Sun Enterprise™ 10000

Just the Facts

Part 1 of 2

Sun microsystems

Copyrights
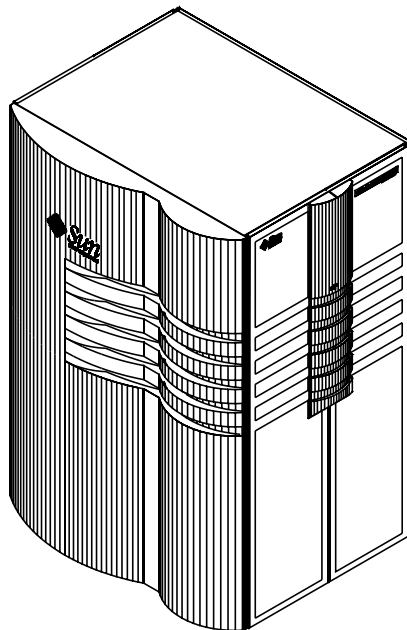
Positioning

Overview

Figure 1. The Sun Enterprise 10000 System

The Sun Enterprise™ 10000 system is a SPARC™ processor-based, scalable symmetrical multiprocessing (SMP) computer system running on the Solaris™ Operating Environment (UNIX® System V, Release 4). It is an ideal, general-purpose application and data server for host-based or client-server applications such as on-line transaction processing (OLTP), decision support systems (DSS), data warehousing, communications services, or multimedia services.

The Sun Enterprise 10000 can be configured with up to 64 processors, up to 64 GB of memory, over 60 TB of on-line disk storage, and a wide range of UNIX application software. All functional areas of the Sun Enterprise 10000 system are field upgradeable, and most upgrades can be performed without disrupting users or halting the system.

The Sun Enterprise 10000 system incorporates RAS features that are unique for a UNIX server. Two examples are Dynamic Reconfiguration (the ability to on-line service boards) and Dynamic System Domains (whereby the Sun Enterprise 10000 can be logically partitioned into multiple smaller servers).

Overview *(cont.)*

The Sun Enterprise 10000 system combines the power of Gigaplane-XB™ interconnect technologies with hardware and software based on UltraSPARC™ technology. By using the Gigaplane-XB interconnect at the core of the system, the Sun Enterprise 10000 system provides a data bandwidth of up to 12.8 GB per second. In addition to a large system memory, each Sun Enterprise 10000 processor utilizes an efficient, fully coherent, local cache to support scalable multiprocessing in an SMP environment.

The physical package of the Sun Enterprise 10000 system houses the system and control boards, the centerplane, the DC power supplies, and the cooling fans. There is also room in the system cabinet for more than 288 GB of disk storage. The I/O expansion cabinets can hold additional disks and tape drives.

Each Sun Enterprise 10000 system has an external system service processor (SSP) to perform system management functions while monitoring the Sun Enterprise 10000 host for problems and to take corrective action when needed.

The Sun Enterprise 10000 may be clustered for Availability (failover) or Scalability. Up to four nodes can be clustered.

Target Markets and Users for the Sun Enterprise 10000 System

The Enterprise family of servers is targeted at strategic markets chosen by Sun: Manufacturing, finance, telecommunications, government, education, health care, retail, design automation, and oil and gas. The following positions the Sun Enterprise 10000 versus the Sun Enterprise 6000 in the target markets.

The Sun Enterprise 10000 system is one of the highest performing SMP system on the market. It offers enhanced scalability and performance in a large-scale, centralized, enterprise server for parallel processing of commercial and technical applications. Commercial parallel and technical applications will use the scalability of the Sun Enterprise 10000 system along with its standard operating environment and commodity hardware components. Also, technical parallel applications rely heavily on the floating point performance of the Sun Enterprise 10000. Commercial parallel applications include large-scale data warehousing, high-volume OLTP, server consolidation, and financial analytic applications.

OLTP customers are now facing high-volume issues associated with providing short response times and high availability for thousands of users. The Sun Enterprise 10000 system addresses these issues by providing mainframe-like RAS capabilities and the ability to handle very high transaction volumes and thousands of concurrent users with a better than two second response time.

Data warehousing customers appreciate the Sun Enterprise 10000 system's ability to provide greater levels of delivered bandwidth where fast query performance is desired. Additionally, the Sun Enterprise 10000 system's large data volume, commodity RDBMS solutions, and mission-critical, high availability make it an even more attractive solution to their needs.

Target Markets and Users for the Sun Enterprise 10000 System *(cont.)*

Because the Sun Enterprise 10000 system supports a larger number of processors, memory, and I/O than the Sun Enterprise 6500 system, the Sun Enterprise 10000 system is recommended for those customers who require 24 or more processors at the time of purchase, or within 18 months of the time of purchase, and expect large, simultaneous I/O and processing operations.

The Sun Enterprise 10000 system offers more than two times the CPU and memory expandability of the Sun Enterprise 6500 server. It delivers the highest level of compute and I/O performance in the Sun Enterprise family of servers. The Sun Enterprise 10000 system should be recommended when the Sun Enterprise 6500 does not offer the necessary amount of growth in CPUs, memory, or I/O bandwidth. In addition, if higher levels of availability are needed compared to the Sun Enterprise 6500, customers should be guided to the Sun Enterprise 10000 system.

Technical Computing Customers

Technical computing customers seeking high performance compute servers are typically not divided by markets or applications, but by response time, room for growth, and cost.

The Sun Enterprise 10000 system has a peak performance of up to 51.2 GFLOPS. Computationally intensive applications, where the Sun Enterprise 10000 system is preferred are those that are highly parallelized or those where large numbers of users are accessing particular applications. Common technical vertical markets include CAD, EDA, petroleum, and computational chemistry to name a few.

For technical customers, the Sun server family is positioned as listed below. For further information, refer to the *High Performance Computing 2.0 Just The Facts*.

- HPC 450: A flexible workgroup server delivering exceptional performance at an affordable price
- HPC 3500: A powerful mid-range server with exceptional availability
- HPC 4500: A highly expandable mid-range server with a compact design
- HPC 5500/6500: Expandable, high-performance systems with mission critical availability and integrated storage
- HPC 10000: Sun's most powerful and highly available server for high-performance computing which incorporates additional RAS capabilities, like dynamic system domains, dramatically increasing system availability for compute intensive applications

Performance

The Sun Enterprise 10000 architecture is designed to offer balanced system performance. These systems feature outstanding integer and floating-point performance, supporting up to sixty-four, 250-MHz, 336-MHz, or 400-MHz UltraSPARC CPUs with 4 MB external caches. The Gigaplane-XB runs up to 12.8 GB per second. High-speed networking is supported by 10/100-Mb Ethernet, FDDI, ISDN, token ring, and ATM interfaces. Fast I/O capability is supported through 64-bit SBus, fast/wide SCSI, UltraSCSI, and fibre channel arbitrated loop (FC-AL) interfaces. Optionally available for selected uses is the PCI I/O bus. This can support 32-bit or 64-bit-wide adapters at a clocking frequency of 33 MHz or 66 MHz. The table below shows a performance and feature comparison of the Sun Enterprise 6500 and the Sun Enterprise 10000 systems.

Performance Type	Sun Enterprise 6500	Sun Enterprise 10000 System
Processor performance • SPECint_rate95 (4 MB Ecache) • SPECfp_rate95 (4 MB Ecache)	3480/30 (336 MHz) CPUs 3021/30 (336 MHz) CPUs	4945/64 (250 MHz) CPUs 6013/64 (250 MHz) CPUs
TPC-D benchmark (300 GB) Informix AD/XP 8.21UD1	QppD: 3270.6 $/QphD: $1553 (24 processors)	QppD: 8113.2 $/QphD: $1400 (64 processors)
TPC-D benchmark (1000 GB)	QppD: 12,931.9 $/QphD: $1,353 Informix IDS AD/XP (96 processors)	QppD: 8870.6 $/QphD: $1508 Oracle 8 v8.0.4.2 (64 processors)
Sustained system bus throughput	2.68 GB per second	12.8 GB per second
Memory latency	approximately 300 ns	approximately 500 ns
Networking performance	Up to 622 Mb per second	Up to 622 Mb per second
I/O performance • SBus • PCI	3–45 SBuses 100 MB per second sustained 2–12 PCI buses 132 to 528 MB per second	2–32 SBuses 100 MB per second sustained 0–32 PCI buses 132 to 528 MB per second
SCSI performance	20 MB per second	20 MB per second
UltraSCSI performance	40 MB per second	40 MB per second
Fibre channel arbitrated loop	100 MB per second	100 MB per second

Sun microsystems

Markets and Applications

The following chart illustrates how the Sun Enterprise 10000 system fits into the current line of Sun server products.

Product	Positioning	Applications	Markets
Sun Enterprise 10000 System	Enhanced scalability, availability, and performance in a large-scale, mission-critical, centralized, enterprise server for commercial and technical parallel processing applications.	• Data warehousing • Data mining • Business applications • Customer management systems • High-volume OLTP • Engineering • Design automation • Analytics/ commercial compute intensive • Inter/Intranet • LAN consolidation	• Manufacturing • Finance • Telecommunications • Government • Education • Health care • Retail • Oil and gas • Pharmaceuticals • Chemical • Internet commerce
Sun Enterprise 6500 Server	High-end scalable and expandable Sun server, offering the performance and availability required for mainframe-class, mission-critical applications	• Data warehousing • Data mining • Business applications • Customer management systems • OLTP • NFS™ • Design automation • Analysis and simulation • Video	
Sun Enterprise 5500 Server	Affordable data center system designed to deliver high performance and high availability for enterprise-wide applications supporting thousands of users		

Specifications of the Sun Enterprise 5500, 6500, and 10000 Systems

Specifications	Sun Enterprise 5500	Sun Enterprise 6500	Sun Enterprise 10000 System
Packaging	Rack	Rack	Rack
Number of CPUs **Clock Rate**	1–14 250, 336, or 400 MHz	1–30 250 or 336 MHz	4–64 250, 336, or 400 MHz
Ecache per CPU	4 MB @250 4 MB @336 4 MB @400	4 MB @250 4 MB @336	4 MB @250 4 MB @336 4 MB @400
Maximum memory	14 GB	30 GB	64 GB
System bandwidth	2.6 GB per second	2.6 GB per second	12.8 GB per second*
Maximum SBus slots	21	45	64
Maximum PCI slots	12	12	32
Maximum internal disk	509.6 GB	382.2 GB	288 GB
Maximum total disk	greater than 6 TB	greater than 10 TB	greater than 60 TB
RAS Features	• Hot-plug boards • Hot-swap power and cooling • Redundant power • ASR • Remote control • ECC-protected data paths • ECC memory	• Hot-plug boards • Hot-swap power and cooling • Redundant power • ASR • Remote control • ECC-protected data paths • ECC memory	• On-line hot swap of boards, power, and cooling components • Fault-tolerant power and cooling • Redundant AC line cords and breakers • Monitoring tools • Automatic system recovery • Domains • ECC on memory and Interconnect • Complete parity checking • Environmental monitoring • Remote console support • Redundant consoles • Interconnect data path resiliency • Interconnect address path resiliency • Redundant "housekeeping" functions • Redundant option for all hardware components
Operating System	Solaris 2.6	Solaris 2.6	Solaris 2.6 (May 1998)
Warranty	one year (hardware) four hours on site	one year (hardware) four hours on site	one year (hardware and software) four hours on site

* Shipments with the 83.3-MHz interconnect and 250-MHz or 336-MHz CPUs will have a 10.6-GB-per-second peak system bandwidth; shipments with 100-MHz interconnect and 400-MHz CPUs will achieve a 12.8-GB-per-second peak system bandwidth.

Channels and Support

The Sun Enterprise™ 10000 uses the same selling channels as the rest of the Sun server line: direct and indirect worldwide. The principal support provider is Sun Enterprise Services, which uses all their standard mechanisms for the product. The Sun Enterprise 10000 warranty level is one year for the hardware and software at the SunSpectrum℠ Gold℠ level. Installation of the system, ServerStart℠, is included in the purchase price.

Key Selling Factors

- **Expandability.** Sun's Enterprise servers expand from entry-level configurations to system configurations that can handle TB of data and thousands of users. The Sun Enterprise 10000 system is configured from 4 to 64 CPUs, 512 MB to 64 GB of memory, and to over 60 TB of on-line disk storage. There are no slot trade-offs between processors, memory and I/O.

- **Scalability.** The Sun Enterprise 10000 system is highly modular. Customers can easily configure these systems to meet their application and performance requirements by simply adding UltraSPARC™ modules, memory, or I/O boards. The high-throughput Gigaplane-XB™ and I/O architecture eliminates system bottlenecks and provides balanced system performance, even in systems with the maximum number of UltraSPARC modules and I/O devices.

- **Investment Protection.** All of the 250-MHz, 336-MHz, or 400-MHz processor modules, SIMMs, and SBus boards used in the Sun Enterprise 3500, 4500, 5500, and 6500 servers are common to the Sun Enterprise 10000 system. Therefore, when upgrading to the larger Sun Enterprise 10000, customers can move these components from the existing chassis to the new chassis, protecting their investment. The Sun Enterprise 10000 system uses the same peripherals in the same expansion cabinets as the rest of the family.

- **Solaris™ applications**. The Sun Enterprise 10000 runs the standard Solaris Operating Environment. Therefore all 12000-plus Solaris applications are binary compatible and will run on the Sun Enterprise 10000 without any conversion.

- **Upgrade program**. There is a trade-in program available to move customers to the Sun Enterprise 10000 from Sun's other servers and from selected servers from Sun's competitors.

- **Upgradability.** The modular design of the Sun Enterprise 10000 system means easy upgrading to new technologies and higher performance. The Sun Enterprise 10000 system will support future generations of UltraSPARC processors, disk arrays, tape devices, SBus cards, and networking interface cards.

Key Selling Factors *(cont.)*

- **Reliability, availability and serviceability features that result in uptimes greater than 99.95 percent**
  - No single points of hardware failure: No single component (with the exception of the control board) will prevent a properly configured Sun Enterprise 10000 system from automatically reconfiguring itself to resume execution after a failure.
    - Achieved through a combination of redundancy and alternate pathing architecture.
  - Error correction interconnect: Data and address buses are protected by a combination of error correcting codes and parity.
  - Dynamic system domains: Groups of system boards can be arranged in multiprocessor system domains that can run independent copies of the Solaris Operating Environment concurrently.
    - Each system domain is completely isolated from all software errors, and most hardware failures that might occur in another system domain.
  - Dynamic reconfiguration: Enables the system administrator to add, remove, or replace system components or create/remove system domains on line without disturbing production usage
  - Hot swapping: Power supplies, fans, and most board-level system components can be exchanged while "hot"; that is, while the system is on line and in operation.
- **Manageability.** Using Network Console (netcon) over the network, system administrators can remotely login to the SSP to control the Sun Enterprise 10000 system.

Selling Highlights

Technology

Four principal areas of technology used in the design of the Sun Enterprise™ 10000 give Sun a significant competitive advantage. They are:

- **The UltraSPARC™ microprocessor family.** This is a high-performance 64-bit processor with features that allow workstations and servers to compute fast.

- **Custom ASICs.** These represent a huge investment in time and money. The benefits compared to discrete logic are: faster internal speed, improvements in system availability, and lower manufacturing cost. The Sun Enterprise 10000 has three ASICs common to Sun's other servers (the SPARC™ microprocessor, the data buffer on the processor module, and the SBus I/O chip), as well as seven designs that are custom to the product.

- **Enormous system bandwidth.** The Sun Enterprise 10000 system uses a crossbar router instead of a bus to interconnect processors, memory and I/O. System scalability and low latency are a function of having sufficient internal bandwidth. A router's bandwidth scales up as system hardware is added which is exactly what one wants. The crossbar router is packaged on the centerplane. Its manufacture requires use of state of the art manufacturing processes and procedures.

- **The Solaris™ Operating Environment.** Without a stable and well-proven operating system, the best hardware in the world is useless. The Solaris Operating Environment has been enhanced over the past few years to be able to address very large memories and to scale up the 64 processors—both important features for the Sun Enterprise 10000.

Introduction

The Sun Enterprise™ 10000 system is a shared-memory SMP computer that can be configured with up to 64 UltraSPARC™ processors, 64 SBus boards (or 32 PCI boards), and 64 GB of on-line memory. The Sun Enterprise 10000 system is comprised of system boards, a centerplane, centerplane support boards, control boards, peripherals and power and cooling subsystems. These components and their relationships are illustrated in Figure 2 below and their functions are listed in the following table.
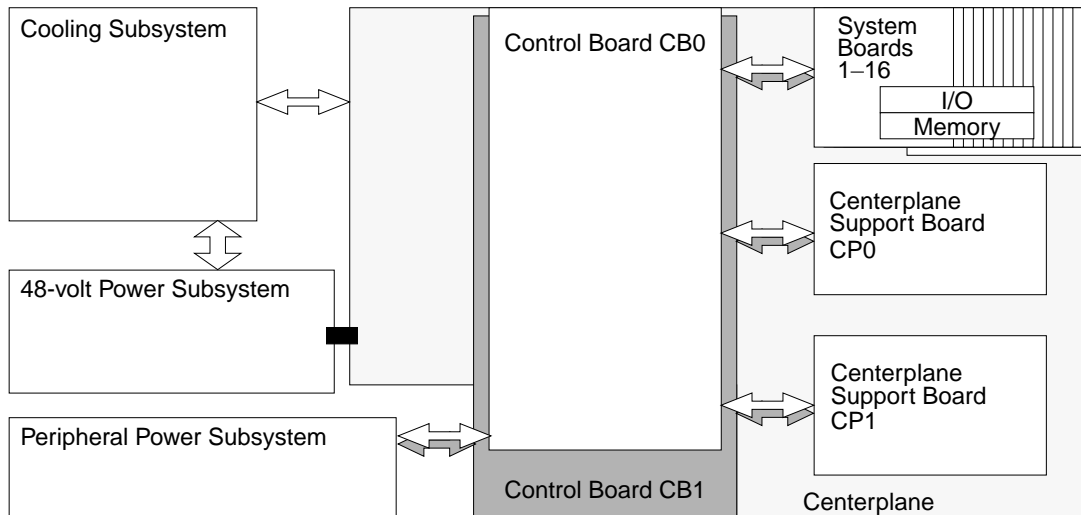

Figure 2. System block diagram

Introduction *(cont.)*

Component	Function	Quantity
System board	Contains processors, memory, I/O subsystem, SBus boards, and power converters	up to 16
Processor modules	Mezzanine boards that contain the UltraSPARC processor and support chips	up to 64
Memory	Removable SIMMs	up to 512
I/O	Removable SBus or PCI boards	up to 64
Control board	Controls the system's JTAG, clock, fan, power, serial interface, and Ethernet interface functions	up to 2
Centerplane	Contains address and data interconnect to all system boards	1
Centerplane support board	Provides the centerplane's JTAG, clock, and control functions	up to 2
48-volt power system		
AC power controller	Receives 220 VAC, monitors it, and passes it to the power supplies	3 or 4
Power supply	Converts AC power to DC	5 or 8
Circuit breaker panel	Interrupts power to various components within the system	1
19-inch rack-mount power sequencer	Receives 220 VAC, monitors it, and passes it to the peripherals. This unit can be operated in either master or slave mode allowing the on/off function to be controlled by another power sequence.	1 or more
Remote Power Control Unit	Connects the remote control line between two control boards and passes it to one or more power sequencer units	up to 5
Fan centerplane	Provides power to the pluggable fan trays	2
Fan Trays	Contains two fans for system cooling	10 or 16

Sun
microsystems

System Interconnect

The Sun Enterprise 10000 system uses the Gigaplane-XB™ interconnect which adheres to the Ultra™ Port Architecture (UPA) standard. A combination of improvements have been utilized to increase interconnect bandwidth over previous generation bus-based systems. This amount of bandwidth is enough to keep memory latency nearly constant for data-intensive processing on full 64-processor configurations— with some headroom left over for faster processors in the future.

The UPA bus, the primary bus for the Ultra 1 desktop workstations and Sun Enterprise servers, is used as an intermediate bus to connect CPU/memory boards and I/O boards to the Gigaplane™ bus. The UPA bus runs at 83.3 MHz, with a peak bandwidth of 1.3 GB per second.

The following design elements increase system throughput and reduce memory latency:

- **The Gigaplane-XB interconnect uses separate address and data lines:** The UPA defines a separate address and data interconnect. Usually on a bus-based system, only about 70 percent of the wire bandwidth is available for data, with the rest being used for address and control. Separating the functions lets both addresses and data each have 100 percent of the wire bandwidths on their separate paths, and lets the wire topology of each function be optimized differently. Snoop addresses need to be broadcast simultaneously to all the boards, while data packets can be sent point-to-point.

  Figure 3 illustrates how the Gigaplane-XB interconnect is used to transfer four 16-byte blocks of data from the memory of one system board to a single 64-byte block of cache memory on the processor module of another system board. In the Sun Enterprise 10000 system, each system board is connected to all other system boards via the Gigaplane-XB interconnect.

- **The Gigaplane-XB interconnect datapath width is 16 bytes.**

	Sun Enterprise 10000 System	Sun Enterprise 5500 and 6500
Memory data bus	576 bits	576 bits
Data bus	144 bits per board	288 bits
CPU data bus	144 bits	144 bits

- **Sixteen data paths:** To meet the Sun Enterprise 10000 system's bandwidth goals, the 16 data paths allow a separate connection to each board.

- **The Sun Enterprise 10000 system contains four snoop paths:** Sixteen data paths require sufficient performance on the address bus to achieve maximum system performance. The Sun Enterprise 10000 system uses four snoop paths to supply enough address bandwidth to match the data bandwidth.

- **Point-to-point wires versus multi-drop buses:** In a multi-drop bus, all the processors, I/O devices, and memory modules attach to a single set of wires. As the number of connections rises, the clock rate must be lowered to maintain reliability in the face of increasing electrical load. A failure of any component on the bus may bring down the entire bus, not just the connections to the failing component.

- **Multistage address and data routers:** The Sun Enterprise 10000 system has a two-stage routing topology based upon the physical board partitioning. Local many-to-one routers gather on-board requests, and connect them to one off-board port. A global-data crossbar connects one port from each board together. Four point-to-point address buses broadcast addresses to all the boards.

- **The system clock rate is designed for 100 MHz:** The UltraSPARC-I processor requires the system clock rate and the processor clock rate be an integer multiple. The initial Sun Enterprise 10000 system used an 83.3-MHz system clock and 250-MHz processors (a 3X ratio). Sun sells 336-MHz processors that run at 333-MHz (a 4X ratio). However, the design is capable of using a 100-MHz system clock that is utilized with the addition of the 400-MHz processors.

System Board 11
System Board 12
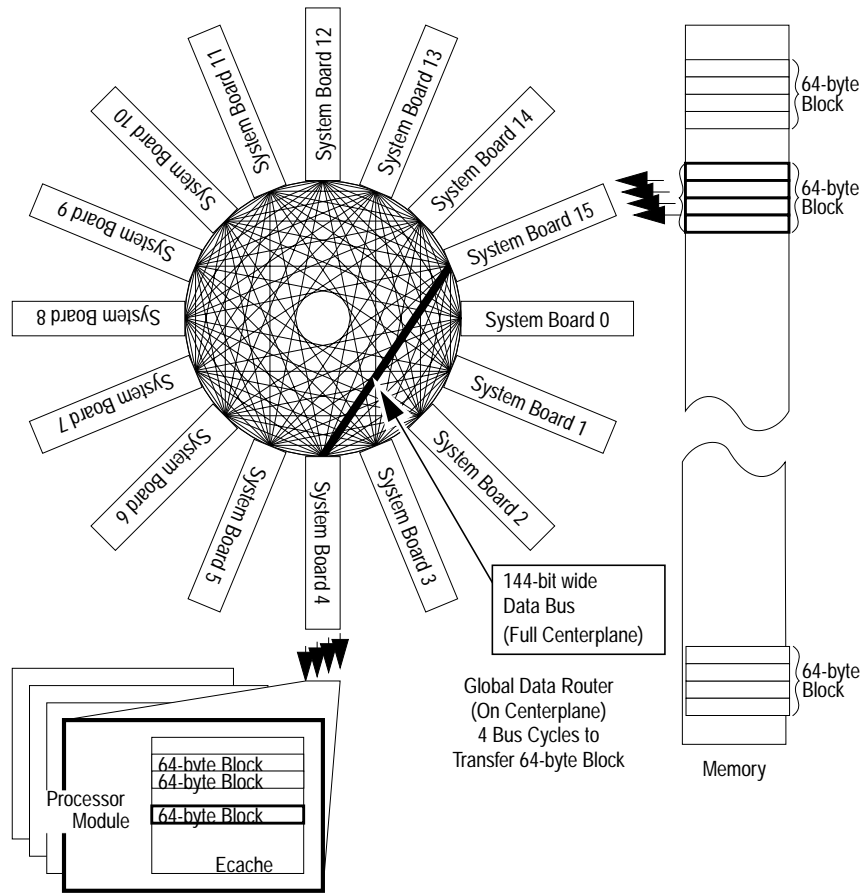System Board 13
System Board 10
System Board 14
System Board 9
System Board 15
System Board 8
System Board 0
System Board 7
System Board 1
System Board 6
System Board 2
System Board 5
System Board 3
System Board 4

64-byte Block

64-byte Block

64-byte Block

Memory

144-bit wide
Data Bus
(Full Centerplane)

Global Data Router
(On Centerplane)
4 Bus Cycles to
Transfer 64-byte Block

Processor
Module

64-byte Block
64-byte Block
64-byte Block

Ecache

Figure 3. Data Routing

System Architecture

System Interconnect *(cont.)*

Figure 4 shows the System Board architecture. Data routing in the Sun Enterprise 10000 system is conducted at two levels: global and local. The global data router (located on the centerplane) is an 18-byte wide, 16 x 16 crossbar that steers data packets between the 16 system boards. With the 16 x 16 crossbar, any port can be connected to any other throughout the centerplane. Of the 18 bytes, 16 are for data and the remaining 2 bytes are for error correction.

Address routing on the Sun Enterprise 10000 is implemented over a separate set of four global address buses. Although called "address buses" to convey that addresses are broadcast, the implementation is as a point-to-point router. The significance of this is that routers have more inherent reliability than a bus. The buses are 48 bits wide including error correcting code bits. Each bus is independent, meaning that there can be four distinct address transfers simultaneously. An address transfer takes two clock cycles, equivalent to a snoop rate of 167 million snoops per second on all four address buses. Should an uncorrectable failure occur on an address bus, degraded operation is possible using the remaining buses.
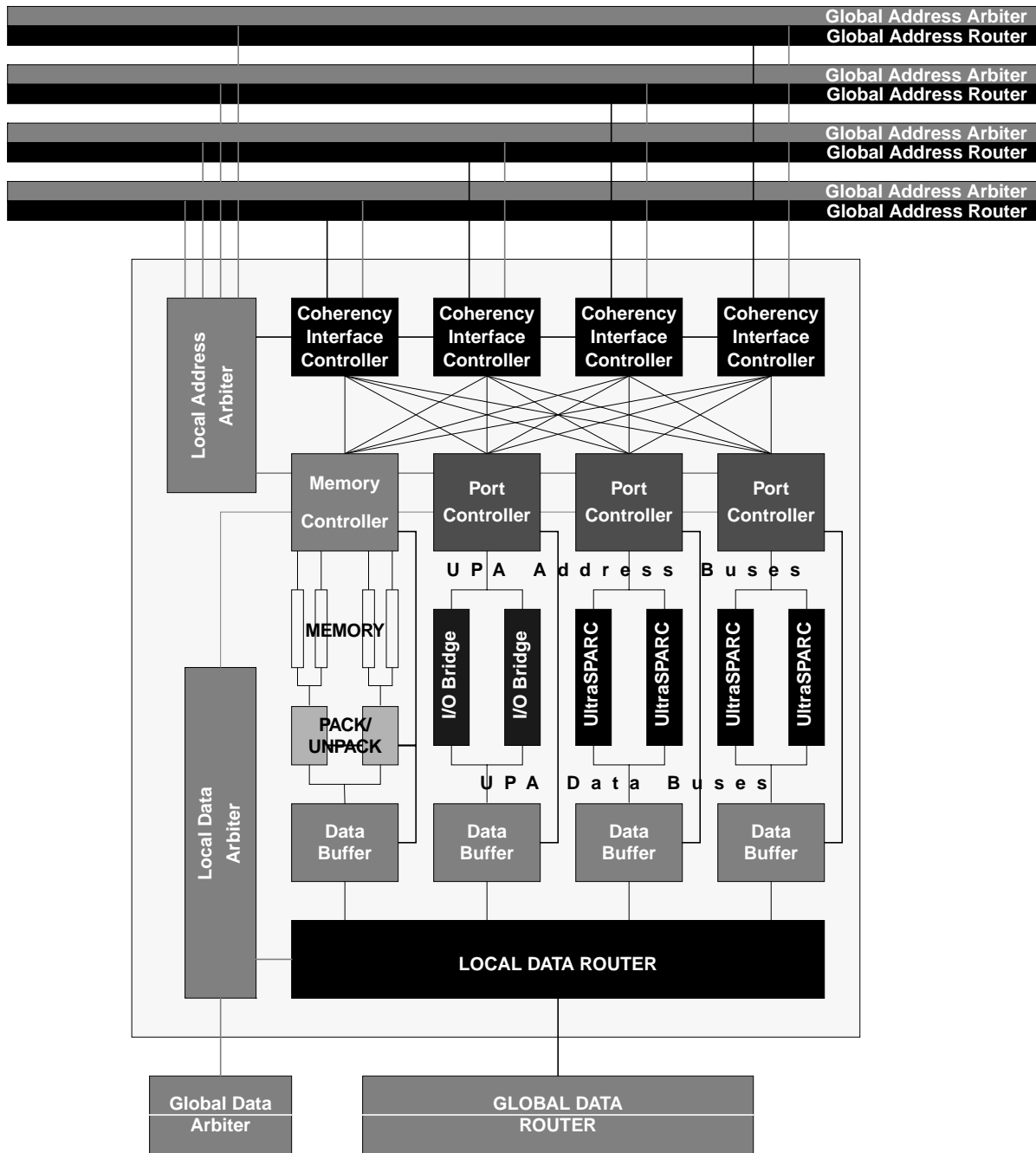
System Architecture

Figure 4. System Board Architecture

System Boards (with SBus I/O)

The Sun Enterprise 10000 system consists of multiple system boards (refer to Figure 5) interconnected by a centerplane. A single system cabinet holds up to 16 of these system boards, each of which can be independently configured with processors, memory, and I/O channels, as follows:

- One-to-four 250-MHz, 336-MHz, or 400-MHz UltraSPARC microprocessor modules per system board. Processor clock frequencies may not be mixed within any one Sun Enterprise 10000 system.

- Four memory banks with a capacity of up to 4 GB per system board (64 GB per system). Each memory bank consists of eight SIMMs. Sun supplies low density (32-MB) SIMMs or high density (128-MB) SIMMs. Therefore a fully populated system board will have 1 GB or 4 GB of storage. System boards can have just two banks of SIMMs giving 512 MB or 2 GB of storage.

- Two SBuses per module each with slots for up to two adapters for networking and I/O (32 SBuses or 64 slots per system)

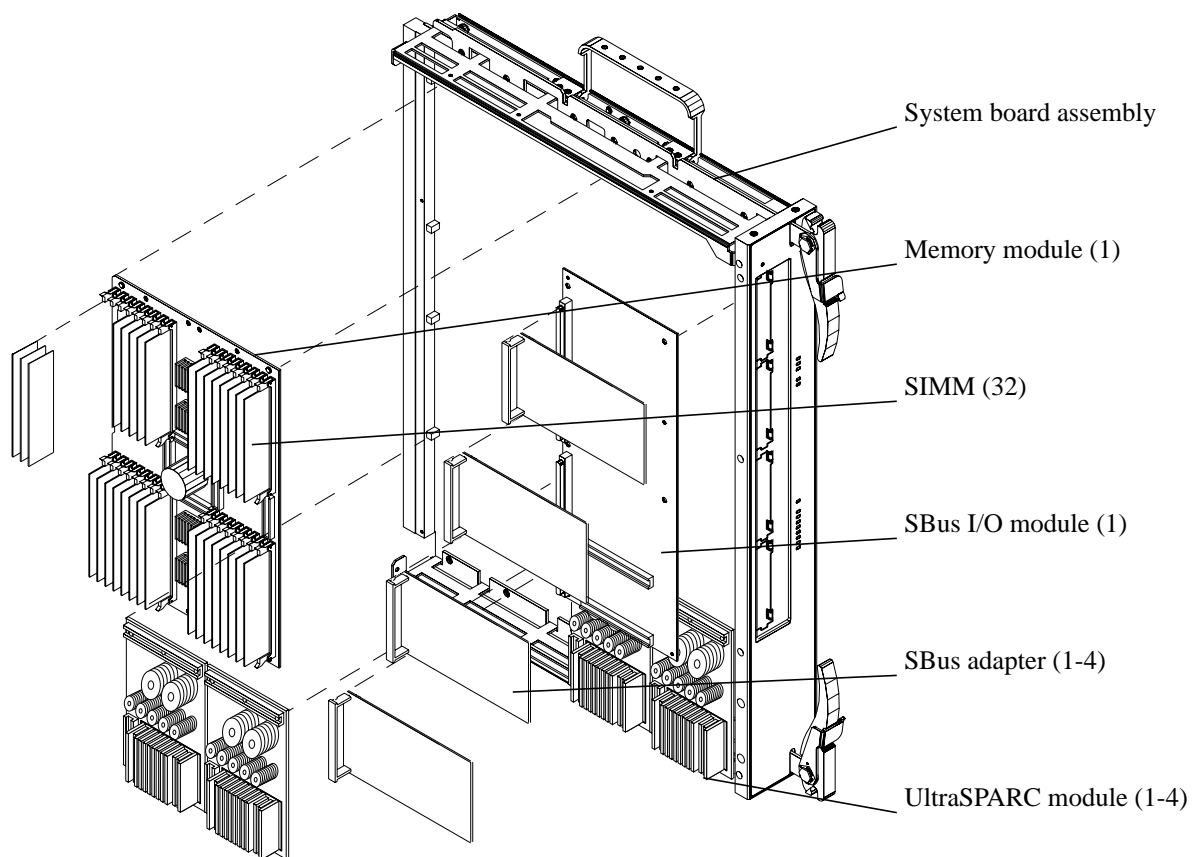The mechanical assembly of the System Board is as shown in

Figure 5. System Board Assembly (with SBus I/O)

CPU/memory boards have temperature sensors located under the UltraSPARC modules. This allows the actual temperature of individual boards to be monitored through the SSP's GUI called Hostview.

UltraSPARC™ Processor Module

The Sun Enterprise 10000 system houses up to 64 UltraSPARC processors which can execute four instructions per clock cycle. The processor mounts on a small daughterboard, the UltraSPARC module, which also houses the 4-MB, second-level cache and the UltraSPARC data buffer (UDB) circuitry. The second-level cache handles cache misses from the processor's on-chip data cache memory. In total, the architectural elements on the processor chip and the module support the Sun Enterprise 10000 system's ability to execute two floating point instructions, add or subtract, and two integer instructions during a single clock cycle.

Memory Subsystem

Large-scale systems must provide sufficient memory capacity to sustain high performance from the processors and I/O channels. Additionally, memory must be quickly accessible in order to avoid interfering with other subsystem activities. Finally, the large concentration of data in today's data center systems and production environments necessitates a highly reliable design. The Sun Enterprise 10000 system is designed to meet all of these requirements. Using currently-available 64-Mbit DRAM chips, a fully configured system offers 64 GB of system memory.

The memory in the Sun Enterprise 10000 system is located on the memory board mounted as a daughter board on the system board. Up to 4 GB of RAM can be installed on each system board. The memory subsystem in the Sun Enterprise 10000 system is designed to offer fast, reliable data access.

- The memory controller manages four banks of memory on each memory module. Each bank of memory consists of eight standard JEDEC SIMM modules, implemented in 3.3v CMOS.

- The Sun Enterprise 10000 system supports up to eight way memory interleave, but normally only four-way interleaving is used. Going beyond this would not allow Dynamic Reconfiguration to be used.

- The interleaved memory banks can be different sizes on different system boards.

- The Solaris Operating Environment has been enhanced to provide scalability consistent with this memory capacity

The entire memory data path is protected by ECC mechanisms, and SIMM organization is specifically designed so each DRAM chip contributes only one bit to one half-byte of data. In this way, the failure of a DRAM chip will result in correctable memory errors in four successive words.

System Architecture

I/O Subsystem, Devices and Networking (for SBus)

The Sun Enterprise 10000 I/O module is a mezzanine card that plugs into the system board and connects the UPA to a pair of SBuses. Each of these SBuses can in turn be populated with one or two single-width, SBus adapters, or one double-width.

- Incrementally expandable I/O by configuring up to 64 SBus slots on 32 independent buses
- Each SBus interface includes its own memory management unit to translate between virtual and physical addresses
- SBus supports 32-bit or 64-bit data transfers
- The following disk/tape adapters and peripherals are available on the Sun Enterprise 10000 system:
  - SCSI adapter: For connections to discrete disks or tape devices
  - Fiber adapters for the Sun StorEdge™ A5000 fiber channel arbitrated loop (FC-AL) array
  - UltraSCSI adapter: For connection to the Sun StorEdge A3500 and D1000 arrays
- The following network adapters are supported by the Sun Enterprise 10000 system:
  - Ethernet adapter: Basic networking at 10 or 100 Mb per section with one or four 10BASE-T ports per board. There is auto-speed detection.
  - Gigabit Ethernet
  - HiPPI for HPC applications
  - FDDI: Faster networking at 100 Mb per section over fiber cabling
  - ATM: This adapter allows the Sun Enterprise 10000 system to interface to 155 or 622 Mb per section asynchronous transfer mode, local or wide area networks
  - ISDN to allow connection to wide-area networks (WAN) that support this standard
  - Token Ring to allow the Sun Enterprise 10000 system to participate in mainframe networks
  - High-speed, serial interface for attachment to traditional wide-area networks
- A wide range of peripheral options are available for the Sun Enterprise 10000 system. These include:
  - The Sun StorEdge A7000 array (capacity up to 2.93 TB)
  - The Sun StorEdge A5000 FC-AL array (capacity up to 127.4 GB)
  - The Sun StorEdge A3500: A high availability hardware RAID solution with up to 720 GB of storage
  - The Sun StorEdge D1000 array (capacity up to 144 GB)
  - The Sun UniPack disks (mounted in the system cabinet only) for booting the Solaris Operating Environment
  - Tape devices, including EXABYTE 8-mm tape drive and digital linear tape
  - Sun Enterprise Tape Library™: Stores up to 3.5 TB of uncompressed data

System Boards with PCI I/O

PCI is an open I/O standard to which Sun products are moving. The main advantage of using PCI, in a server application, is the higher speed of PCI when compared to SBus. PCI adapters are available with 32-bit or 64-bit data paths and run at a clock frequency or 33 MHz or 66 MHz.

System Boards for the Sun Enterprise 10000 are available with a PCI module in place of the standard SBus module. This PCI module has two 66-MHz buses and each can accommodate one PCI adapter. (These adapters are the 6.5-inch cards, not the 11-inch cards).
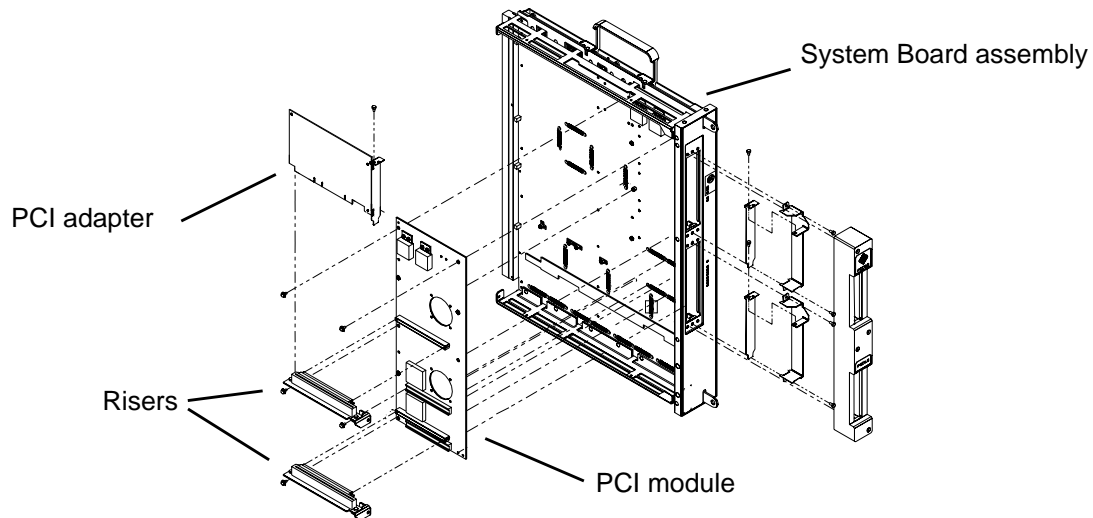

Figure 6. System Board with PCI module

The above figure shows how the PCI module is mounted to a system board. The "risers" allow the PCI adapters to be mounted in the same plane as the PCI module.

Because it is only possible to package two PCI adapters per system board (in contrast to four SBus adapters), PCI is not cost effective where there is not a performance requirement. Therefore the Sun Enterprise 10000 will remain basically as an SBus-based system with PCI available for selected uses. For instance, customers will prefer to use SBus for the interfaces detailed in the previous section (for example, SCSI, Ethernet, fiber channel, FDDI, and ATM). PCI will be used for fast data transfer situations such as HIPPI. PCI will also be used for customer-supplied adapters.

Reliability, Availability and Serviceability (RAS)

Strategy

The Sun Enterprise 10000 offers excellent reliability, availability, and serviceability (RAS). These RAS features result in the Sun Enterprise 10000 being the highest fault-resistant systems in its class. Customers want the highest possible uptime. Reliability and Maintainability are features designed into the system for delivering the greatest possible uptime ("Availability").

The following is a list of the RAS features found in the Sun Enterprise 10000 system.

Reliability

- Current-sharing power circuitry supports redundant power capability.

- ECC-protected data throughout the system increases data integrity.

- Parity-protected address and control signals increase the integrity of those signals.

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

- Built-in, self-test logic in all the ASICs applies pseudo-random patterns at system clock rate providing at least 80 percent single-stuck-at-fault coverage of combinatorial logic.

- The power-on self-test, controlled from the SSP, tests each logic block first in isolation, then with progressively more and more of the system. Failing components are electrically isolated from the centerplane. The result is that the system is booted only with logic blocks that have passed this self-test and which should operate without error.

- All Sun Enterprise 10000 ASICs have *paranoid* logic which checks for anomalous conditions indicating an error has occurred, such as queue overflows, invalid internal states, and missing events, rather than let the error propagate and become corrupted data or access timeouts that would be difficult to correlate with the actual failure.

- The Sun Enterprise 10000 system uses a highly reliable distributed power system. Each system, control, or centerplane support board within the system has DC-to-DC converters for that board only, with multiple converters for each voltage.

- The internal temperature of the system is monitored at key locations as a fail-safe mechanism. If an over-temperature threshold is reached on a System Board, that board is excluded from its domain following an auto-reboot. Other domains (if any) do not require a reboot.

- Tachometers detect that the cooling system is moving air into the system. A failed fan will trigger the SSP to log a warning message.

- Additional sensing is performed by the Sun Enterprise 10000 system in order to enhance the reliability of the system by allowing constant "health" checks. DC voltages are monitored at key points within the Sun Enterprise 10000 system and DC current from each power supply is monitored and reported to the SSP.

- The reset signals in the Sun Enterprise 10000 system are sequenced with the DC power levels in order to guarantee stability of voltage throughout the cabinet prior to removing reset and allowing normal operation of any of the Sun Enterprise 10000 system's logic.

Reliability, Availability, and Serviceability (RAS) Features *(cont.)*

Availability

- Sophisticated system diagnostics minimize downtime.

- Multiple UltraSPARC modules provide redundancy.

- Fan trays each have two fans. Should one of a pair fail, the survivor provides sufficient cooling. A warning message is logged.

- Remote administration control allows remote reboots and power-cycling.

- Redundant components can be added to augment the system's already high reliability and availability. There are no components in the system which cannot be configured redundantly if the customer so desires.

- Intelligent SSP identifies system and component errors and then takes corrective action.

- During an automatic reboot, the system uses power-on self-test (POST) to automatically reconfigure around a hard failure prior to bringing the system up.

- Each side of the centerplane has its own 48V distribution bus, and each system board develops its own low-voltage supplies locally with on-board regulators. Should a regulator fail, the system adapts automatically by reconfiguring itself to exclude the offending board.

- The base Sun Enterprise 10000 system uses three line cords, each fed by a separate 220V, single-phase, 50/60-Hz, AC circuit, to deliver the required input power to the bulk DC supplies. A fourth discrete line cord serves the system's I/O space. This level of redundancy ensures against a system-wide power loss and also reduces the current through any one circuit.

- Multiple operating systems and/or diagnostics can be co-hosted by the hardware using several, independent system domains. This keeps development work isolated from production, thereby improving the production availability.

- For even higher availability, a pair of Sun Enterprise 10000 systems can be configured in a redundant fashion so, should the primary system fail, processing continues with the secondary Sun Enterprise 10000 system. All this is under control of Sun Cluster failover software that effects a rapid and seamless switchover from one machine to the other.

- The Solaris Operating Environment panics and hangs result in an auto-boot of the system.

- In the event of a centerplane data crossbar component failure, one half of the crossbar is disabled and the system will again be operational following a reboot. This holds true for the address router as well.

- Error correction on the interconnect ensures that transient errors do not affect availability.

System Architecture

Reliability, Availability, and Serviceability (RAS) Features *(cont.)*

Serviceability

- Modular system design makes it easy to replace failed components.

- Most hardware maintenance can be performed without taking the system off line; only the components actually being worked on are taken out of service. This uses the Sun Enterprise 10000 server's dynamic reconfiguration (DR) and hot-swap capabilities.

- The ability of Hostview to notify a system administrator of a failure allows the system administrator to know immediately which components have failed and need service.

- All centerplane connections are point-to-point making it possible to logically isolate system boards by dynamically reconfiguring the system.

- Improved remote administration control allows users to reboot and power-cycle in a "lights out" environment.

- Sun Validation Test Suite (SunVTS™) allows users to perform UNIX®-level diagnostics.

The DR capability allows concurrent servicing of the system. It also allows system boards to be upgraded with different processors, more memory or have SBus cards added—all without materially disturbing a production system.

- When uncorrectable errors occur, information about the error is saved to help with further isolation.
  - The Sun Enterprise 10000 system has extensive error logging capabilities

- Connectors are keyed so that boards may not be plugged in upside down

- Special tools are *not* required to access the inside of the system for changing of field-replaceable units.
  - No jumpers are required for configuration of the Sun Enterprise 10000 system.

- Air filters are replaceable while the system is operational.

- The Sun Enterprise 10000 system uses a distributed DC power system with each system board having its own power supply.
  - This type of power system allows each system board to be powered on/off individually.

- All ASICs that interface to the centerplane have a loop-back mode, which allows a system board to be verified before it is dynamically reconfigured into the system.

System Architecture

Sun microsystems

RAS Summary Table

Reliability Features	Availability Features	Serviceability Features
• ECC-protected data	• Redundant UltraSPARC modules	• Modular system design
• Parity-protected address and control signals	• Redundant CPU/memory boards	• Hot-swap system boards
• Current-sharing power circuitry	• Redundant power supplies	• Hot-swap control boards
• Environment monitors and controls	• Twin fans in each cooling unit	• Remote booting and power-cycling
• Connectors, cables, and guides all designed for robustness	• Dual disk array host interfaces	• Hot-swap disk drives
• Point-to-point routers to maintain bus integrity over multi-drop buses.	• Automatic reboot	• Hot-swap power/cooling modules
	• Multiple operating system support using dynamic system domains	• SunVTS
	• Compatible with commercial battery-backup systems	• Several internal self-tests for error reporting
	• Fault-tolerant AC power system	• Dynamic reconfiguration for trouble isolation and repair on line
	• Four independent address buses	
	• Sixteen-by-sixteen data interconnect with two independent routers	

Replacing or Upgrading Concurrently Serviceable Components

Concurrently serviceable components, those that can be removed and replaced while the system is running, include all field-replaceable units except the fan centerplane and system centerplane. Concurrently serviceable components must be configured for redundancy prior to removal to prevent system interruption. This can be done while the system is running.

If an UltraSPARC module, SIMM, SBus board, memory module, I/O module, system board, control board, centerplane support board, power supply, or fan fails, the system attempts to recover without any service interruption. After the failed CPU/memory or I/O board is deconfigured from the system, the failed board may be removed, replaced, and reconfigured into the system, again, while the system is on line. This uses the Dynamic Reconfiguration capability of the Sun Enterprise 10000.

System Architecture