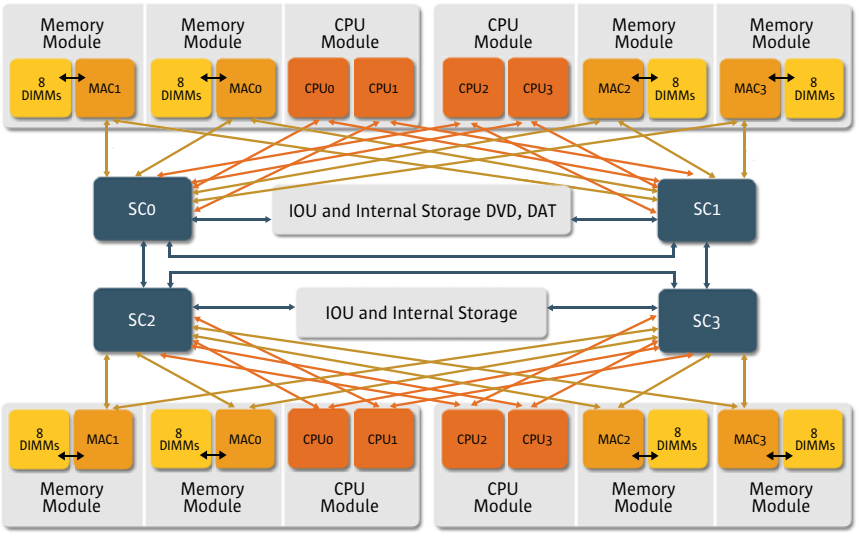

Figure 3-2. Sun SPARC Enterprise M5000 system interconnect diagram.

Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 System Interconnect Architecture

Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers feature multiple system boards which connect to a common crossbar. Each system board contains four system controllers. Each system controller connects to every CPU module. For improved bandwidth, every memory access controller connects to two system controllers, and each system controller connects to every other system controller within the system board. The system controllers also provide a connection to each crossbar unit, enabling data transfer to other system boards (Figure 3-3).

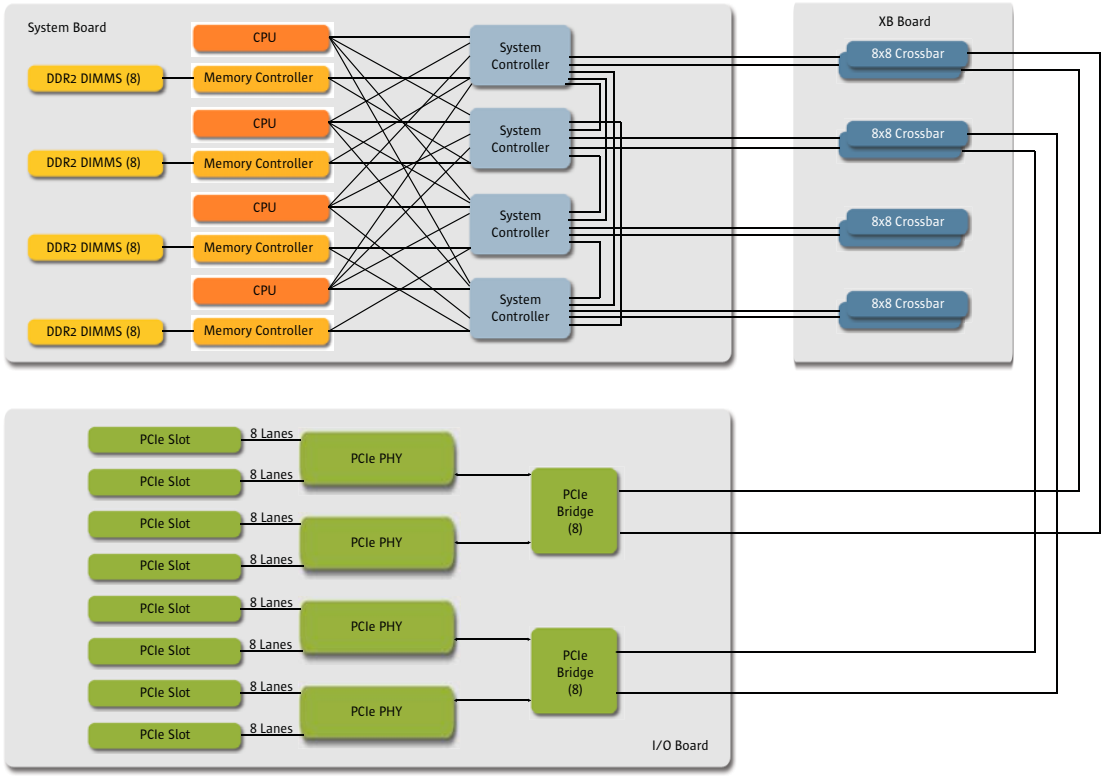

Figure 3-3. Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 system interconnect diagram.

System Interconnect Reliability Features

Built-in redundancy and reliability features of Sun SPARC Enterprise server system interconnects enhance the stability of these servers. The Jupiter interconnect protects against loss or corruption of data with full ECC protection on all system buses and in memory. When a single-bit data error is detected in a CPU, Memory Access Controller, or I/O Controller, hardware corrects the data and performs the transfer. Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers feature degradable crossbar switches and bus routes. In the rare event of a hardware failure within the

interconnect, the system uses the surviving bus route on restart, isolating the faulty crossbar and enabling operations to resume.

Scalable Performance

The high bandwidth and overall design of the Jupiter system interconnect contributes to the scalable performance of Sun SPARC Enterprise servers. Theoretical peak system throughput, snoop bandwidth, and I/O Bandwidth numbers, as well as Stream benchmark results for all Sun SPARC Enterprise servers are found in Table 3-1. In Sun SPARC Enterprise servers, the CPUs, memory address controllers, and IOUs are directly connected by a high-speed broadband switch for data transfer, enabling a relatively even latency to be maintained between individual components. As components are added, processing capability and latency are not degraded. In fact, the crossbar interconnect implementation in Sun SPARC Enterprise high-end servers results in increased interconnect bandwidth every time a system board is added to the server.

Table 3-1. Theoretical peak system and I/O bandwidth, snoop bandwidth and Streams benchmark results for Sun SPARC Enterprise servers.

	Theoretical Peak System Bandwidth[a] (GB/second)	**Snoop Bandwidth (GB/second)**	**Stream Benchmark Triad Results (GB/second)**	**Stream Benchmark Copy Results (GB/second)**	**Theoretical Peak I/O Bandwidth[b] (GB/second)**
Sun SPARC Enterprise M4000 Sever	32	129	12.7	12.5	8
Sun SPARC Enterprise M5000 Server	64	129	25.2	24.8	16
Sun SPARC Enterprise M8000 Server	184	245	69.6	60.3	61
Sun SPARC Enterprise M9000-32 Server	368	245	134.4	114.9	122
Sun SPARC Enterprise M9000-64 Server	737	245	219.3	212.9	244

a.Theoretical Peak System Bandwidth is calculated by multiplying the bus width by the frequency of the bus between the system controller and the memory access controller.
b.Theoretical Peak I/O Bandwidth is calculated by multiplying the bus width by the frequency of the bus between the system controller and the PCI bridge.

Chapter 4
SPARC64 VI Processor

Sun SPARC Enterprise servers utilize the Fujitsu SPARC64 VI Processor which boasts an innovative design that incorporates the latest in dual-core and multithreaded technology, and provides extensive reliability features. In addition, the SPARC64 VI processor is SPARC V9 level 2 compliant, enabling support for thousands of existing software applications. Providing for investment protection, the Sun SPARC Enterprise servers will also support upgrades to future quad-core SPARC64 VII processors. The use of SPARC64 VI processors in Sun SPARC Enterprise servers enables enterprises to benefit from exceptional reliability and application choice in addition to outstanding processing power.

Next Generation Processor Technology

The past decade introduced major changes to processor architectures as system design engineers found that increases to CPU clock rates began exhibiting diminishing returns on performance, and creating power and heat concerns. Innovations such as Chip Multiprocessing (CMP) and Vertical Multithreading (VMT) technologies now dominate plans for improving compute capacity.

Chip Multiprocessing

Chip Multiprocessing technology is an architecture in which multiple physical cores are integrated on a single processor module. Each physical core runs a single execution thread of a multithreaded application independently from other cores at any given time. With this technology, dual-core processors often double the performance of single-core modules. The ability to process multiple instructions at each clock cycle provides the bulk of the performance advantage, but improvements also result from the short distances and fast bus speeds between chips as compared to traditional CPU to CPU communication.

Vertical Multithreading

VMT technology lets a single physical core host multiple threads, each viewed by the operating system as a virtual CPU. Multiple threads on the same core run in a time-sliced fashion, with only one executing at any given moment. A thread does not run if it is idle, or has encountered a cache miss and is waiting for main memory. A thread switch occurs on events such as an L2 cache miss, hardware timer, interrupt, or specific instruction to control threads. In this way, VMT improves system performance by maximizing processor utilization and effectively mitigating the impact of a cache miss. VMT is enabled automatically to improve performance when the number of threads in the system exceeds the number of cores.

SPARC64 VI Processor Architecture

Steady enhancements and changes rather than design overhauls are key to the current success of SPARC64 series processors. Toward that end, the SPARC64 VI CPU module reuses the proven SPARC64 V core with no major pipeline changes. Specifications for the SPARC64 VI processor are detailed in Table 4-1.

Table 4-1. SPARC64 VI processor specifications.

Speed	• 2.15 GHz • 2.28 GHz • 2.4 GHz
Architecture	• Dual-core • SPARC V9 • sun4u • 90nm process technology
L1 Cache	• 128 KB L1 I-cache per core • 128 KB L1 D-cache per core
L2 Cache	• 5 MB (2.15 GHz and 2.28 GHz) and 6 MB (2.4GHz) • 10-way associative (2.15 and 2.28 GHz) and 12-way associative (2.4 GHz) • 256 byte line size (four 64 byte sublines) • ECC tag and data protection
Power	• 120 watts at 2.4 GHz

CPU Implementation

Design changes implemented in the SPARC64 VI processor are geared to keep in step with changes in the software environment. The SPARC64 VI processor implements a combination of CMP and VMT technologies. This processor consists of two physical cores where each core supports two VMT threads, enabling four threads to execute in parallel. The operating system views each thread as a virtual processor. For example, the Solaris OS `psradm` command can be used to set each virtual CPU as spare, on-line, or off-line as desired.

Two threads that belong to the same physical core share most of the core's resources, such as the ALU and instruction pipeline, while the two physical cores only share the L2 cache and system interface. Using coarse grained multithreading techniques, a single thread occupies the full resources of the core until a long latency event. Specifically, a thread switch is triggered on an L2 cache miss or passage of a periodic time interval. This mitigates the effect of cache misses by scheduling an unblocked thread, while maintaining fairness so all threads make progress.

Support for multiple threads in this manner requires duplicating general purpose registers (GPR), floating point registers (FPR), the program counter (PC), and the control (state) registers. A copy of GPR called current window register (CWR) realizes one-cycle register file access time. In addition, a fast, high-bandwidth path from the GPR to the CWR speeds operations when a register window move or thread switch is required.

Performance Improvements

The advanced design of the SPARC64 VI processor speeds system performance. Built with two cores, the SPARC64 VI processor provides double the thread throughput capability of the SPARC64 V processor. Furthermore, VMT takes advantage of L2 cache misses to enable even more threads to run in parallel on the SPARC64 VI cores.

While the CMP and VMT innovations specifically enhance multithreaded performance, single-threaded throughput is not compromised. The SPARC64 VI processor delivers approximately two times the single threaded performance of SPARC64 V running at 1.35 GHz. Other enhancements to the SPARC64 VI processor include a refined core with specific floating point improvements and a doubled Translation Lookaside Buffer (TLB) which reduces the TLB miss rate to improve both integer and floating-point application performance.

Reliability Features

The design of the SPARC64 VI processor module increases system reliability by delivering improved fault avoidance and error correction capabilities. In fact, much of the area on the SPARC64 VI processor is dedicated to error detection and data correction within the CPU. All RAM units are ECC protected or duplicated, and most latches and execution units are parity protected. Rather than force the loss of operation of the entire processor, a single bad core can be isolated and taken offline.

Other reliability features of the SPARC64 VI CPU include support for error marking, instruction retry, and preventive maintenance. When memory read data has a multibit error, a special mark identifying the source of the error is written into the data and the ECC syndrome becomes a special value, providing valuable information for identifying the exact source of the fault. In addition, when a hardware error is detected, all instructions that are currently in execution are cancelled, and retried automatically to prevent transient errors. Error data generated by the SPARC64 VI processor is also sent to the service processor to support preventive maintenance.

Chapter 5

I/O Subsystem

A growing reliance on compute systems for every aspect of business operations brings along the need to store and process ever-increasing amounts of information. Powerful I/O subsystems are crucial to effectively moving and manipulating these large data sets. Sun SPARC Enterprise servers deliver exceptional I/O expansion and performance, enabling organizations to scale systems and accommodate evolving data storage needs.

I/O Subsystem Architecture

The use of PCI technology is key to the performance of the I/O subsystem within Sun SPARC Enterprise servers. A PCIe bridge supplies the connection between the main system and all components of the I/O unit, such as PCI-X slots, PCIe slots, and internal drives. The PCI Express bus also enables the connection of external I/O devices by using internal PCI slots or connecting an External I/O Expansion Unit.

In order to facilitate hot-plug of PCIe and PCI-X adapter cards, Sun SPARC Enterprise servers utilize PCI cassettes. PCI cards which support PCI Hot Plug can be mounted by administrators into a PCI cassette and inserted into an internal PCI slot or External I/O Expansion Unit of a running Sun SPARC Enterprise server.

Sun SPARC Enterprise Midrange Server I/O Subsystem

The Sun SPARC Enterprise M4000 server supports one IOU, while the Sun SPARC Enterprise M5000 server supports two IOUs. A single PCIe bridge connects each IOU to the system controllers and a PCIe to PCI-X bridge enables Sun SPARC Enterprise midrange servers to include on-board PCI-X slots. The single IOU on the Sun SPARC Enterprise M4000 server contains four PCIe slots and one PCI-X slot. The two IOUs on the Sun SPARC Enterprise M5000 server contain a total of eight PCIe slots and two PCI-X slots (Figure 5-1 and Figure 5-2). In addition, an External I/O Expansion Unit increases the number of available PCI slots on midrange Sun SPARC Enterprise servers (Table 5-2).

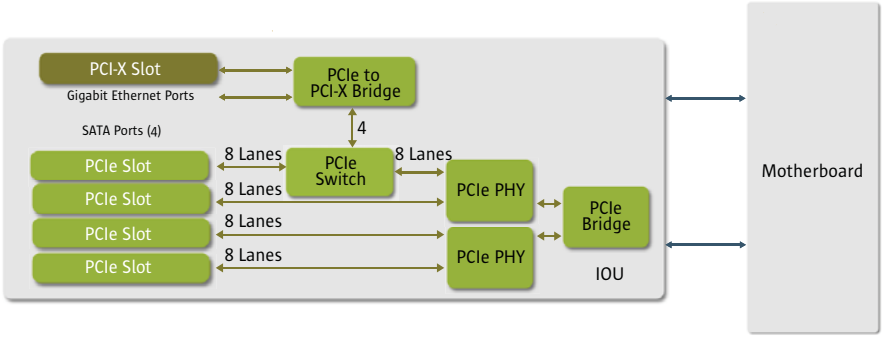

Figure 5-1. Sun SPARC Enterprise M4000 server I/O subsystem architecture.

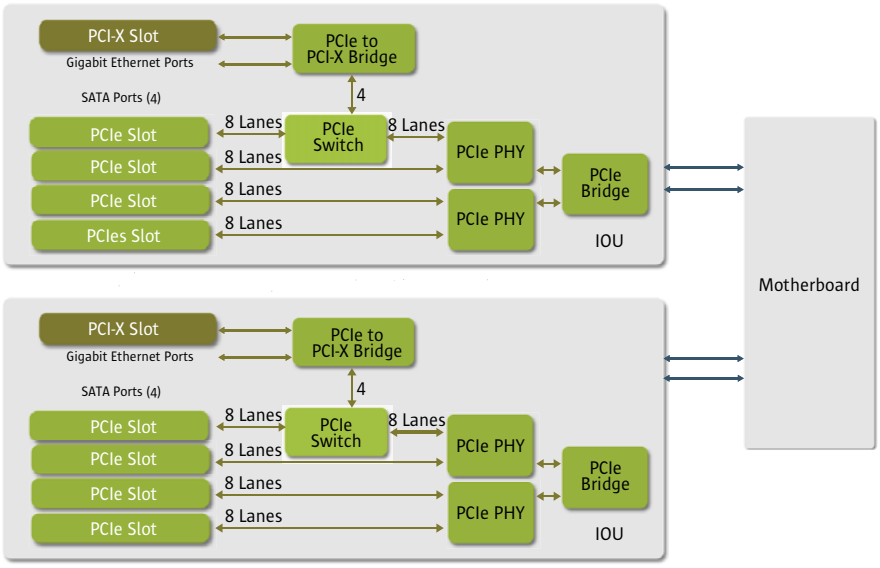

Figure 5-2. Sun SPARC Enterprise M5000 server I/O subsystem architecture.

Sun SPARC Enterprise High-end Server I/O subsystem

A system board on a Sun SPARC Enterprise high-end server can contain one IOU. Two PCIe bridges connect the IOU on each system board to a crossbar switch. Each PCIe bridge also controls communications to four PCIe slots on the system board (Figure 5-3).

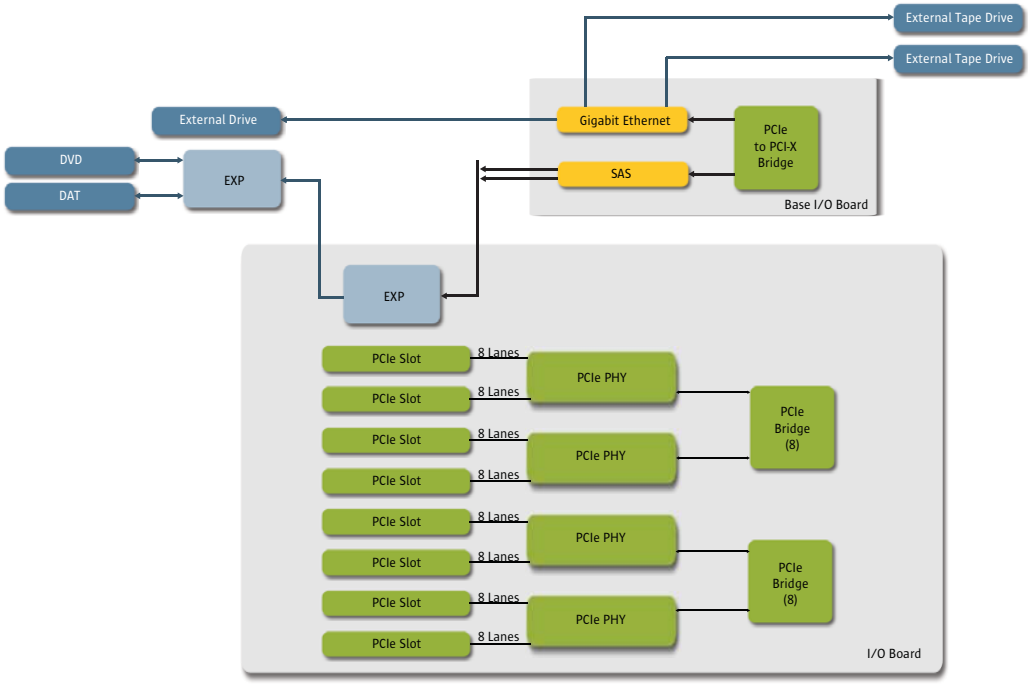

Figure 5-3. Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 server I/O subsystem.

A Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 IOU contains eight PCI Express slots with the total number of PCI slots for these servers dependent upon the number of mounted system boards. The maximum number of internal PCIe slots for Sun SPARC Enterprise high-end servers is listed in Table 5-1. In addition, an External I/O Expansion Unit can be added to a Sun SPARC Enterprise server in order to increase the total number of available PCI slots (Table 5-2).

Table 5-1. Sun SPARC Enterprise high-end server internal PCI slot counts.

Sun SPARC Enterprise Server Model	Maximum Number of Internal PCIe slots
M8000	32
M9000-32	64
M9000-64	128

Internal Peripherals

While disk and tape devices are directly integrated into Sun SPARC Enterprise midrange servers, an add-on base I/O card enables access to internal devices on high-end Sun SPARC Enterprise servers. All Sun SPARC Enterprise servers support one internal DVD drive and an optional DAT tape drive. A Sun SPARC Enterprise M9000 server with an expansion cabinet supports two internal DVD drives and the option for two internal DAT tape drives. Sun SPARC Enterprise servers each support multiple internal Serial Attached SCSI (SAS) 2.5-inch hard disk drives.

External I/O Expansion Unit

Sun SPARC Enterprise servers support the attachment of an optional External I/O Expansion Unit to provide additional I/O connectivity. The External I/O Expansion Unit is a four RU rack mountable device which accommodates up to two IOUs with six PCIe or PCI-X slots. By using PCI cassettes, the external I/O chassis supports active replacement of hot-plug PCI cards.

An I/O Link card mounted in the host provides connectivity to the Sun SPARC Enterprise External I/O Expansion Unit and enables host management control via sideband signals. The I/O link card is available as a low height copper or full height fibre card and includes a single 8-lane PCIe bus with 4GB/second bandwidth. The architecture of the Sun SPARC Enterprise Expansion unit provides high-throughput I/O performance, supporting maximum data rates for many types of PCIe cards and bursty traffic from additional PCIe cards (Figure 5-4).

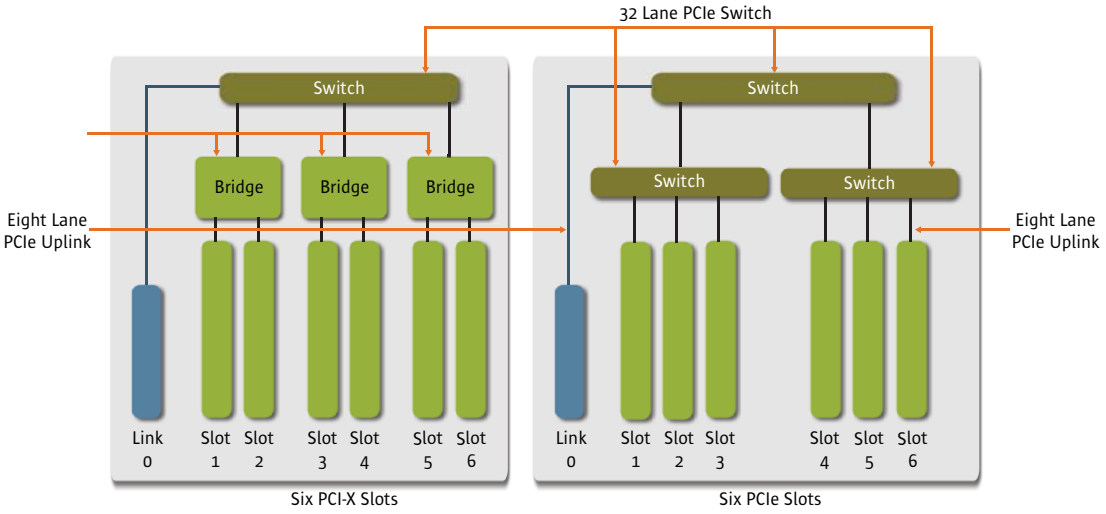

Figure 5-4. External I/O Expansion Unit architecture diagram.

External I/O Expansion Units are added to Sun SPARC Enterprise servers by inserting a link card into an internal PCI Express slot and using a cable to connect the link card. The link card options include a low height copper link card kit or full height fibre link card kit. Sun SPARC Enterprise servers support the connection of multiple External I/O Expansion Units as shown in Table 5-2.

Table 5-2. Sun SPARC Enterprise servers support massive expansion using the optional External I/O Expansion Unit.

Sun SPARC Enterprise Server Model	Maximum Number of External I/O Expansion Units	Maximum Number of PCI Slots
M4000	2	25
M5000	4	50
M8000	8	112
M9000-32	16	224
M9000-64	16	288

To ease management of the External I/O Expansion Unit, I/O Manager software is included with Sun SPARC Enterprise servers and provides the following command line accessible functions.

- Discovers External I/O Expansion Units and FRUs when PCIe slots are powered on
- Collects environmental, voltage, status information
- Logs External I/O Expansion Unit error data

Chapter 6

Reliability, Availability, and Serviceability

Reducing downtime — both planned and unplanned — is critical for IT services. System designs must include mechanisms that foster fault resilience, quick repair, and even rapid expansion, without impacting the availability of key services. Specifically designed to support complex, network computing solutions and stringent high-availability requirements, the systems in the Sun SPARC Enterprise server family include redundant and hot-swap system components, diagnostic and error recovery features throughout the design, and built-in remote management features. The advanced architecture of these reliable servers enables high levels of application availability and rapid recovery from many types of hardware faults, simplifying system operation and lowering costs for enterprises.

Redundant and Hot-Swap Components

Today's IT organizations are challenged by the pace of non-stop business operations. In a networked global economy revenue opportunities remain available around the clock, forcing planned downtime windows to shrink and in some cases disappear entirely. To meet these demands, Sun SPARC Enterprise servers employ built-in redundant and hot-swap hardware to help mitigate the disruptions caused by individual component failures or changes to system configurations. In fact, these systems are able to recover from hardware failures — often with no impact to users or system functionality.

Sun SPARC Enterprise servers feature redundant, hot-swap power supply and fan units, as well as the option to configure multiple CPUs, memory DIMMs, and I/O cards. Administrators can create redundant internal storage by combining Sun SPARC Enterprise server hot-swap disk drives with disk mirroring software. High-end servers also include redundant, hot-swap service processors, and Sun SPARC Enterprise M9000 servers include degradable Crossbar Units and redundant Clock Control Units. If a fault occurs, these duplicated components can enable continued operation. Depending upon the component and type of error, the system may continue to operate in a degraded mode or may reboot — with the failure automatically diagnosed and the relevant component automatically configured out of the system. In addition, hot-swap hardware within the Sun SPARC Enterprise servers speeds service and allows for simplified replacement or addition of components, without a need to stop the system.

Dynamic System Domains

In order to reduce costs and administrative burden, many enterprises look to server consolidation. However, organizations require tools that increase the security and effectiveness of hosting multiple applications on a single server. Dynamic System Domains enable IT organizations to divide a single large system into multiple,

fault-isolated servers each running independent instances of the Solaris OS. With proper configuration, hardware or software faults in one domain remain isolated and unable to impact the operation of other domains. Each domain within a single server platform can even run a different version of the Solaris OS, making this technology extremely useful for pre-production testing of new or modified applications. The maximum number of Dynamic System Domains by server is itemized in Table 6-3.

Table 6-3. Dynamic System Domains limits for Sun SPARC Enterprise servers.

Sun SPARC Enterprise Server Model	Maximum Number of Domains
M4000	2
M5000	4
M8000	16
M9000	24

eXtended System Board

Dynamic System Domains provide a very effective tool for consolidation and enable the ideal separation of resources. In order to achieve this high level of isolation, previous generations of Sun servers designated entire system boards as the smallest unit assignable to a domain. However, some organizations do not require complete hardware isolation and can benefit from the ability to create a higher number of domains with compute power that more precisely matches current workloads. To meet these needs, the Sun SPARC Enterprise server family introduces eXtended System Boards (XSB).

To use a physical system board, the hardware resources on the board are divided, reconfigured as eXtended System Boards, and assigned to a Dynamic System Domain. There are two types of eXtended System Boards. An XSB that consists of an entire system board is called a Uni-XSB. Alternatively, a system board or motherboard that is logically divided into four parts is called a Quad-XSB. The following diagrams depict the logical division lines within each type of Sun SPARC Enterprise server (Figure 6-1, Figure 6-2, and Figure 6-3).

Using eXtended system boards enables granular, sub-system board assignment of compute resources to Dynamic System Domains. A Dynamic System Domain can consist of any combination of Uni-XSBs and Quad-XSBs, providing enterprises the ability to perform sophisticated asset allocation. Determining the exact number and type of XSBs for inclusion in a domain requires balancing the need for fault isolation against the desire to maximize resource utilization. In additions to XSBs, DVD and DAT devices connected to an I/O unit are also assignable to Dynamic System Domains. By using Dynamic System Domains and XSBs, enterprises can better optimize the use of hardware resources while still providing isolated and secure data and programs to customers.

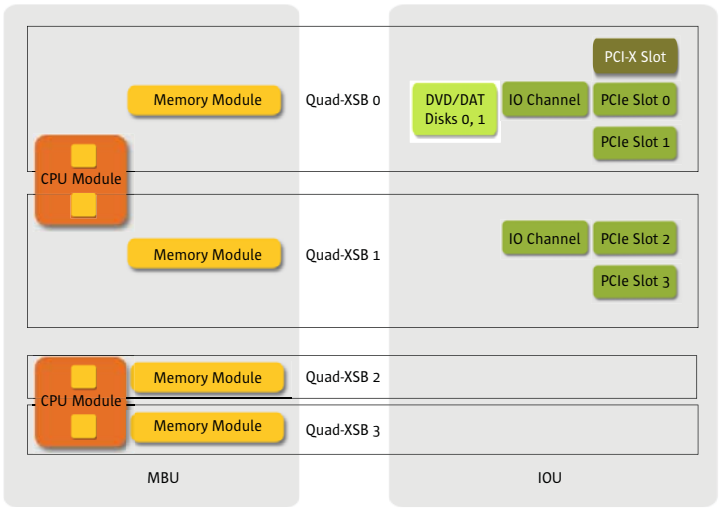

Figure 6-1. Sun SPARC Enterprise M4000 server Quad-XSB configuration.

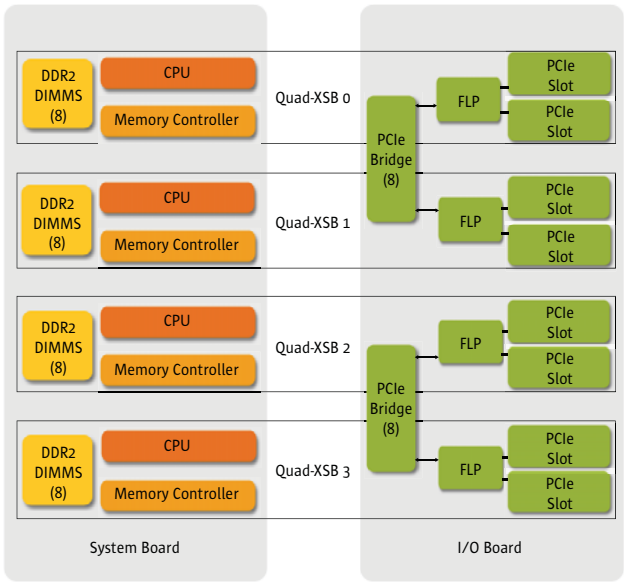

Figure 6-2. Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 system board Quad-XSB configuration.

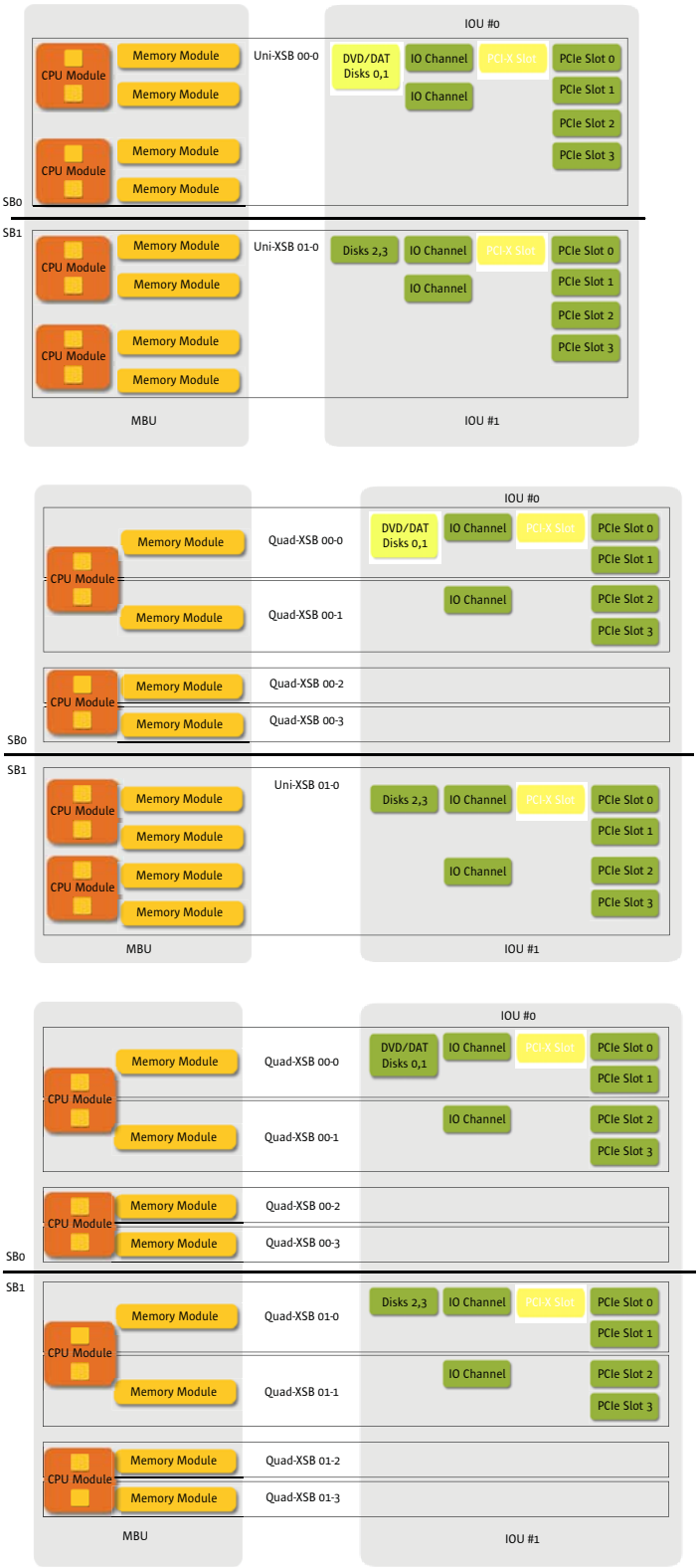

Figure 6-3. Sun SPARC Enterprise M5000 server Uni-XSB and Quad-XSB configurations.

Dynamic Reconfiguration

Dynamic Reconfiguration technology provides added value to Dynamic System Domains by providing administrators with the ability to shift resources without taking the system offline. This technology enables enterprises to perform maintenance, live upgrades, and physical changes to system hardware resources, while the server continues to execute applications. Dynamic Reconfiguration even enables multiple simultaneous changes to hardware configurations without interrupting critical systems.

The ability to remove and add components such as CPUs, memory, and I/O subsystems from a running system helps reduce system downtime. Using Dynamic Reconfiguration simplifies maintenance and upgrades by eliminating the need for system reboots after hardware configuration changes.

Advanced Reliability Features

Advanced reliability features included within the components of Sun SPARC Enterprise servers increase the overall stability of these platforms. For example, all members of the Sun SPARC Enterprise server family include multiple system controllers, and high-end servers include degradable crossbar switches to provide redundancy within the system bus. Reduced component count and complexity within the server architecture contributes to reliability. In addition, advanced CPU integration and guaranteed data path integrity provide for autonomous error recovery by the SPARC64 VI processor, reducing the time to initiate corrective action and subsequently increasing uptime.

Solaris Predictive Self Healing software further enhances the reliability of Sun SPARC Enterprise servers. The implementation of Solaris Predictive Self Healing software for Sun SPARC Enterprise servers enables constant monitoring of all CPUs and memory. Depending upon the nature of the error, persistent CPU soft errors can be resolved by automatically offlining either a thread, core, or entire CPU. In addition, the memory page retirement capability enables memory pages to be taken offline proactively in response to multiple corrections to data access for a specific memory DIMM.

Error Detection, Diagnosis, and Recovery

Sun SPARC Enterprise servers feature important technologies that correct failures early and keep marginal components from causing repeated downtime. Architectural advances which inherently increase reliability are augmented by error detection and recovery capabilities within the server hardware subsystems. Ultimately, the following features work together to raise application availability.

- End-to-end data protection detects and corrects errors throughout the system, ensuring complete data integrity.
- State-of-the-art fault isolation enables the Sun SPARC Enterprise servers to isolate errors within component boundaries and offline only the relevant chips instead of

the entire component. Isolating errors down to the chip improves stability and provides continued availability of maximum compute power. This feature applies to CPUs, memory access controllers, crossbar ASICs, system controllers, and I/O ASICs.

- Constant environmental monitoring provides a historical log of all pertinent environmental and error conditions.
- The host watchdog feature of Sun SPARC Enterprise server family periodically checks for operation of software, including the domain operating system. This feature also uses the XSCF firmware to trigger error notification and recovery functions.
- Sun SPARC Enterprise servers support dynamic CPU resource deallocation which includes processor fault detection, isolation, and recovery. This feature dynamically reallocates CPU resources into an operational system using Dynamic Reconfiguration without interrupting the applications that are running.
- Periodic component status checks are performed to determine the status of many system devices to detect signs of an impending fault. Recovery mechanisms are triggered to prevent system and application failure.
- Error logging, multistage alerts, electronic FRU identification information, and system fault LED indicators all contribute to rapid problem resolution.

Chapter 7
System Management

Providing hands-on, local system administration for server systems is no longer realistic for most organizations. Around the clock system operation, disaster recovery hot sites, and geographically dispersed organizations lead to requirements for remote management of systems. One of the many benefits of Sun servers is the support for *lights-out* datacenters, enabling expensive support staff to work in any location with network access. The Sun SPARC Enterprise system design, combined with a powerful eXtended System Control Facility (XSCF), XSCF Control Package, and Sun N1™ software enables administrators to remotely execute and control nearly any task that does not involve physical access to hardware.These management tools and remote functions lower administrative burden, saving organizations time and reducing operational expenses.

Extended System Control Facility

The eXtended System Control Facility provides the heart of remote monitoring and management capabilities in midrange and high-end Sun SPARC Enterprise servers. The XSCF consists of a dedicated processor that is independent of the server system and runs the XSCF Control Package. The Domain to Service Processor Communication Protocol (DSCP) is used for communication between the XSCF and the server. The DSCP protocol runs on a private TCP/IP-based or PPP-based communication link between the service processor and each domain. While input power is supplied to the server, the XSCF constantly monitors the system even if all domains are inactive.

The XSCF regularly monitors the environmental sensors, provides advance warning of potential error conditions, and executes proactive system maintenance procedures as necessary. For example, the XSCF can initiate a server shutdown in response to temperature conditions which might induce physical system damage. The XSCF Control Package running on the service processor enables administrators to remotely control and monitor domains, as well as the platform itself.

Using a network or serial connection to the XSCF, operators can effectively administer the server from anywhere on the network. Remote connections to the service processor run separately from the operating system and provide the full control and authority of a system console.

Redundant XSCF

On Sun SPARC Enterprise M8000 servers and M9000 servers, one XSCF is configured as active and the other is configured as a standby. The XSCF network between the two service processors enables the exchange of system management information. In case of failover, the service processors are already synchronized and ready to change roles.

DSCP Network

The Domain to Service Processor Communication Protocol service provides a secure TCP/IP and PPP-based communications link between the service processor and each domain. Without this link, the XSCF cannot communicate with the domains. The service processor requires one IP address dedicated to the DSCP service on its side of the link, and one IP address on each domain's side of the link. In a system with more than one XSCF, the standby XSCF does not communicate with the domains. In the event of a failover of the XSCF, the newly active XSCF assumes the IP address of the failed-over service processor.

XSCF Control Package

The XSCF Control Package enables users to control and monitor Sun SPARC Enterprise server platforms and individual Dynamic System Domains quickly and effectively. The XSCF Control Package provides a command line interface (CLI) and Web browser user interface that gives administrators and operators access to all system controller functionality. Password-protected accounts with specific administration capabilities also provide system security for domain consoles. Communication between the XSCF and individual domains uses an encrypted connection based on Secure Shell (SSH) and Secure Socket Layer (SSL), enabling secure, remote execution of commands provided with the XSCF Control Package.

The XSCF Control Package provides the interface for the following key server functions.

- Execution of Dynamic Reconfiguration tasks to logically attach or detach installed system boards from the operating system while the domain continues to run applications without interruption.
- Domain administration which consist of creating logical system boards comprised of Uni-XSB and Quad-XSB units.
- Audit administration includes the logging of interactions between the XSCF and the domains.
- Monitor and control of power to the components in all Sun SPARC Enterprise servers.
- Interpretation of hardware information presented, and notification of impending problems such as high temperatures or power supply problems, as well as access to the system administration interface.
- Integration with the Fault Management Architecture of the Solaris OS to improve availability through accurate fault diagnosis and predictive fault analysis.
- Execution and monitoring of diagnostic programs, such as the Open Boot Prom (OBP) and power-on self-test (POST).

- Execution of Sun Capacity on Demand operations which provide the ability to stage and then later activate additional processing resources.
- Monitoring of the dual XSCF configuration on Sun SPARC Enterprise M8000 and Sun SPARC Enterprise M9000 servers for failure and performing an automatic failover if needed.

Role Based System Management

The XSCF Control Package enables the independent administration of several autonomous domains by different system administrators and operators — all cooperating within a single Sun SPARC Enterprise platform. This management software supports multiple user accounts which are organized into groups. Different privileges are assigned to each group. Privileges allow a user to perform a specific set of actions on a specific set of hardware, including physical components, domains, or physical components within a domain. In addition, a user can possess multiple, different privileges on any number of domains.

Sun™ Management Center and Sun N1™ Software

Controlling a rapidly changing IT infrastructure requires intelligent management tools and an ability to provision servers efficiently. Sun™ Management Center software simplifies administrative tasks and Sun N1 software automates complex software installation and configuration, helping to ease operations.

Sun Management Center

Sun Management Center software improves management efficiency by enabling an aggregate view of the entire network of Sun components from the heart of the datacenter to remote locations at the edge of the network. By using Sun Management Center software, a single interface provides IT administrators with the ability to proactively manage and monitor remote Sun systems, storage components, the Solaris OS, and applications. Remote access services enable system administrators, onsite or offsite, to gain protected access through administrative networks. From this interface technicians can monitor system health, perform remote bring up, and restart or take down individual machines.

Consolidating management views even further, integration with Solaris Cluster systems also allows visibility and control over cluster resources. In addition, to support legacy networks and heterogeneous environments, Sun Management Center software tightly integrates with all major management frameworks including CA Unicenter TNG, HP Open-View, IBM Tivoli, and BMC Patrol.

Sun Management Center software is based on a three-tiered design with an agent-based framework that provides a single point of management for the enterprise (Figure 7-1). This approach simplifies systems management by delivering a higher level of scalability and availability while reducing cost and complexity. Sun Management Center software provides two separate agents for the Sun SPARC Enterprise servers. A *domain agent* provides monitoring and management for individual domains. A separate *platform agent* runs on the XSCF.

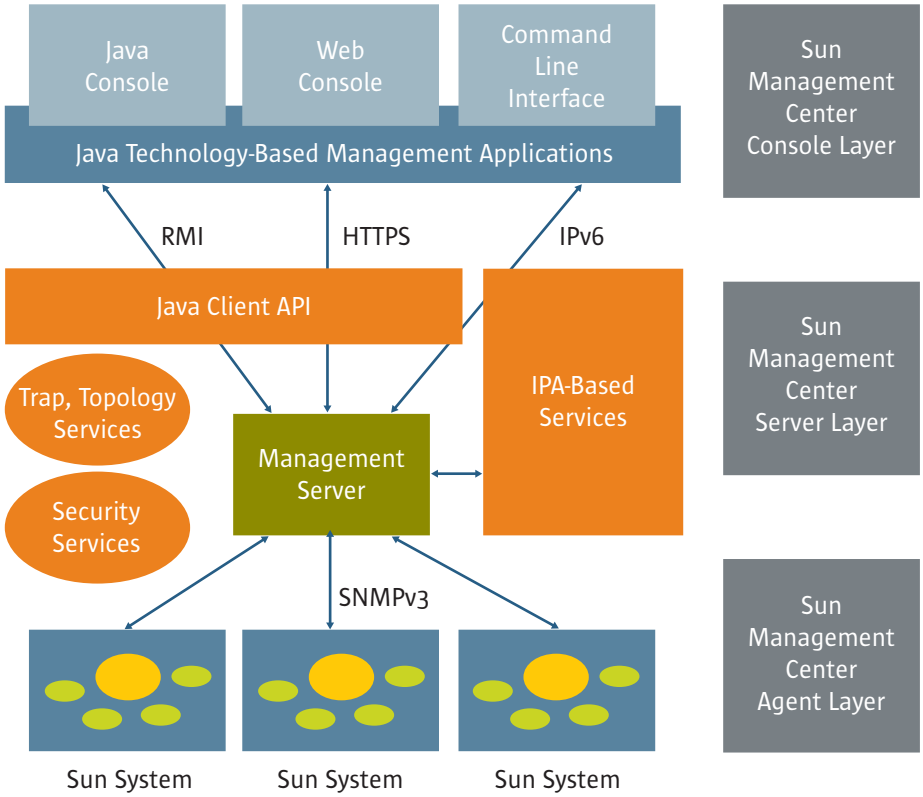

Figure 7-1. Sun Management Center software three-tier architecture.

A major strength of Sun Management Center software is that these autonomous agents continue to operate even when contact with the central software server is lost. A disaster could wipe out the central server, but remote agents continue to collect data and take action if an event is triggered. When the server for Sun Management Center software is restarted, perhaps at a backup site, the agents are reconnected. The central server collects data on all events that occurred during the downtime, allowing administrators to retain historical perspectives.

Sun Management Center software can also manage Dynamic Reconfiguration within Sun SPARC Enterprise servers, including domain creation, dynamic board attach, and dynamic detach. (The actual attach and detach operations take place within the affected domain, but Sun Management Center software provides the Graphical User

Interface (GUI) for initiating these activities.) Sun Management Center software is able to communicate with the XSCF to retrieve status information not only for active domains, but also for boards not currently assigned to any domain — an important requirement for managing the deployment of resources between domains.

Sun N1 Service Provisioning System

Sun N1 Service Provisioning System software simplifies application life cycle management by rapidly provisioning new business services and enabling quick replication of a complete software stack on servers throughout the IT infrastructure. This software offers organizations a standard method of deploying business services and tracking changes throughout the deployment process and includes an audit trail to help enterprises meet regulatory compliance.

Sun N1 Service Provisioning System provides pre-built models for deploying industry leading Java™ 2 Platform, Enterprise Edition (J2EE™ platform) application servers, Web servers, and databases and enables administrators to create custom XML models to deploy, install, upgrade, start and stop other applications. Every action taken by administrators across applications and managed servers is recorded, simplifying rollback to previous states in the event of a service incident.

Sun N1 System Manager

Managing large enterprise infrastructures provides challenge for even the most sophisticated IT organizations. The Sun N1 System Manager software enables rapid discovery, configuration, and provisioning of hundreds of Sun servers from a single console. Designed to simplify datacenter management tasks, Sun N1 System Manager software enables remote power control, operating system deployment and patching, firmware updates, event logging and notification, and hardware and operating system monitoring. A hybrid user interface that integrates a GUI and a CLI in one console executes management tasks. In addition, this software features the ability to create logical groups of systems and perform actions across these groupings as easily as performing actions on a single node. By providing fast and easy access to systems for monitoring and maintenance, the Sun N1 System Management software helps organizations reduce total cost of ownership and increase operational efficiency.

Chapter 8

The Solaris™ 10 Operating System

With mission-critical business objectives on the line, enterprises need a robust operating environment with the ability to optimize the performance, availability, security, and utilization of hardware assets. In a class by itself, the Solaris 10 OS offers many innovative technologies to help IT organizations improve operations and realize the full potential of Sun SPARC Enterprise servers.

Observability and Performance

Organizations need to make effective use of the power of hardware platforms. The Solaris OS supports near linear scalability from 1 to 72 CPUs (144 cores) and memory addressability that reaches well beyond the physical memory limits of even Sun's largest server. The following advanced features of the Solaris 10 OS provide IT organizations with the ability to identify potential software tuning opportunities and maximize raw system throughput.

- Solaris Dynamic Tracing framework (DTrace) is a powerful tool that provides a true, system-level view of application and kernel activities, even those running in a Java Virtual Machine. DTrace software safely instruments the running operating system kernel and active applications without rebooting the kernel or recompiling — or even restarting — software. By using this feature, administrators can view accurate and concise information in real time and highlight patterns and trends in application execution. The dynamic instrumentation provided by DTrace enables organizations to reduce the time to diagnose problems from days and weeks to minutes and hours, resulting in faster data-driven fixes.
- The highly scalable, optimized TCP/IP stack in the Solaris 10 OS lowers overhead by reducing the number of instructions required to process packets. This technology also provides support for large numbers of connections and enables server network throughput to grow linearly with the number of CPUs and network interface cards (NICs). By taking advantage of the Solaris 10 OS network stack, organizations can significantly improve application efficiency and performance.
- The memory handling system of the Solaris 10 OS provides multiple page size support in order to enable applications to access virtual memory more efficiently, improving performance for applications that use large memory intensively. In addition, Solaris 10 OS Memory Placement Optimization (MPO) works to ensure that data is stored in memory as close as possible to the processors that accesses it while still maintaining enough balance within the system. MPO can boost performance in business workloads by as much as 20 percent and as much as 50 percent in some High Performance Computing workloads.

- The Solaris OS multithreaded execution model plays an important role in enabling Sun servers to deliver scalable performance. Improvements to the threading capabilities in the Solaris OS occur with every release, resulting in performance and stability improvements for existing applications without recompiliation.

Availability

The ability to rapidly diagnose, isolate, and recover from hardware and application faults is paramount for meeting the needs of non-stop business operations. Long standing features of the Solaris OS provide for system self-healing. For example, the kernel memory scrubber constantly scans physical memory, correcting any single-bit errors in order to reduce the likelihood of those problems turning into un-correctable double-bit errors. The Solaris 10 OS takes a big leap forward in self-healing with the introduction of Solaris Fault Manager and Solaris Service Manager technology. With this software, business-critical applications and essential system services can continue uninterrupted in the event of software failures, major hardware component breakdowns, and software misconfiguration problems.

- Solaris Fault Manager software reduces complexity by automatically diagnosing faults in the system and initiating self-healing actions to help prevent service interruptions. The Solaris Fault Manager diagnosis engine produces a fault diagnosis once discernible patterns are observed from a stream of incoming errors. Following error identification, the Solaris Fault Manager provides information to agents that know how to respond to specific faults. Problem components can be configured out of a system before a failure occurs — and in the event of a failure, this feature initiates automatic recovery and application re-start. For example, an agent designed to respond to a memory error might determine the memory addresses affected by a specific chip failure and remove the affected locations from the available memory pool.
- Solaris Service Manager software converts the core set of services packaged with the operating system into first-class objects that administrators can manipulate with a consistent set of administration commands. Using Solaris Service Manager, administrators can take actions on services including start, stop, restart, enable, disable, view status, and snapshot. Service snapshots save a service's complete configuration, giving administrators a way to roll back any erroneous changes. Snapshots are taken automatically whenever a service starts to help reduce risk by guarding against erroneous errors. The Solaris Service Manager is integrated with Solaris Fault Manager. When a low-level fault is found to impact a higher-level component of a running service, Solaris Fault Manager can direct Solaris Service Manager to take appropriate action.

In addition to handling error conditions, efficiently managing planned downtime greatly enhances availability levels. Tools included with the Solaris OS, such as Solaris

Flash and Solaris Live Upgrade software, can help enterprises achieve more rapid and consistent installation of software, upgrades, and patches, leading to improved uptime.

- The Solaris Flash facility enables IT organizations to quickly install and update systems with an OS configuration tailored to enterprise needs. This technology provides tools to system administrators for building custom rapid-install images—including applications, patches, and parameters—that can be installed at a data rate close to the full speed of the hardware.
- The Solaris Live Upgrade facility provides mechanisms to upgrade and manage multiple on-disk instances of the Solaris OS. This technology enables system administrators to install a new operating system on a running production system without taking it offline, with the only downtime for the application being the time necessary to reboot the new configuration.

Security

Today's increasingly connected systems create benefits and challenges. While the global network offers greater revenue opportunities, enterprises must pay close attention to security concerns. The most secure OS on the planet, the Solaris 10 OS provides features previously only found in Sun's military-grade Trusted Solaris OS. These capabilities enable the strong controls required by governments and financial institutions but also benefit all enterprises focused on security concerns and requirements for auditing capabilities.

- User Rights Management and Process Rights Management work in conjunction with Solaris Container virtualization technology to enable multiple applications to securely share the same domain. Security risks are reduced by granting users and applications only the minimum capabilities needed to perform assigned duties. Best yet, unlike other solutions on the market, no application changes are required to take advantage of these security enhancements.
- Solaris Trusted Extensions extend the existing Solaris 10 OS security policy with labeling features previously only available in highly specialized operating systems or appliances. These extensions deliver true multilevel security within a commercial-grade operating system, beneficial to civilian organizations with specific regulatory or information protection requirements.
- Core to the Solaris 10 OS are features which fortify platforms against compromise. Firewall protection technology included within the Solaris 10 OS distribution protects individual systems against attack. In addition, file integrity checking and digitally signed binaries within the Solaris 10 OS enable administrators to verify platforms remain untouched by hackers. Secure remote access capabilities also increase security by centralizing the administration of system access across multiple operating systems.

Virtualization and Resource Management

The economic need to maximize the use of every IT asset often necessitates consolidating multiple applications onto single server platforms. Virtualization techniques enhance consolidation strategies one step further by enabling organizations to create administrative and resource boundaries between applications within each domain on a server. Solaris Containers technology provides a breakthrough approach to virtualization and software partitioning, enabling the creation of many private execution environments within a single instance of the Solaris OS. Using this technology, IT organizations can quickly harness and provision idle compute power into a secure, isolated runtime environment for a new deployments without increasing the number of operating system instances to manage. In addition, hosting applications within individual Solaris Containers provides administrators the ability to exert fine-grained control over rights and resources within a consolidated server.

In addition, Solaris Resource Manager software enables the allocation of computing resources within Solaris Containers and among individual tasks and users in a structured, policy-driven fashion. Using the Solaris OS resource management facilities to proactively allocate, control, and monitor system resources —such as CPU time, processes, virtual memory, connect time, and logins— on a fine-grained basis helps organizations obtain more predictable service levels. As business needs change, Solaris Resource Manager software helps enterprises to regularly set new priorities for the use of compute resources. By taking advantage of Solaris Containers and Solaris Resource Manager software, organizations can improve resource utilization, reduce downtime, and lower solution costs.

Chapter 9

Summary

To support the endless demand for scalability, reliability, and manageability in the datacenter, infrastructures need to provide ever-increasing performance and capacity while becoming simpler to deploy, adjust, and manage. Sun SPARC Enterprise servers outfitted with SPARC64 VI processors, large memory capacity, an inherently reliable architecture, and an eXtended system control facility deliver new levels of power, availability, and ease-of-use to enterprises. The sophisticated resource control enabled by Dynamic System Domains, eXtended system boards, and Dynamic Reconfiguration further increase the value of these servers by enabling enterprises to optimize the use of these hardware assets. By taking advantage of fast, scalable Sun SPARC Enterprise servers, organizations gain extraordinary power and flexibility — a strategic asset in the quest to get ahead and stay ahead of the competition.

For More Information

To learn more about innovative Sun products and the benefits of Sun SPARC Enterprise servers and the Solaris OS, contact a Sun sales representative or consult the related Web sites listed in Table 1 below.

Table 1. Related Web sites.

Web Site URL	Description
sun.com/servers	Sun Servers
sun.com/servers/sparcenterprise	Sun SPARC Enterprise Servers
sun.com/software/solaris	The Solaris Operating System
sun.com/software/n1gridsystem	Sun N1 Software
sun.com/servicesolutions	Sun Services and Solution

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN (9786) **Web** sun.com

Printed in USA 04/07