An Oracle White Paper
March 2010

Sun Netra T5440 Server Architecture

ORACLE

Introduction

In a period of intense competition and emerging technologies, telecommunications companies must deliver and support a variety of new voice and data solutions. The integration of wireless, video, multimedia, and messaging solutions—along with escalating growth in user volumes—is forcing underlying IT infrastructures to expand rapidly to meet application demands. In many cases, IT managers must quickly and dramatically scale data center resources to address this growth trend while managing to the tight constraints of operating budgets and available space.

For this reason, Oracle has focused on engineering innovative systems with robust compute, memory, and I/O densities that provide levels of throughput higher—by many orders of magnitude—than what is available in existing systems. The company's focus on ecocentric design has also drastically changed assumptions about power and cooling requirements. In keeping with this focus, Oracle has now introduced a third generation of UltraSPARC processors that leverage chip multithreading (CMT)—UltraSPARC T2 Plus processors with CoolThreads technology—to deliver high throughput within a small power envelope. The UltraSPARC T2 Plus processor design enables multisocketed system implementations, allowing Oracle's Sun Netra T5440 server to provide a total of 128 threads in a single chassis. As a result, the carrier-grade Sun Netra T5440 server delivers breakthrough application throughput with dual processors, supporting twice as many threads as previous-generation systems using a single UltraSPARC T2 processor.

Sun Netra T5440 servers combine the certified ruggedness and reliability of other Sun Netra server designs with the performance and ecoefficiency of UltraSPARC T2 Plus processors. The Sun Netra T5440 system uses a 4U chassis that features remarkable expansion capabilities for memory, I/O, and storage. The resulting server offers a dramatic boost in density and throughput capacity while conserving energy and space, thus lowering operating costs while enhancing application availability.

Managing Complexity

As next-generation technologies for wireless, video, and messaging emerge, the telecommunications marketplace is becoming increasingly competitive. Network equipment providers (NEPs) and service providers are placing increasing demands on the IT infrastructure to support expanding voice and data services and skyrocketing transaction volumes. At the same time, many providers are moving toward a consolidated framework to converge telecom, internet, and multimedia services using a single, open standards–based architecture. These trends of growth and consolidation put additional strain on existing IT infrastructures, which strive to deliver growing service volumes within the constraints of budget and time to market.

Even as demand increases for an expanded range of services, economic realities are reshaping the IT infrastructure. Real-estate constraints and rapidly rising energy costs are discouraging the once-common practice of adding racks upon racks of servers. What's more, the cost and complexity of managing large numbers of systems create formidable challenges, especially because low utilization levels sometimes result, reducing return on investment. Exacting service-level agreements (SLAs) are increasingly common because service availability is frequently linked to economic success. In managing the IT infrastructure, organizations strive to control capital and operational costs more precisely, making deployment decisions that take into account overall system lifecycle and total cost of ownership (TCO).

Introducing the Sun Netra T5440 Server

To help telecommunications providers address the challenge of scaling throughput while managing TCO and service levels, Oracle offers the Sun Netra T5440 server—a robust carrier-grade system that features multiple powerful, energy-efficient UltraSPARC T2 Plus processors. Providing extreme reliability and high throughput, the server also offers unprecedented processor core, memory, storage, and I/O densities to enable dramatic expansion and configuration flexibility within a compact 4U form factor.

CoolThreads technology in the UltraSPARC T2 Plus processor helps deliver high throughput for demanding telecommunications applications while minimizing power consumption. (The next major section—"The UltraSPARC T2 Plus Processor and the Evolution of Throughput Computing"—describes the evolution of CMT technology and explains how CMT improves application throughput while maintaining low power consumption.) The UltraSPARC T2 Plus processor in the Sun Netra T5440 server is a third-generation CMT processor design. It leverages the same CoolThreads technology proven in Oracle's Sun Netra T2000 and Sun Netra T5220 servers (which use previous-generation UltraSPARC T1 processors and UltraSPARC T2 processors, respectively). Dual UltraSPARC T2 Plus processors in the Sun Netra T5440 server—each supporting eight cores and 64 concurrent processor threads—take throughput-per-watt to a new level, doubling potential throughput of the earlier Sun Netra T5220 server. The UltraSPARC

T2 Plus processor design adds cache coherency capabilities to support multisocketed implementations, allowing Sun Netra T5440 servers to address extremely throughput-intensive workloads. Combined with breakthrough levels of expansion for memory, storage, and I/O, these carrier-grade servers form an ideal solution for consolidation initiatives that address growing application and user volumes.

In addition to delivering breakthrough levels of throughput and scalability, the Sun Netra T5440 server is certified to meet Telecordia Network Equipment Building Specification (NEBS) Level 3 requirements for operating environments that demand continuous availability and simplified management. NEBS Level 3 certification indicates that the system has undergone a series of rigorous, independently conducted tests to prove that it can continue to operate under severe environmental conditions.

Key Features

With a design that complements the rest of Oracle's carrier-grade server family, Sun Netra T5440 servers are engineered to address the needs of today's emerging telecommunications applications, including

- **Efficient application scalability.** With support for 128 threads and large memory capacities, Sun Netra T5440 servers take advantage of the cache coherency logic, floating-point, and I/O capabilities built in to the UltraSPARC T2 Plus processor chip. The highly threaded design contributes to leading levels of scalability while maintaining high throughput efficiency.
- **Leading ecoresponsibility.** Sun Netra T5440 servers continue a tradition of ecoresponsibility by offering optimal performance—and performance per watt—across a wide range of workloads. The UltraSPARC T2 Plus processor incorporates unique power management features at both processor core and memory levels. High-efficiency power supplies and an innovative chassis design (which optimizes airflow) minimize the need for cooling, lowering energy-related costs.
- **Continuous system availability.** Reliability features help to increase availability, reduce costs, and meet SLA targets—even in harsh operating environments. NEBS Level 3 certification demonstrates that the system has been rigorously tested and can withstand severe operating conditions. With high levels of integration provided by the UltraSPARC T2 Plus processor design, the Sun Netra T5440 server has a relatively small part count and is designed for high reliability, availability, and serviceability (RAS). Redundant, hot-swappable DC or AC power supplies and disk drives enhance reliability and serviceability. Low power consumption reduces generated heat loads and the potential for temperature-related problems.
- **Simplified management.** Each Sun Netra T5440 server includes Oracle Integrated Lights Out Manager, which provides a command-line interface (CLI), a Web-based graphical user interface (GUI), and Intelligent Platform Management Interface (IPMI) functionality to support out-of-band monitoring and administration.

- **Industry-leading tools for virtualization and consolidation.** The Sun Netra T5440 server is ideal for consolidation, providing low-level multithreading support for virtualization at multiple layers of the technology stack. Oracle VM Server for SPARC (previously called Sun Logical Domains) exploits the thread-intensive model of the UltraSPARC T2 Plus processor by allowing multiple instances of Oracle Solaris to run concurrently, while Oracle Solaris Containers provide virtualization within a single Oracle Solaris.
- **Zero-cost security.** With electronic intrusion and theft at all-time highs, secure communications and data protection have never been more important. Because each UltraSPARC T2 Plus processor features up to eight integrated cryptographic accelerators, there's no need to transmit plain text on the network or to store plain text in storage systems. Sun Netra T5440 servers can support more cryptographic operations per second than many competitive systems and dedicated cryptographic accelerator cards, with minimal impact to system overhead.
- **Accelerated time to deployment.** Sun Netra T5440 servers running Oracle Solaris 10 provide full binary compatibility with earlier UltraSPARC systems, preserving investments and speeding time to deployment. Sun Netra T5440 servers come preinstalled with Oracle Solaris 10 to expedite deployment and provide a robust and secure operating environment.

Table 1 summarizes the features of this new system.

TABLE 1. SUN NETRA T5440 SERVER FEATURES

FEATURE	SUN NETRA T5440 SERVER
Chassis size (in rack units)	4U
Number of CPU sockets	2
CPU	UltraSPARC T2 Plus processor at 1.2 GHz with 4, 6, or 8 cores
Threads	Up to 128 threads (2 CPUs, 8 cores per CPU, 8 threads per core)
Number of cores and threads per 40U rack enclosure	Up to 160 cores (1,280 threads) per rack
Memory capacity	32 slots with up to 128 GB (using 4-GB FB-DIMMs)
Memory type	667-MHz FB-DIMM
Internal storage	Up to twelve 2.5-inch SAS HDDs
Removable media	Optional DVD-RW
PCI expansion	• 2 PCI-X (2 full-length/full-height, 64-bit @ 133 MHz)

	• 6 PCIe (2 full-length/full-height, 4 MD2 LP) • 2 PCIe/XAUI combination slots
Gigabit Ethernet (GbE)	4 onboard GbE ports (10/100/1,000) (2 are shared as 10-GbE ports via optional XAUI cards in PCIe/XAUI combination slots)
USB	4 USB 2.0 ports (2 front/2 rear)
Power supplies	4 x 650 W DC or AC (N+2 redundancy)
RAS components	Hot-swappable and redundant PSUs and disk drives, Integrated Lights Out Manager (accessible via network or serial management ports)
Telco features	Dry Contact Alarm (DCA) indicators, Watchdog Timer (WDT), NEBS Level 3 certified
Operating system (OS)	Oracle Solaris 10 (preinstalled) or Ubuntu Linux

Figure 1 shows the Sun Netra T5440 server enclosure.

Figure 1. The Sun Netra T5440 Server addresses the needs of today's telecommunications applications.

Typical Applications

The Sun Netra T5440 server delivers scalability, energy efficiency, and reliability for a variety of demanding applications:

- Media gateway controllers
- Operations and maintenance systems for telecommunications networks
- Signaling gateways
- Intelligent networks

- Multimedia messaging services (MMS)/short messaging services (SMS) and unified messaging
- Defense, military, and intelligence applications, including shipboard command and control, mobile weapons control, and remote intelligence access servers
- Embedded original equipment manufacturer (OEM) applications, such as industrial process control, semiconductor test equipment, and network imaging systems
- Application servers
- Web servers
- Content caching, network proxy servers
- Home/visitor location registries (HLR/VLR)
- Base station controllers (BSCs)
- Content distribution networks
- DNS services
- Firewalls for virtual private network/IP security (VPN/IPSEC)
- IP traffic management systems
- Security systems
- Streaming media systems

Innovative System Design

Beyond the capabilities of individual systems, Oracle understands that data centers have unique and pressing requirements that system designers must take into account. Density, performance, and scalability are all essential design considerations, but systems must also be serviceable and operate within modern data center strategies for power, cooling, and serviceability. Sun Netra T5440 servers share an innovative design philosophy that extends across the Sun Netra and rackmount server platforms. Design principles behind this philosophy include

- **Optimal compute density.** Sun Netra and volume rackmount servers provide leading density in terms of CPU cores, memory, storage, and I/O. This focus on density often allows customers to replace competitive servers with Oracle's Sun servers, saving space and energy.
- **Leading storage capacity.** Sun Netra servers provide leading density and flexible RAID options. Smaller disk drives and innovations in structure, airway, and carrier design allow more disk capacity in smaller spaces while enhancing airflow.
- **Common, shared management.** Sun Netra T5440 servers are designed for ease of management and serviceability, sharing service processors used in other Oracle volume server

platforms. Systems and components are designed for easy identification, and hot-swappable components facilitate easy online replacement.

- **Continued investment protection.** Oracle strives to protect customer investments in applications and training. Even with breakthrough technology such as multisocketed and multithreaded processors, applications can run on new SPARC processor–based systems and Oracle Solaris without requiring modification.
- **Common chassis design.** Shared chassis designs leverage key innovations across multiple architectures, enabling the use of common components and subassemblies and simplifying administrative complexity.

Chassis Design Innovations

Sun Netra T5440 servers share many common chassis design characteristics with other Sun Netra and Oracle volume server platforms. The result is a consistent look and feel across product lines and simplified administration and serviceability (thanks to consistent placement and shared components).

Enhanced System and Component Serviceability

Finding and identifying servers and components in a modern data center can be a challenge. This is not the case with Sun Netra T5440 servers, which are optimized for lights-out data center configurations with easy-to-identify servers and modules. Color-coded operator panels provide easy-to-understand diagnostics, and systems are designed for deployment in hot-isle/cold-isle multiracked deployments, with both front and rear diagnostic LEDs to pinpoint faulty components. Consistent connector layouts for power, networking, and management make moving between Oracle systems straightforward. Hot-pluggable components are tool-less and easily available for serviceability. (For instance, a hinged front bezel panel provides easy access to disk drives so that drives can be inserted or replaced without exposure to sensitive components.)

Robust Chassis, Component, and Subassembly Design

Many of Oracle's volume servers share chassis designs that are carefully engineered to provide reliability and cool operation. Even features such as the honeycomb-shaped chassis ventilation holes help to provide the best compromise for strength, maximum airflow, and maximum electronic attenuation. Despite extreme computational, I/O, and storage density, Sun servers from Oracle are able to maintain adequate cooling using conventional technologies. Efficient modular fan assemblies keep the chassis within an effective operating temperature range. One fan assembly resides directly behind the disk drives in Sun Netra T5440 systems, allowing airflow to be directed effectively both above and below the disk drives. Next-generation drive carriers enhance chassis ventilation, enabling greater storage density while increasing system airflow. In addition, system fans force air from the front of the chassis to the back, directing airflow over the motherboard, CPUs, and fully buffered dual inline memory modules (FB-DIMMs) to cool

system components efficiently. Fan modules are isolated from the chassis to prevent them from transferring rotational vibration to other components.

Minimized DC-to-DC power conversions also contribute to overall system efficiency. By providing 12-V power to the motherboard, power conversion stages are eliminated. This approach generates less heat and introduces further system efficiencies.

Minimized Cabling for Maximized Airflow

To minimize cabling, enhance airflow, and increase reliability, the system uses a variety of smaller boards and riser cards:

- Power distribution boards (PDBs) distribute system power from the power supplies to the motherboard and the disk backplane.
- Connector boards eliminate the need for many discrete cables, providing direct interconnects to distribute control and data signals to the disk backplane, the fan boards, and the PDB.
- An optional memory mezzanine tray provides 16 memory slots in addition to 16 memory slots on the motherboard.
- A disk backplane mounts to the disk cage, delivering data through one or more four-channel discrete miniSAS cables from the motherboard. The Sun Netra T5440 server supports a 12-disk backplane. Also provided via the disk backplane are two USB connections to the front of the system.
- PCI risers and mezzanine boards enable robust PCI expansion capabilities. PCI mezzanine boards contain components to fan out PCIe and PCI-X buses and enable slot expansion. XAUI risers support combination PCIe/XAUI slots that access on-chip 10-GbE interfaces on the onboard Sun Neptune networking chip. Each PCIe/XAUI combination slot accepts either an optical/copper XAUI card or an industry-standard low-profile PCIe card with up to an x8 form-factor edge connector.
- A telecommunications alarm card allows Dry Contact Alarm (DCA) conditions to be easily monitored and reported via front-panel indicators.

For more details on the chassis, system features, and internal components, see the "Sun Netra T5440 Server Architecture" section.

Leading Reliability, Availability, and Serviceability (RAS)

Numerous architectural features—including redundant and hot-swappable components—enable Sun Netra T5440 servers to deliver the outstanding levels of availability and reliability that contribute to continuous system operation. Redundant hot-swappable DC or AC power supplies and hot-pluggable disk drives enhance system uptime. Highly reliable parts and a relatively low total component count minimize the opportunity for system errors. Dual PCIe root complexes

and the ability to configure multiple processors add to resiliency. In addition, a built-in service processor with ILOM capabilities enables remote monitoring and system management, providing the kind of reliability and serviceability needed by telecommunications companies or military organizations that operate in severe environments.

The following design elements are key to improving service continuity:

- Reduced parts count
- Processor thread and core off-lining
- Built-in RAID capabilities
- Redundant and hot-swappable components
- Parity protection and error correction capabilities
- System monitoring
- ILOM service processor
- Superior energy efficiency
- Robust built-in virtualization technology
- Comprehensive fault management

Space, Watts, and Performance: The SWaP Metric

Sun Netra T5440 servers deliver leading performance across a range of multithreaded workloads and benchmarks. However, with the pressure to conserve energy and real estate and reduce associated costs, it is not enough to measure performance in isolation. Delivering the required level of throughput in a fixed space and power envelope is critical. Traditional system-to-system benchmarks provide a valuable means of comparing the performance of one system to another, but they are limited when it comes to understanding the power and density attributes of the systems being compared. For this reason, the space, watts, and performance (SWaP) metric was designed. It provides a simple and transparent measure of overall server efficiency, SWaP is calculated using the following formula:

SWaP = Performance / (Space * Power Consumption) where

- Performance is measured by industry-standard benchmarks
- Space refers to the height of the server in rack units (U)
- Power is measured by watts used by the system, taken during actual benchmark runs or from vendors' site planning guides

The UltraSPARC T2 Plus Processor and the Evolution of Throughput Computing

Traditional high-frequency, single-threaded processors are often limited in their ability to scale application performance. In many cases, ramping processor clock rates yields only small performance gains—at the same time consuming more power and generating large heat loads. Higher power and cooling requirements usually translate into higher energy costs, which take an increasingly large bite out of today's IT operating budgets.

For this reason, Oracle's processors and systems provide higher throughput (orders of magnitude higher than for previous systems), but reduce the need for power and cooling. This focus on throughput computing represents a composite strategy targeted at improving the performance of key workloads:

- CMT processor designs (such as the UltraSPARC T2 Plus processor) provide massive thread-level parallelism (TLP) and increase application throughput while at the same time maintaining attractive power and cooling profiles.
- The innovative and secure Oracle Solaris 10 effectively delivers the resources of CMT processors to applications, facilitating fine-grained system virtualization and high utilization while maintaining binary compatibility for Oracle Solaris applications running on SPARC processor–based platforms.
- Oracle's compilers, development tools, APIs, and system engineering efforts leverage CMT processor capabilities, resulting in significant application performance gains.

This section describes the limitations of traditional processor designs and demonstrates how Oracle's CMT emphasis is helping to increase real-world application performance while optimizing energy and space efficiency.

Diminishing Returns of Traditional Processor Designs

An often-quoted tenet of Moore's Law states that the number of transistors that fit in a square inch of integrated circuitry will double approximately every two years. For more than three decades, the pace of Moore's Law has held, driving processor performance to new heights. Processor manufacturers have long exploited gains in chip real estate to build increasingly complex processors, with instruction-level parallelism (ILP) as a goal. Today, these traditional processors employ very high frequencies along with a variety of sophisticated tactics to accelerate a single instruction pipeline, including

- Large caches
- Superscalar designs
- Out-of-order execution

- Very high clock rates
- Deep pipelines
- Speculative prefetches

Although these techniques have produced faster processors with impressive-sounding multigigahertz frequencies, they have largely resulted in complex, hot, and power-hungry processors that don't fit well within the constraints of today's data centers. In fact, many workloads are simply unable to take advantage of the hard-won ILP provided by these processors. Applications with high shared memory and data requirements are typically more focused on processing a large number of simultaneous threads (or TLP) than on running a single thread as quickly as possible, as in ILP-centric workloads.

Making matters worse, the majority of ILP in existing applications has already been extracted, meaning that further gains are likely to be small. With higher clock speeds, each successive processor generation has seemingly demanded more power than the last, and microprocessor frequency scaling has leveled off in the range of 2 to 4 GHz as a result. Deploying pipelined superscalar processors requires more power, which limits the approach because of a fundamental need for processor cooling.

Chip Multiprocessing (CMP) with Multicore Processors

To address this limitation, some microprocessor engineers have used the transistor budget provided by Moore's Law to group two or even four conventional processor cores on a single physical die, creating multicore processors (or CMPs). The individual processor cores introduced by many CMP designs provide no greater performance than previous single-processor chips; in fact, they've been observed to run single-threaded applications more slowly than single-core processor versions. However, aggregate chip performance increases because multiple programs (or multiple threads) can be accommodated in parallel (as in TLP).

Unfortunately, most currently available (or soon-to-be-available) CMPs simply replicate cores from existing single-threaded processor designs. This approach typically yields only slight improvements in aggregate performance because it ignores key performance issues such as memory speed and hardware thread context switching. As a result, these designs provide some additional throughput and scalability, but they can still consume considerable power and generate significant heat.

Chip Multithreading (CMT)

Engineers were early to recognize the disparity between processor speeds and memory access rates. Although processor speeds continue to double every two years, memory speeds have typically doubled only every six years. As a result, memory latency now mostly dominates in

determining application performance, erasing impressive gains in clock rates. This growing disconnect is the result of memory suppliers focusing on density and cost rather than speed.

Unfortunately, this relative gap between processor and memory speeds can leave ultrafast processors idle as much as 85 percent of the time, waiting for memory accesses to return required data. Ironically, as traditional processor execution pipelines get faster and more complex, the impact of memory latency grows, with fast, expensive processors spending more cycles waiting and doing nothing. Worse still, idle processors continue to draw power and generate heat. Thus, it's easy to see that frequency (that is, gigahertz) is truly a misleading indicator of real performance.

First introduced with the UltraSPARC T1 processor, CMT takes advantage of CMP evolution but adds the ability to scale with threads rather than frequency. Unlike traditional single-threaded processors and even most current multicore (CMP) processors, hardware multithreaded processor cores allow rapid switching between active threads while other threads stall for memory. Figure 2 illustrates the difference between CMP, fine-grained hardware multithreading (FG-MT), and CMT. The key to the CMT approach is that each core is designed to switch between multiple threads on each clock cycle. As a result, the processor's execution pipeline remains active doing useful work, even as memory operations for stalled threads continue in parallel.

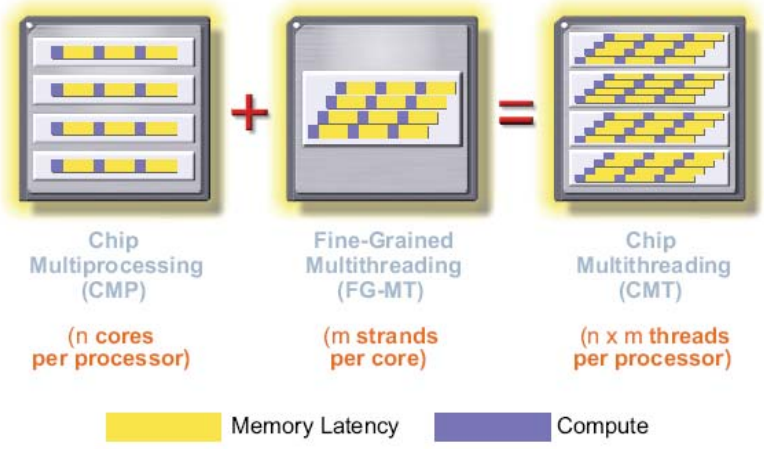

Figure 2. Chip multithreading combines chip multiprocessing and fine-grained hardware multithreading.

CMT provides value by increasing the execution pipeline's ability to do useful work on any given clock cycle. Because a number of execution threads now share resources, pipeline utilization is greatly enhanced. In addition, the negative effects of memory latency are effectively masked because the processor and memory subsystems remain active in parallel to the execution pipeline. Because individual processor cores implement much simpler pipelines (emphasizing TLP over ILP), they are also substantially cooler and require significantly less electrical energy to operate. This innovative approach results in CoolThreads technology—multiple physical instruction

execution pipelines (one for each core), with multiple active thread contexts available per core. UltraSPARC T2 and UltraSPARC T2 Plus processors feature two execution pipelines per core to boost scalability.

UltraSPARC Processors with CoolThreads Technology

The first processor to implement CoolThreads technology was the UltraSPARC T1 processor, shown on the right in Figure 3.

Figure 3. The UltraSPARC T2 Plus, UltraSPARC T2, and UltraSPARC T1 processors all take advantage of CoolThreads technology.

The UltraSPARC T2 processor (see Figure 3, center)—the second-generation processor to use CoolThreads technology)—was the industry's first system-on-a-chip (SoC), integrating key system functions directly on the processor. In addition, the UltraSPARC T2 processor supplied more cores and threads than any general-purpose system processor available at the time. The UltraSPARC T2 Plus processor (see Figure 3, left) adds cache coherency capabilities to the SoC design, enabling multisocketed system implementations.

TABLE 2. ULTRASPARC T1, ULTRASPARC T2, AND ULTRASPARC T2 PLUS PROCESSOR FEATURES

FEATURE	ULTRASPARC T1 PROCESSOR	ULTRASPARC T2 PROCESSOR	ULTRASPARC T2 PLUS PROCESSOR
Year introduced	2005	2007	2008
Cores per processor	Up to 8	Up to 8	Up to 8
Threads per core	4	8	8

Threads per processor	32	64	64
Sockets supported	1	1	2 or 4[1]
Hypervisor	Yes	Yes	Yes
Chip memory characteristics	4 memory controllers, 4 DIMMs per controller	4 memory controllers, 16 DIMMs per controller	2 memory controllers, 16 or 32 DIMMs per controller
Caches	16-KB instruction cache, 8-KB data cache, 3-MB L2 cache (4 banks, 12-way associative)	16-KB instruction cache, 8-KB data cache, 4-MB L2 cache (8 banks, 16-way associative)	16-KB instruction cache, 8-KB data cache, 4-MB L2 cache (8 banks, 16-way associative)
Technology	90-nm technology	65-nm technology	65-nm technology
Floating-point unit (FPU)	1 FPU per chip	1 FPU per chip, 8 FPUs per core	1 FPU per chip, 8 FPUs per core
Integer resources	Single execution unit	2 integer execution units per core	2 integer execution units per core
Cryptography	Accelerated modular arithmetic operations (RSA), support for 2 popular ciphers	Stream-processing unit per core, support for the 10 most popular ciphers	Stream-processing unit per core, support for the 10 most popular ciphers
Additional on-chip resources	None	Dual 10-GbE interfaces, integrated PCIe controller	Integrated PCIe controller, cache coherency logic and links (4.8 Gb/sec)

System-on-a-Chip (SoC) Design

By integrating computing, security, and I/O on a single chip, the UltraSPARC T2 and UltraSPARC T2 Plus processors minimize the need for additional, expensive custom hardware and software. Binary compatible with earlier UltraSPARC processors, these later processor generations provide the agility needed to rapidly scale new network services with maximum efficiency and predictability.

Unlike complex single-threaded processors, CMT processors use the available transistor budget to implement multiple threaded processor cores on a single chip die. The UltraSPARC T2 and

[1]The Sun Netra T5440 server is a two-socket implementation.

UltraSPARC T2 Plus processors provide up to eight cores per processor, with each core supporting up to eight threads via two independent pipelines. This design effectively doubles UltraSPARC T1 processor throughput without raising the clock frequency. In addition, the UltraSPARC T2 and UltraSPARC T2 Plus processors use the increased transistor budget (afforded by 65-nm silicon technology) to implement a threaded SoC design. A single UltraSPARC T2 or UltraSPARC T2 Plus processor die hosts

- 8 threads per core for a maximum of 64 threads per processor
- On-chip Level 1 and Level 2 caches
- On-chip memory controllers
- Per-core pipelined floating-point units (FPUs)
- Per-core cryptographic acceleration
- Two on-chip 10-GbE interfaces (UltraSPARC T2 processor only)
- On-chip PCI-Express (PCIe) interface
- On-chip cache coherency logic and links (UltraSPARC T2 Plus processor only)

The CMT design recognizes that memory latency is truly the bottleneck to improving performance, allowing each UltraSPARC T2 and UltraSPARC T2 Plus core to switch between up to eight threads. In addition, each core provides two integer execution units (EXUs), so that a single core is capable of executing two threads at a time. Figure 4 provides a simplified high-level illustration of the thread model supported by the eight-core UltraSPARC T2 and UltraSPARC T2 Plus processors.

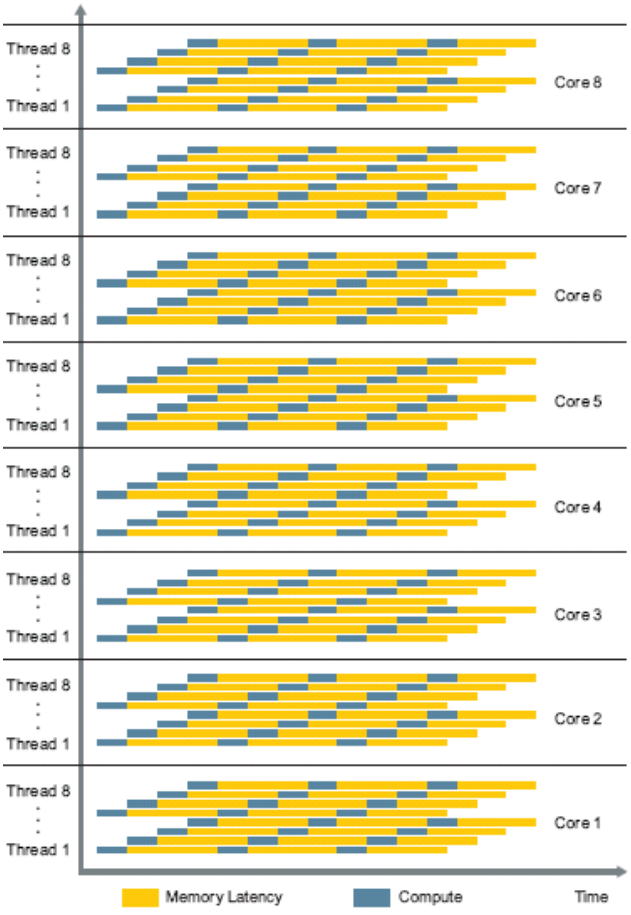

Figure 4. A single eight-core UltraSPARC T2 or UltraSPARC T2 Plus processor can run up to 64 threads, with up to 2 threads running in each core simultaneously.

UltraSPARC T2 Plus Processor Architecture

Figure 5 provides a block-level diagram of the UltraSPARC T2 Plus processor. As in the UltraSPARC T2 design, the eight cores on the UltraSPARC T2 Plus processor are interconnected with a full on-chip nonblocking 8 x 9 crossbar switch. The crossbar connects each core to the eight banks of L2 cache and to the system interface unit for I/O. The crossbar provides 300 Gb/sec of bandwidth and supports 8-byte writes from a core to a bank and 16-byte reads from a bank to a core. The system interface unit connects networking and I/O directly to memory through the individual cache banks. Using FB-DIMM memory supports dedicated northbound and southbound lanes to and from the caches to accelerate performance and reduce latency. This approach provides higher bandwidth than with DDR2 memory, with up to 42.4 GB/sec of read bandwidth and 21 GB/sec of write bandwidth.

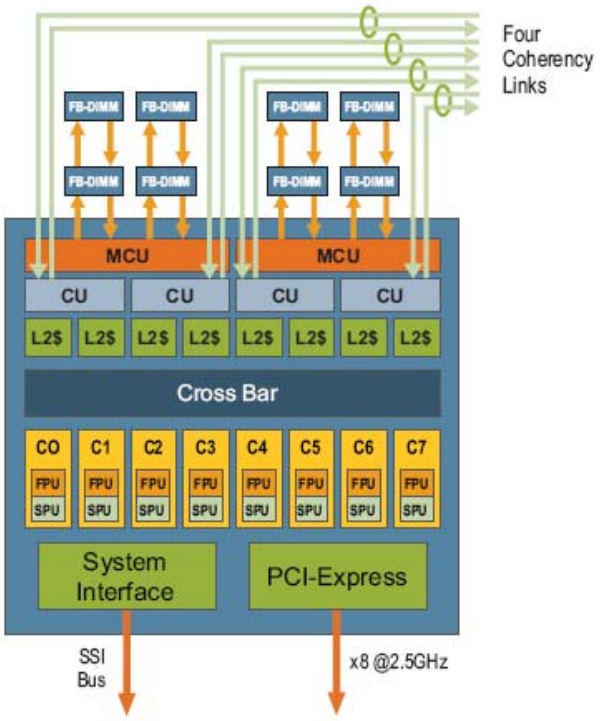

Figure 5. The UltraSPARC T2 Plus processor uses a design similar to the UltraSPARC T2 processor but adds four coherency links to interconnect up to four processors.

Each core provides its own fully pipelined floating-point/graphics unit, as well as a stream-processing unit (SPU). The FPUs greatly enhance floating-point performance over that of the UltraSPARC T1, enhancing the performance of multimedia processing including image encoder/decoder (codec) operations. The SPUs provide wire-speed cryptographic acceleration and support more than 10 popular ciphers, including DES, 3DES, AES, RC4, SHA-1, SHA-256, MD5, RSA to 2048 key, ECC, and CRC32. (In contrast, the UltraSPARC T1 processor supports only 2 ciphers, DES and 3DES.) Embedding hardware cryptographic acceleration for many ciphers in the UltraSPARC T2 and UltraSPARC T2 Plus processors allows end-to-end encryption with no penalty in either performance or cost.

The UltraSPARC T2 Plus architecture is similar in design to the earlier UltraSPARC T2 processor, but it adds support for multisocketed implementations. To do this, it omits dual on-chip 10-GbE interfaces from the UltraSPARC T2 processor design and uses the on-chip real estate to provide four coherency units (CUs). It also replaces two memory channels with four coherency channels (or coherency links)—one provided by each CU. These links run a cache coherency (snoopy) protocol over an FB-DIMM–like physical interface to provide up to 4.8 GT/sec, providing 204 Gb/sec in each direction. (Note that the memory link speed of the UltraSPARC T2 Plus processor is 4.8 GT/sec, compared with 4.0 GT/sec for the earlier-generation UltraSPARC T2 processor.)

Because of the additional coherency capabilities, the UltraSPARC T2 Plus processor can support either two- or four-socket implementations. Figure 6 shows a typical two-socket implementation (like that of the Sun Netra T5440 server). Dual-socket UltraSPARC T2 Plus implementations interconnect the processor's four coherency links, without the need for any additional circuitry.

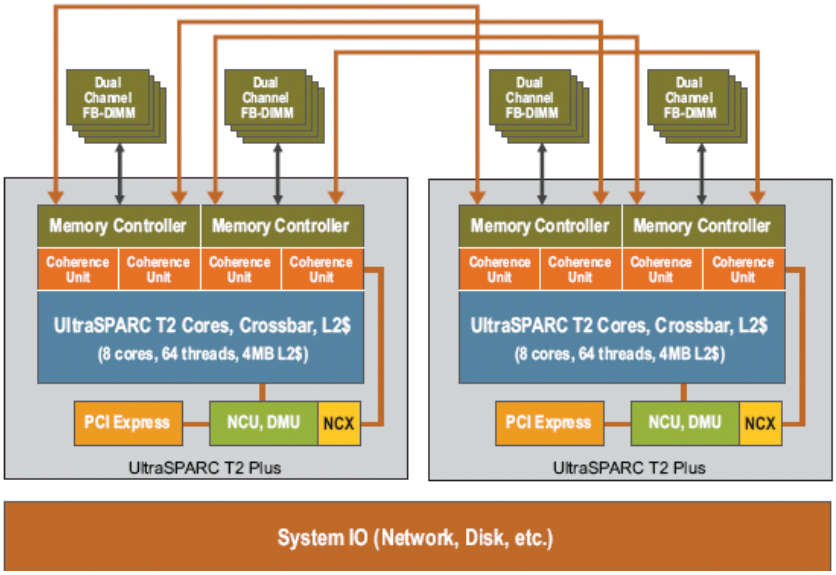

Figure 6. The Sun Netra T5440 server features a dual-socket UltraSPARC T2 Plus configuration like this one.

UltraSPARC T2 Plus Core Architecture and Pipelines

Figure 7 provides a block-level diagram representing a single core on the UltraSPARC T2 Plus processor.

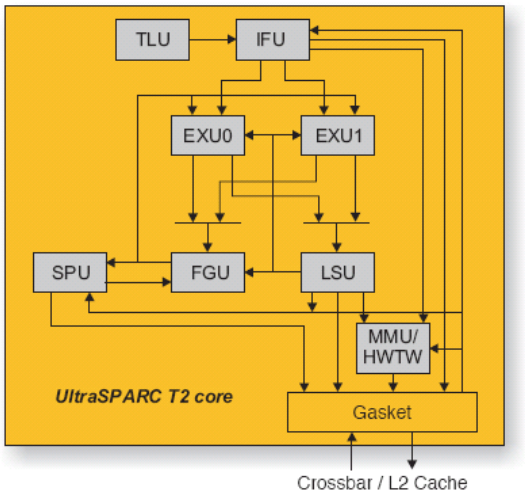

Figure 7. This diagram represents a single core on the UltraSPARC T2 Plus processor.

Components implemented in each core include

- **Trap logic unit (TLU).** The TLU updates the machine state and handles exceptions and interrupts.
- **Instruction fetch unit (IFU).** The IFU includes the 16-KB instruction cache (32-byte lines, eight-way set associative) and a 64-entry fully associative instruction translation lookup buffer (ITLB).
- **Integer execution units (EXUs).** Dual integer EXUs are provided per core, with four threads sharing each unit. Eight register windows are provided per thread, with 160 integer register file (IRF) entries per thread.
- **Floating-point/graphics unit.** A floating-point/graphics unit is provided within each core and is shared by all eight threads assigned to the core. Thirty-two floating-point register file entries are provided per thread.
- **Stream-processing unit (SPU).** Each core contains an SPU that provides cryptographic co-processing.
- **Memory management unit (MMU).** The MMU provides a hardware table walk (HWTW) and supports 8-KB, 64-KB, 4-MB, and 256-MB pages.

An 8-stage integer pipeline and a 12-stage floating-point pipeline are provided by each UltraSPARC processor core (Figure 8). The "pick" pipeline stage chooses two threads (out of the eight possible per each core) to execute in each cycle.

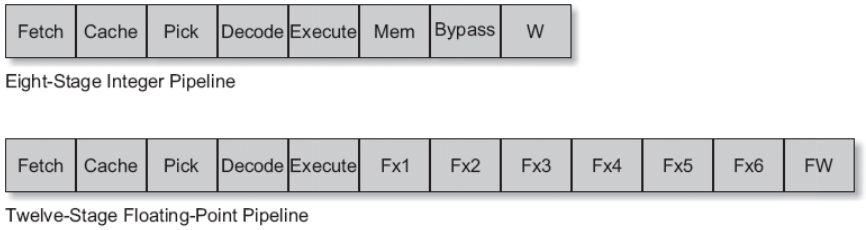

Figure 8. Each UltraSPARC T2 processor core provides integer and floating-point pipelines.

To illustrate how the dual pipelines function, Figure 9 depicts the integer pipeline with the load-store unit (LSU). All eight threads within the core share the instruction stack. A least-recently-fetched algorithm is used to select the next thread to fetch. Each thread is written into a thread-specific instruction buffer (IB), and each of the eight threads is statically assigned to one of two thread groups within the core.

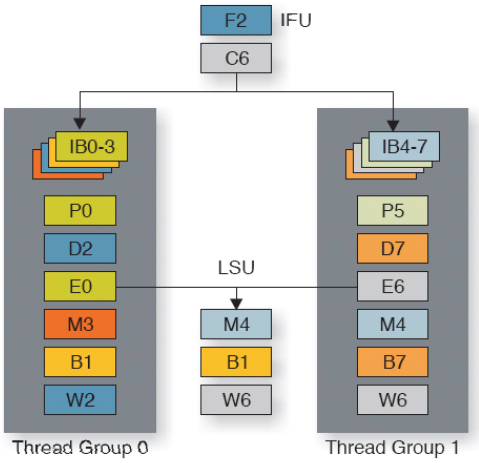

Figure 9. Threads are interleaved between pipeline stages with very few restrictions.

The pick stage chooses one thread each cycle within each thread group. Picking within each thread group is independent of the others, and a least-recently-picked algorithm is used to select the next thread to execute. The decode state resolves resource conflicts that are not handled during the pick stage. As shown in the illustration, threads are interleaved between pipeline stages with very few restrictions. Any thread can be at the fetch or cache stage before being split into either of the two thread groups. LSUs and FPUs are shared between all eight threads. Only one thread from either thread group can be scheduled on such a shared unit.

Integral PCI-e Support

The UltraSPARC T2 Plus processor provides an on-chip PCIe controller that operates at 4 GB/sec per lane bidirectionally through a point-to-point dual-simplex chip interconnect. An integral IOMMU supports I/O virtualization and process device isolation by using the PCIe BDF number. The total I/O bandwidth is 3 to 4 GB/sec, with maximum payload sizes of 128 to 512 bytes. An x8 serializer/deserializer interface is provided for use with off-chip PCIe switches.

Power Management

The UltraSPARC T2 processor was the first processor to incorporate unique power management features at both the core and memory levels. Now the UltraSPARC T2 Plus processor takes advantage of these features, including reduced instruction rates, parking of idle threads and cores, and the ability to turn off clocks in cores and memory to reduce power consumption. Substantial innovation is present, such as

- Limiting speculation such as conditional branches not taken
- Extensive clock gating in the data path, control blocks, and arrays
- Power throttling that allows extra stall cycles to be injected into the decode stage

Sun Netra T5440 Servers

Carrier-grade Sun Netra T5440 servers are designed to take advantage of the powerful resources and features of the UltraSPARC T2 Plus processor, including eight cores, 64 threads, large on-chip memory capacities, cryptographic acceleration, integrated on-chip I/O technology, and power management. In a compact, energy-efficient 4U package, Sun Netra T5440 servers provide robust, scalable performance within significant power, cooling, and space constraints—at the same time enabling remarkable reliability, density, and expandability.

Sun Netra T5440 Server Architecture

The Sun Netra T5440 server is designed to deliver breakthrough performance while maximizing reliability, minimizing power consumption, and optimizing expandability. This section details physical and architectural aspects of the system.

System-Level Architecture

Figure 10 shows the Sun Netra T5440 system block diagram. Key features include:

- Compact, rack-optimized 4U chassis design
- Dual sockets for UltraSPARC T2 Plus processors, connected by four coherency links, which deliver up to 128 simultaneous execution threads using CoolThreads technology. (System configurations include one or two processors with four, six, or eight processor cores.)
- Exceptional expandability and I/O performance, including two PCI-X slots, two combination PCIe/XAUI slots for 10-GbE networking, six other PCIe slots, four onboard 10/100/1,000-Mb/sec Ethernet ports, and an optional memory mezzanine tray that supplies an additional 16 slots of memory expansion (for a total of 32 system memory slots)
- Motherboard integration of Oracle's Neptune Ethernet chip, which can provide 10-GbE functionality in conjunction with PCIe/XAUI cards
- Onboard FPGA that operates as the ILOM service processor, providing simplified state-of-the-art remote maintenance
- Onboard serial-attached SCSI (SAS) disk controller and large internal storage capacity, with a maximum of 12 internal, 2.5-inch, 146-GB, hot-pluggable SAS disk drives
- I/O subsystems that support USB, DVD-RW, quad GbE, and PCI expansion
- Preloaded Oracle Solaris to shorten time to deployment
- Enhanced system uptime with DCA indicators, on-chip RAS features, redundant DC or AC power supplies, hot-pluggable disk drives, and support for hardware RAID (0 + 1)

- Telecordia NEBS Level 3 certification for extremely reliable operation in harsh environmental conditions

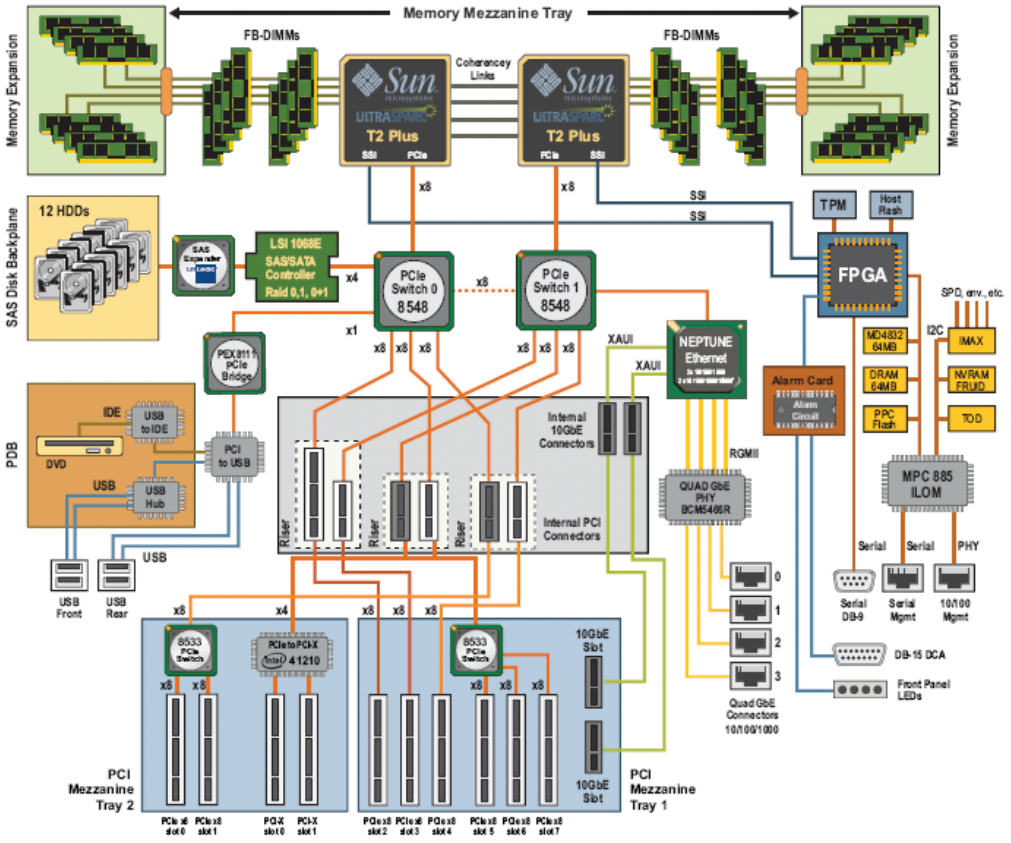

Figure 10. This block-level diagram provides an overview of the Sun Netra T5440 server motherboard.

The motherboard interconnect is greatly simplified compared to prior-generation servers. Power is distributed to the motherboard through a pair of metal bus bars connected to a PDB. A single flex-circuit connector routes all critical power control and DVD drive signaling to the PDB. One or two miniSAS cables (depending on the number of drives configured) connect the motherboard to the disk drive backplane, enabling access to internal SAS hard drives.

Sun Netra T5440 Server Overview and Subsystems

The Sun Netra T5440 server is engineered to deliver high throughput and breakthrough CPU, memory, storage, and I/O densities. With considerable expansion capabilities and many high-reliability features, this server provides an ideal platform for consolidating throughput-intensive, business-critical workloads.

Enclosure

The Sun Netra T5440 server features a compact yet expandable 4U rackmount chassis (see Table 3) that provides the flexibility to scale processing and I/O needs without wasting valuable data center space.

TABLE 3. DIMENSIONS AND WEIGHT

SERVER/ DIMENSION	UNITED STATES	INTERNATIONAL
Height	6.86 inches (4U)	174.2 millimeters
Width	17.5 inches (including bezel)	445 millimeters (including bezel)
Depth	• 20.87 inches (maximum to PSU Depth handles) • 19.88 inches (to rear I/O)	• 530 millimeters (maximum to PSU handles) • 505 millimeters (to rear I/O)
Weight	72.6 pounds (fully configured without PCI cards)	33 kilograms (fully configured without PCI cards)

Major Components

The Sun Netra T5440 server includes the following major components:

- One or two UltraSPARC T2 Plus processors running at 1.2 GHz and configured with four, six, or eight cores
- Up to 128 GB of memory in 32 FB-DIMM slots (populated in pairs of 1-GB, 2-GB, or 4-GB FB-DIMMs)
- Two PCI-X slots, two combination PCIe/XAUI slots, and six additional PCIe slots
- Four onboard 10/100/1,000-Mb/sec Ethernet ports, two of which can be configured to support 10 GbE with optional PCIe/XAUI cards
- Four USB 2.0 ports
- Up to 12 internal SAS drives
- An optional internal, slim-line DVD-RW device
- An onboard service processor with ILOM
- Four hot-pluggable/hot-swappable, high-efficiency DC or AC power supplies (N+2 redundancy)
- Three separate fan assemblies that operate under environmental monitoring and control

Front and Rear Perspectives

Figure 11 illustrates the Sun Netra T5440 server's front and back panels.

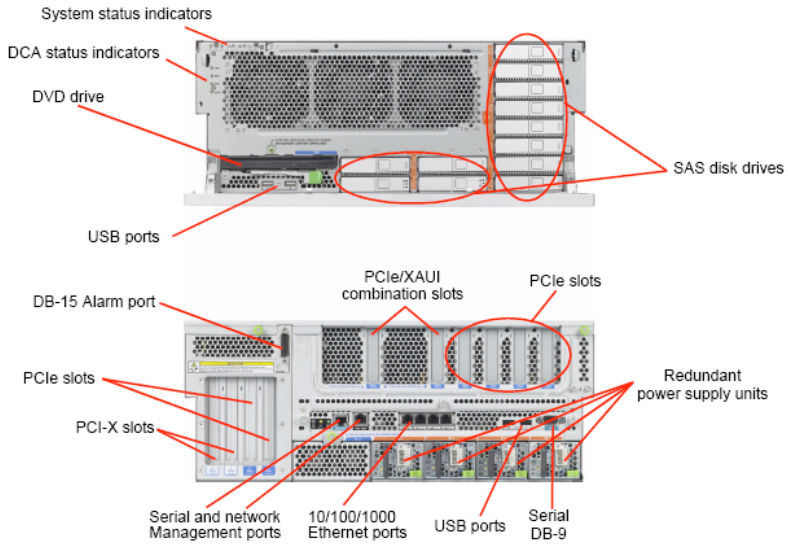

Figure 11. These front- and rear-panel views of the Sun Netra T5440 server reveal its external features.

External features include the following:

- Front and rear system status indicator lights provide "Locator" (white), "Service Required" (amber), and "Activity Status" (green) for the system.
- Four additional front-panel indicators ("Critical," "Major," "Minor," and "User") signify application-defined DCA conditions.
- A hinged access door opens to provide access to the front panel, where hot-swappable SAS disk drives and the optional DVD-RW device are accessible.
- Two USB ports are provided on the rear panel along with two on the front panel.
- Four hot-pluggable/hot-swappable DC or AC power supply units (PSUs) insert from the rear (rear indicator lights convey the status of each PSU).
- Four autosensing 10/100/1,000 Base-T Ethernet ports are available on the rear panel. If optional PCIe/XAUI networking cards are installed, two of the four ports connect to the Neptune networking chip to provide 10-GbE functionality.
- PCI expansion slots support up to 10 PCI cards (two PCI-X slots, two combination PCIe/XAUI slots for either optional XAUI networking cards or PCIe cards, and six other PCIe card slots).

- Two RJ-45 management ports (serial and network) on the rear panel are for use with Oracle Integrated Lights Out Manager. The serial port is the default connection to the controller. (Optionally, the network port supports a 10/100 Base-T connection to the controller.)
- One ttya serial port (DB-9) is provided on the rear panel for serial device connection.
- One DCA output port (DB-15) is provided on the rear panel, which connects to the service processor for reporting application-specific conditions.

Memory Subsystem

Each UltraSPARC T2 Plus processor provides two memory controller units (MCUs). The MCUs communicate directly to FB-DIMM memory through high-speed serial links, transferring data at an aggregate rate of 4.8 GT/sec.

Sixteen motherboard memory socket locations house two rows of 667-MHz FB-DIMMs per channel. Sun Netra T5440 servers support an optional memory mezzanine tray (see Figure 12) that can be added to support an additional 16 FB-DIMM slots, for a total of 32 slots per system. On both the motherboard and the memory mezzanine tray, FB-DIMM modules must be populated in like pairs—with either 1-GB, 2-GB, or 4-GB modules—enabling a maximum system memory capacity of 128 GB using 4-GB FB-DIMM modules.

Figure 12. An optional memory mezzanine tray doubles memory capacity of the Sun Netra T5440 server.

I/O Subsystem

Each UltraSPARC T2 Plus processor incorporates a single, eight-lane (x8) PCIe port capable of operating at 4 GB/sec bidirectionally. This port natively interfaces to I/O devices through a series of PLX Technology PCIe expander chips, connecting either to PCIe card slots or to bridge devices that interface with PCIe, including

- **Disk controller.** Disk control is managed by a single LSI Logic SAS1068E SAS controller chip on the motherboard that interfaces to a four-lane (x4) PCIe port.
- **Modular disk backplanes.** A 12-disk backplane is attached to the LSI disk controller by one or more x4 SAS links, depending on the drive configuration.

- **Quad GbE.** Oracle's Neptune Ethernet chip provides two 10/100/1,000 Base-T ports and two 10/100/1,000/10,000 Base-T interfaces, exposed as four RJ-45 connectors on the rear panel.
- **Dual 10 GbE.** With the addition of optional PCIe/XAUI cards, Oracle's Neptune Ethernet chip can support two 10-GbE XAUI connections. When XAUI cards are used and 10-GbE ports are connected, two of the four 10/100/1,000 Base-T ports become unavailable.
- **USB and DVD.** A single-lane PCIe port connects to a PLX PEX8111 PCI bridge device. A second bridge chip converts the 32-bit 33-MHz PCI bus into multiple USB 2.0 ports. These ports drive the rear-panel USB interconnects, the front-panel USB connections (via a hub), and a USB-to-IDE bridge that supports the DVD device.

PCI Expansion

Each UltraSPARC T2 Plus incorporates a single eight-lane PCIe port capable of operating at 2 GB/sec in either direction. This port natively interfaces to the system I/O devices through two PLX Technology 8548 PCIe expander chips, which branch out to support system PCI expansion. Sun Netra T5440 servers include support for legacy PCI-X telecommunications cards and PCIe/XAUI cards, as well as PCIe cards.

The Sun Netra T5440 server features expansion for two PCI-X slots and a maximum of eight PCIe slots. Operating at 2.5 GHz, PCIe is a next-generation system bus interconnect that replaces the original PCI bus design. It is a high-speed, point-to-point dual simplex chip interconnect. Two of the eight PCIe slots are combination PCIe/XAUI slots and can be populated with XAUI cards for 10-GbE networking connections. Otherwise, the combination slots can be used for PCIe card expansion.

To facilitate PCI expansion, the system chassis houses dual PCI mezzanine boards and three PCI risers that connect to the system motherboard. The first PCI mezzanine board uses a PLX Technology PEX8533 PCIe switch to expand PCIe buses and provide support for six PCIe slots that accept low-profile cards. (Two of these six slots are the combination slots for PCIe/XAUI cards. Two PCIe/XAUI risers enable XAUI card support.)

The second PCI mezzanine board incorporates a PLX Technology PEX8525 PCIe switch (to fan out PCIe buses) and an Intel 41210 bridge chip, which converts four-lane PCIe buses to PCI-X buses to support two PCI-X slots. The second mezzanine board also supports two full-length, full-height PCIe slots (for a maximum of eight PCIe slots in the system).

PCI slots are all oriented vertically in the Sun Netra T5440 server. From the rear, slots are numbered starting at the bottom left and continuing to the right. The two leftmost slots (Slots 0 and 1) are PCI-X slots for full-length, full-height PCI-X cards. The next two full-length, full-height slots (Slots 2 and 3) are PCIe slots, x8 electrically and x16 mechanically. Moving to the upper-right bank of slots, Slots 4 and 5 are the combination PCIe/XAUI slots, which accept either an optical/copper XAUI 10-GbE networking card or an industry-standard, low-profile

PCIe card. The remaining bank of four PCIe slots (Slots 6 through 9) accepts industry-standard, low-profile cards. Table 4 summarizes numbering, type, and mechanical and electrical characteristics for all 10 PCI slots in Sun Netra T5440 servers.

TABLE 4. SUN NETRA T5440 SERVER PCI SLOT NUMBER, TYPE, AND CHARACTERISTICS

SLOT NUMBER	TYPE	DESCRIPTION
0	PCI-X	133 MHz @ 64-bit, full-length, full-height
1	PCI-X	133 MHz @ 64-bit, full-length, full-height
2	PCIe (full-length, full-height)	• Mechanical: x16 • Electrical: x8[2]
3	PCIe (full-length, full-height)	• Mechanical: x16 • Electrical: x8
4	Combination PCIe/XAUI slot (for 10-Gb networking card)	• Mechanical: x8 • Electrical: x8
5	Combination PCIe/XAUI slot (for 10-Gb networking card)	• Mechanical: x8 • Electrical: x8
6	PCIe (low-profile)	• Mechanical: x8 • Electrical: x8
7	PCIe (low-profile)	• Mechanical: x8 • Electrical: x8
8	PCIe (low-profile)	• Mechanical: x8 • Electrical: x8
9	PCIe (low-profile)	• Mechanical: x8 • Electrical: x8

[2]There is a 25-W maximum load for PCI Slots 0–3 and a 15-W maximum load for PCI Slots 4–9.

System Network Interfaces

Multiple onboard GbE ports promote flexibility as well as network configurations that support interface failover. The Sun Netra T5440 server facilitates advanced networking capabilities, with the flexibility to configure high-speed 10-GbE networking through the addition of optional XAUI networking cards.

Oracle's Neptune Ethernet controller enables on-chip support for both 1-GbE and 10-GbE networking. On the motherboard, one eight-lane PCIe port connects to the Neptune Ethernet chip, which in turn drives the four Ethernet RJ-45 interfaces located on the rear panel.

The ports are numbered in sequence from left to right. Each port autonegotiates its link connection, and LEDs above the port indicate the speed of the established link (green signifies that the link is established at 1,000 Mb/sec). All four ports support 10/100/1,000-Mb/sec; full- and half-duplex operation; IEEE 802.3ab autonegotiation for speed, duplex, and flow control; and PXE boot for network booting.

In addition to standard 1-GbE networking, the Neptune chip supports two 10-GbE connections via the two combination PCIe/XAUI slots, which are implemented using XAUI connectors on the motherboard. In the combination slots, either optical/copper XAUI cards or PCIe cards can be used. Because the Neptune controller is shared, using an XAUI card disables a corresponding 10/100/1,000-Mb/sec port. Installing an XAUI card in Slot 4, for example, disables onboard Network Port 1 (NET1). Installing an XAUI card in Slot 5 disables onboard Network Port 0 (NET0).

Internal Storage

The Sun Netra T5440 server supports up to 12 SAS disk drives. The SAS hard drives are hot-swappable, 146-GB-capacity, 10,000-RPM, 3-Gb/sec, 2.5-inch hard drives. The drives are also 100 percent-duty-cycle, small-form-factor drives that are NEBS certified.

The Sun Netra T5440 server uses an LSI SAS1068E disk controller on the motherboard. The SAS controller supports hardware RAID, providing data redundancy and increased performance without requiring additional cost. The controller supports both RAID 0 (integrated mirroring) and RAID 1 (integrated striping), as well as RAID 0+1. A 28-port LSI SAS expander (the LSISASX28) helps facilitate the server's large internal storage capacity.

An innovative new drive carrier design (which facilitates effective airflow above and below each drive) is partly responsible for the server's ample storage density. Drives insert into a modular disk tray and cable-free disk backplane that increases reliability and serviceability. The carrier includes an ejection handle that simplifies drive removal (drives are hot-swappable when disk mirroring is configured). Drive status lights indicate "Ready to remove," "Fault," and "Status."

The optional internal DVD-RW device provides the ability to read and write to removable DVD media, enabling users to store data without external storage or hard drive requirements. In

addition, the virtual storage feature of the ILOM service processor allows host access to remote CD-ROM ISO images, just as if the CD device were a locally attached USB CD-ROM device, eliminating the cost, complexity, and need for a separate optical disk drive for each individual server.

Power Supplies and Distribution

Engineered for high availability as well as low energy consumption, the Sun Netra T5440 server is equipped with four highly efficient, redundant, hot-swappable DC or AC PSUs. PSU redundancy means that continuous power is supplied if a pair of power supplies fails (two power supplies are sufficient to run a fully populated server). For maximum protection against power supply failures, Oracle recommends that four power supplies be installed at all times. In normal operation, redundant PSUs share system power demands equally.

To reduce power requirements and meet Telecordia NEBS Level 3 certification, the Sun Netra T5440 server can be configured with DC (rather than AC) power supplies. Using DC power reduces overall operating costs by lowering energy use, reducing heat, and increasing reliability.

DC or AC PSUs in both server models are rated at 660 watts. The AC input PSUs automatically adjust to operate on any input voltage between 90 VAC and 264 VAC at frequencies of 47 Hz to 63 Hz. The PSUs typically offer efficiencies in the range of 80 to 85 percent for the AC input units and 87 to 91 percent for the DC input units. Each PSU features a nonremovable internal fan that supplies independent PSU cooling. Three indicator lights display power supply status information ("AC," "Fault," and "OK").

The Sun Netra T5440 server uses a PDB to route connections between power supplies and major system components. The PDB contains a single 10-A 12-V-to-5-V DC-to-DC supply used to power the disk subsystem and the optional integrated development environment (IDE) DVD-RW device.

Cooling and Fan Assemblies

The Sun Netra T5440 server's innovative chassis design helps reduce the need for cooling. Effective front-to-back airflow helps lower component temperatures, reducing the number of fans needed to cool the system.

An onboard service processor, which monitors processor and system ambient air temperatures, controls variable-speed system fans. Based on temperature readings, the fans operate at the lowest possible speeds to provide sufficient cooling—conserving power usage, prolonging fan life, and reducing acoustical noise. A green status light on a fan module indicates proper operation, while an amber light indicates a fan fault.

The system chassis is divided into distinct airflow chambers with separate fan assemblies for cooling. All air chambers are front-to-back cooled. System fans draw air across the motherboard (including CPUs and memory) and vent it through the rear of the system. An air baffle is used to

channel airflow effectively across the motherboard and the upper PCI mezzanine tray. A separate air baffle fits over the system memory mezzanine to enhance cooling efficiency for the memory area. Another fan assembly resides behind the disk drive cage, pulling air across the drives and across the second PCI mezzanine tray. This fan assembly uses a cableless connection and plugs directly into the motherboard, enhancing airflow and increasing reliability. The Sun Netra T5440 server also uses a third fan assembly that plugs into the rear panel to help cool FB-DIMMs.

Rackmounting

The Sun Netra T5440 server ships with a hard rackmount 19-inch, four-post kit. Optional kits are available for 19-inch, two-post; 23-inch, two-post; and 600-millimeter-by-600-millimeter hard rackmount, as well as 19-inch, four-post slide rackmount. Oracle offers the following options to facilitate rackmounting:

- **Rack-Mounting Slide Rail Kit.** This is a four-point mounted slide rail kit (that is, mounting points are located at the rack front and rear).
- **Cable Management Arm (CMA).** The CMA supports and protects cables as the server slides into and out of the rack.

RAS Features

Corporate data and business information represent critical business assets. Enterprise computing technologies strive to furnish a high degree of data protection (reliability), provide virtually continuous application access (availability), and incorporate procedures and components that help resolve problems with minimal business impact (serviceability). Commonly referred to as RAS, these capabilities are a standard part of Oracle's mission-critical computing solutions.

The Sun Netra T5440 servers are engineered for hardware failure prevention, near-continuous operation, fast recovery, system security, and easy serviceability. RAS features for these systems include

- **High CPU density.** Multiple UltraSPARC T2 Plus processors enable extreme thread densities that help increase overall availability.
- **Hot-swappable redundant components.** Hot-swappable, mirrored disks and redundant PSUs can be quickly and easily changed out, increasing system uptime.
- **Accessible components for improved serviceability.** Front-accessible disk drives can be replaced quickly. The optional DVD/RW drive can also be removed without opening the top cover of the chassis. PSUs can be replaced without completely removing a system from a rack.
- **RAID options to balance storage capacity, availability, and cost.** The onboard LSI SAS controller supports RAID 0, 1, and 0+1.

- **Indicator LEDs on the front and back.** Readily visible LEDs make it easy to identify and isolate problems. Diagnostic LEDs are also included on the motherboard.
- **Oracle Integrated Lights Out Manager.** Standard on Sun Netra T5440 servers (and many other Sun servers), an integrated service processor provides powerful tools for local or remote system management. Oracle Integrated Lights Out Manager functionality simplifies administrative tasks, reduces onsite personnel requirements, and lowers operational costs.
- **Trusted Platform Module (TPM).** Platform security is becoming an important factor in enhancing system availability and reliability. Sun Netra T5440 servers include a TPM chip that connects to the service processor. The TPM chip is used to securely store certificates or encryption keys to help perform platform authentication and/or attestation (a process that indicates that a server is trustworthy and has not been breached).
- **DCAs.** Sun Netra T5440 servers are configured with DCAs and Watchdog Timers (WDTs) to detect and report conditions so that administrative action can be taken to avoid more-serious problems. Alarm I/O circuitry provides four "dry contact" relays. Each relay provides one pair of "dry" (that is, no electrical signal is provided by the system) closed-open contacts, which are controlled via user-defined application interfaces and I/O through the DB-15 DCA connector. (Typically, these relay connections are used to connect to an external network management controller to report a component failure.)
- **Telecordia NEBS Level 3 certification.** Telecordia NEBS Level 3 certification indicates that Sun Netra T5440 servers meet stringent reliability requirements, even in extremely harsh operating environments.

The robust design of Sun Netra T5440 servers makes these platforms ideal for critical telecommunications applications that demand continuous availability.

Carrier-Grade Software Support

Delivering reliable telecommunications services requires stable development tools, secure OSs, and robust systems management capabilities. However, with new technologies, it often takes time for software tools and critical applications to emerge, become stable, and evolve. Even though Sun Netra T5440 servers and UltraSPARC T2 Plus processors are relatively new, the processor technology on which they are based is mature and well tested. Moreover, Sun Netra T5440 servers come preloaded with the rock-solid foundation of Oracle Solaris 10, and they share binary compatibility with earlier SPARC systems. Sun Netra T5440 servers and Oracle Solaris 10 include a variety of sophisticated tools that let organizations consolidate and manage workloads easily while taking advantage of new technology and software innovations.

System Management Technology

As the number of systems grows in any organization, it becomes increasingly difficult to manage the infrastructure throughout its lifecycle. Effective system management requires both integrated hardware that can sense and modify the behavior of key system elements and advanced tools that can automate repetitive administrative tasks.

Oracle Integrated Lights Out Manager

As in other rackmounted and Sun Netra carrier-grade servers from Oracle, the Sun Netra T5440 system features a built-in service processor. Integrating the service processor on the motherboard facilitates remote power control and system monitoring, which simplifies system management. The service processor in this system is the same as in some other Sun Netra servers. As a result, Sun Netra T5440 servers integrate easily with existing management infrastructures, including onsite enterprise management frameworks or element managers.

The Oracle Integrated Lights Out Manager service processor

- Provides IPMI 2.0–compliant management functions to the server module's firmware, OS, and applications, as well as to IPMI-based management tools that access the service processor via the ILOM Ethernet management interface. It provides visibility of environmental sensors (both on the server module and elsewhere in the chassis).
- Manages inventory and environmental controls for the server module, including CPUs, FB-DIMMs, and power supplies, and supplies HTTPS, CLI, or SNMP access to this data.
- Supplies remote textual console interfaces.
- Provides a means to download upgrades to all system firmware.

The Oracle Integrated Lights Out Manager service processor allows the administrator to remotely manage the server independently of the OS and without interfering with system processing. The solution can send e-mail alerts of hardware failures and warnings, as well as other system-related events. The circuitry runs independently from the server, using server standby power. As a result, Oracle Integrated Lights Out Manager firmware and software continue to function when the server OS goes offline and even when the server is powered off.

Oracle Integrated Lights Out Manager monitors the following Sun Netra T5440 server conditions:

- CPU temperature conditions
- Hard drive presence
- Enclosure thermal conditions
- Fan speed and status
- Power supply status

- Voltage conditions
- Oracle Solaris watchdog, boot time-outs, and automatic server restart events

Sun Management Center Software

Sun Management Center software is an element management system for monitoring and managing Oracle's Sun systems environments. Sun Management Center software integrates with leading enterprise management systems to provide customers with a unified management infrastructure. The base package is free and provides hardware monitoring. Advanced applications (add-ons) extend the monitoring capability of the base package. Sun Management Center software provides

- Agents for managing Oracle Solaris (SPARC and x64/x86 platforms) and Linux OSs
- In-depth hardware and software diagnostics
- Aggregate CPU utilization reporting
- Event and alarm management for thousands of attributes
- Corrective action automation through scripts triggered by alarm thresholds
- Secure management controls for remote dynamic reconfiguration
- The ability to customize modules with a powerful, easy-to-use GUI

Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center is a family of software technologies that integrate virtualization and management capabilities, simplifying the oversight of physical and virtualized assets across heterogeneous environments. Built on proven technology and leveraging open source derived from the work of the Xen community, Oracle Enterprise Manager Ops Center addresses both desktop and server virtualization.

Oracle Enterprise Manager Ops Center is a cross-platform, high-efficiency, open source hypervisor for hosting multiple OSs. It is built from Oracle VM Server for SPARC capabilities (native to UltraSPARC T1, T2, and T2 Plus processor–based servers) in addition to technology from the Xen open source community.

Oracle Enterprise Manager Ops Center is a complete, highly scalable data center automation tool that simplifies discovery, provisioning, update, and management of physical and virtualized assets in cross-platform Linux and Oracle Solaris–based x86 and SPARC environments. It allows

- Better data center consolidation and resource management
- Updates of guest OSs and monitoring of virtual assets on a network

- Automated provisioning and updating of Linux and Oracle Solaris instances to increase availability and utilization, and to minimize downtime
- More-effective deployment, management, and monitoring of security and compliance in IT operations, both locally and remotely

Oracle Enterprise Manager Ops Center provides an IT infrastructure management platform for integrating and automating management of thousands of heterogeneous systems. To improve lifecycle and change management, the application supports the management of both the applications and the servers on which they run, including Sun Netra T5440 servers. Oracle Enterprise Manager Ops Center takes a step-by-step approach to unraveling the challenges of getting systems operational quickly. Its capabilities include

- **Discover.** As systems are added to the management network, administrators can use Oracle Enterprise Manager Ops Center to discover bare metal systems based on a given subnet address or IP range.
- **Group.** Given the number of systems to manage and the constant repurposing of systems, it is critical for IT organizations to find ways to group resources.Oracle Enterprise Manager Ops Center lets users logically group systems and then perform actions across a group of systems as easily as performing actions on a single system. Systems can be grouped by function (for example, Web servers versus clustered servers), administrative responsibility, or other categories based on organizational needs.
- **Provision.** Oracle Enterprise Manager Ops Center remotely installs OSs onto selected systems. Administrators can use this functionality to provision OSs onto bare metal systems or to reprovision existing systems. As the infrastructure lifecycle continues, Oracle Enterprise Manager Ops Center can update firmware and provision software packages and patches to selected systems.
- **Monitor.** When systems are up and running, administrators can use Oracle Enterprise Manager Ops Center to monitor system health, helping ensure that everything is running at optimal levels. The software provides detailed hardware monitoring for attributes such as fans, temperature, disk, and voltage usage, including bare metal systems. Oracle Enterprise Manager Ops Center also monitors OS attributes such as swap space, CPU, memory, and file systems. Administrators can define specific threshold levels and set preferred notification methods (including e-mail, pager, or Simple Network Management Protocol [SNMP] traps) for each monitored component, as business needs demand.
- **Manage.** Infrastructure lifecycle management extends beyond deployment and monitoring. Oracle Enterprise Manager Ops Center includes lights-out management capabilities (such as powering systems on and off, and remote serial console access) to help IT organizations manage the infrastructure remotely. Leveraging a Role-Based Access Control (RBAC) feature, organizations can grant permissions to specific users to perform specific management tasks.

- **Hybrid user interface.** Oracle Enterprise Manager Ops Center offers users a hybrid user interface (UI), accessible from the Web, that integrates GUI and CLI capabilities into a single console. With this hybrid UI, operations performed in the GUI are simultaneously reflected in the CLI and vice versa.

Scalability and Support for CoolThreads Technology

Oracle Solaris 10 is specifically designed to support the advanced capabilities of UltraSPARC T2 Plus processor–based systems. In fact, Oracle Solaris 10 provides key functionality for virtualization, optimal utilization, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments.

One significant advantage of systems based on the UltraSPARC T2 Plus processor is that they appear as familiar SMP systems to Oracle Solaris and applications. In addition, Oracle Solaris 10 OS incorporates features that help improve application performance on CMT architectures:

- **CMT awareness.** Oracle Solaris 10 is aware of the UltraSPARC T2 Plus processor hierarchy so that the scheduler can effectively balance the load across all the available pipelines. Even though it exposes each eight-core UltraSPARC T2 Plus processor as 64 logical processors, the Solaris OS understands the correlation between cores and the threads they support.
- **Fine-grained manageability.** Oracle Solaris 10 can enable or disable individual processors. In the case of the UltraSPARC T2 Plus processor, this ability extends to enabling or disabling individual cores and threads (logical processors). In addition, standard Oracle Solaris features such as processor sets provide the ability to define a group of logical processors and schedule processes or threads on them.
- **Support for virtualized I/O and accelerated cryptography.** Oracle Solaris contains technology to support and virtualize components and subsystems on the UltraSPARC T2 Plus processor, including support for the on-chip PCIe interface.

End-to-End Virtualization Technology

Virtualization technology is increasingly popular as organizations strive to consolidate disparate workloads onto fewer but more-powerful systems to improve utilization. Sun Netra T5440 servers are specifically designed for virtualization, providing very fine-grained division of multiple resources—from processing to virtualized networking and I/O. Most importantly, Sun's virtualization technology is provided as a part of the system, built into the processor technology itself.

A Multithreaded Hypervisor

As with the UltraSPARC T1 and UltraSPARC T2 processors, the UltraSPARC T2 Plus processor provides a multithreaded hypervisor—a small firmware layer that provides a stable virtual

machine architecture that is tightly integrated with the processor. Multithreading is crucial, because the hypervisor interacts directly with the underlying multithreaded processor. This architecture is able to context-switch between multiple threads in a single core, a task that can require additional software and considerable overhead in competing architectures that do not interface with the hardware directly.

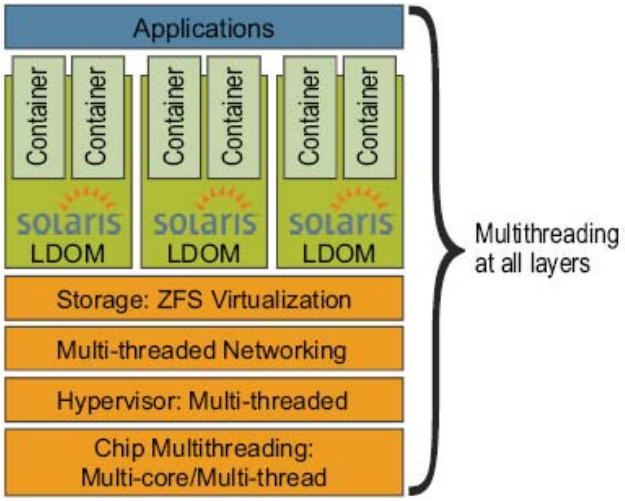

Figure 13. Sun provides parallelization and virtualization at every level of the technology stack.

Corresponding layers of virtualization technology are built on top of the hypervisor, as shown in Figure 13. The strength of Oracle's approach is that all of the layers of the architecture are fully multithreaded—from the processor up through applications that use the fully threaded Java application model. Far from new technology, Oracle Solaris has provided multithreading support since 1992—experience that has shaped technology decisions at other levels, ultimately resulting in a system that parallelizes and virtualizes at every level. In addition to the processor and hypervisor, Oracle provides fully multithreaded networking and the revolutionary, fully multithreaded Oracle Solaris Zettabyte File System (ZFS). Using Oracle Solaris ZFS, Oracle VM Server for SPARC, and Oracle Solaris Containers, multithreaded applications can receive exactly the resources that they need.

Oracle VM Server for SPARC

Supported in all Oracle's Sun servers using CMT technology, Oracle VM Server for SPARC provides full virtual machines that can run an independent OS instance and contain virtualized CPU, memory, storage, console, and cryptographic devices. Within the Oracle VM Server for SPARC architecture, each instance of the Oracle Solaris 10 is written to the hypervisor, which provides a stable, idealized, and virtualizable representation of the underlying server hardware to the OS in each domain. Each domain is completely isolated, and the maximum number of virtual machines created on a single platform relies upon the capabilities of the hypervisor rather than

on the number of physical hardware devices installed in the system. For example, the Sun Netra T5440 server with a single UltraSPARC T2 processor supports up to 128 domains, and each individual domain can run a unique OS instance.[3]

By taking advantage of domains, organizations gain the flexibility to deploy multiple OS instances simultaneously on a single platform. In addition, administrators can leverage virtual device capabilities to transport an entire software stack hosted on a domain from one physical machine to another. Domains can also host Oracle Solaris Containers to capture the isolation, flexibility, and manageability features of both technologies. Tightly integrating Oracle VM Server for SPARC with both the UltraSPARC T2 Plus processor and Oracle Solaris 10 increases flexibility, isolates workload processing, and improves the potential to optimize server utilization.

The Oracle VM Server for SPARC architecture includes underlying server hardware; hypervisor firmware; virtualized devices; and guest, control, and service domains. The hypervisor firmware provides an interface between each hosted OS and the server hardware. An OS instance that is controlled and supported by the hypervisor is called a guest domain. The control domain handles communication with the hypervisor, hardware platform, and other domains for creating and controlling guest domains. Guest domains are granted virtual device access via a service domain that controls both the system and the hypervisor, and also assigns I/O.

To support virtualized networking, Oracle VM Server for SPARC implements a virtual Layer 2 switch, to which guest domains (guest OSs) can be connected. Each guest domain can be connected to multiple vswitches, and multiple guest domains can be connected to the same vswitch. Vswitches can be associated with a real physical network port or exist without an associated port, in which case the vswitch only provides communications between domains in the same server (interdomain networking), thus saving valuable network resources. This approach also gives guest domains a direct communication channel to the network (see Figure 14). Each guest domain believes that it owns the entire NIC and the bandwidth it provides, yet in practice only a portion of the total bandwidth is allotted to the domain. As a result, every NIC can be configured as demand dictates, with each domain receiving bandwidth on an as-needed basis. Dedicated bandwidth can be made available by tying a vswitch device to a dedicated physical Ethernet port.

[3]Although possible, this practice is not recommended.

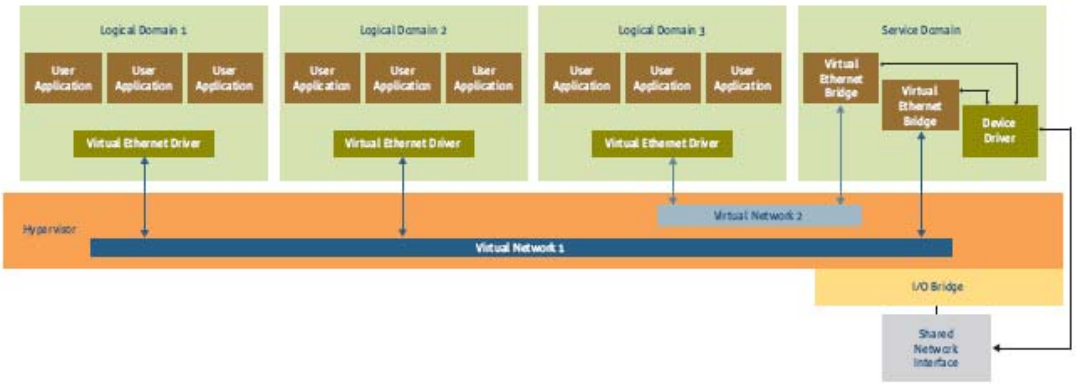

Figure 14. Data moves directly between a domain and a virtualized device.

Oracle Solaris Containers

Providing virtualization at the OS level, Oracle Solaris Containers represent a group of technologies that work together to efficiently manage system resources, virtualize the environment, and provide a complete, isolated, and secure runtime environment for applications.

Oracle Solaris Containers can be used to create an isolated and secure environment for running applications. Oracle Solaris Containers create a virtualized OS environment within a single instance of Oracle Solaris. Oracle Solaris Containers can be used to isolate applications and processes from the rest of the system, thereby enhancing security and reliability because processes in one zone are unable to interfere with processes running in another zone.

Resource management tools provided with Oracle Solaris help allocate resources such as CPUs to specific applications. CPUs in a multiprocessor system (or threads in the UltraSPARC T2 Plus processor) can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to a Solaris zone. Resource pools provide the ability to separate workloads so that CPU resource consumption does not overlap. They also enable a persistent configuration mechanism for processor sets and scheduling class assignment. In addition, the dynamic features of resource pools allow administrators to adjust system resources in response to changing workload demands.

Fault Management and Predictive Self-Healing

Oracle Solaris 10 introduces a new architecture for deploying systems and services that are capable of fault management along with system and service recovery. Oracle Solaris Predictive Self Healing is an innovative capability in Oracle Solaris 10 that automatically diagnoses, isolates, and recovers from many hardware and application faults. As a result, business-critical applications and essential system services can continue uninterrupted in the event of software failures, major hardware component failures, and even misconfigured software.

- **Oracle Solaris fault manager.** The Oracle Solaris fault manager facility collects data relating to hardware and software errors. This facility automatically and silently detects and diagnoses an underlying problem, with an extensible set of agents that automatically respond by taking a faulty component offline. Easy-to-understand diagnostic messages link to articles in Oracle's knowledge base to guide administrators through corrective tasks that require human intervention. The open design of Oracle Solaris fault manager also permits administrators and field personnel to observe the activities of the diagnostic system. With Oracle Solaris fault manager, the overall time from a fault condition to automated diagnosis and any necessary intervention is greatly reduced, increasing application uptime.
- **Oracle Solaris Service Management Facility.** Oracle Solaris Service Management Facility creates a standardized control mechanism for application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services can then be automatically restarted if an administrator accidentally terminates them, a software programming error aborts them, or an underlying hardware problem interrupts them. In addition, Oracle Solaris Service Management Facility reduces system boot time by as much as 75 percent by starting services in parallel according to their dependencies. An Undo feature helps safeguard against human errors by permitting easy change rollback. These capabilities are also simple to deploy: developers can convert most existing applications to take full advantage of SMF features by adding a simple XML file for each application.

Predictive self-healing and fault management provide the following specific capabilities on Sun Netra T5440 servers:

- **CPU "offlining"** takes cores deemed faulty offline, where they're stored in the resource cache and remain offline on reboot unless the processor has been replaced (in which case the CPU is cleared from the resource cache).
- **Memory page retirement** retires pages of memory marked as faulty. Such pages are then stored in the resource cache and remain retired on reboot unless the offending FB-DIMM has been replaced (in which case affected pages are cleared from the resource cache).
- **I/O retirement** logs errors and faults.
- **fmlog** logs faults detected by the system.

Cool Tools for SPARC: Performance and Rapid Time to Market

No matter how compelling new hardware or OS platforms may be, organizations must be assured that the costs and risks of adoption are in line with the rewards they promise. In particular, organizations want to continue to leverage the considerable advantages of popular commercial and open source software. Developers want to use familiar compilers and basic development tools. And administrators do not want to spend valuable time getting applications

to run optimally in a new environment. Oracle's Cool Tools Program is designed to minimize the cost and risk of moving Web-tier environments to new platforms.

Application Selection

Application selection helps identify applications that stand to benefit from CoolThreads technology. The CoolThreads Selection Tool (coolst) helps determine application suitability for both UltraSPARC T1 and UltraSPARC T2 architectures, reducing the risk of investment decisions. The tool measures application floating-point content along with the number of lightweight processes (threads) to determine potential parallelism.

Development

Developers need to be able to build, test, and evaluate applications, producing the most effective code while advancing their productivity with chosen tools. Tools for development on Sun Netra T5440 servers include

- **GCC for SPARC System.** Tuned and optimized for SPARC systems, GCC for SPARC System complements the popular GCC compiler suite, delivering up to three times the performance of compiled applications with even greater levels of reliability. At the same time, GCC for SPARC System is 100 percent compatible with GCC, supporting all ABIs, language extensions, and flags.
- **Oracle Solaris Studio 12.** Oracle Solaris Studio 12 provides developers with record-setting high performance, optimizing C, C++, and FORTRAN compilers for Oracle Solaris on SPARC and x86/x64 platforms. Command-line tools and NetBeans (an integrated development environment for Java-language programming) are provided for application performance analysis and debugging of mixed source language applications. In addition to providing multiplatform support, Sun Studio 12 compilers are compatible with GCC, Visual C++, C99, OpenMP, and FORTRAN 2003.

In addition, developers may take advantage of numerous third-party and Oracle APIs to address application requirements, simplify common programming tasks, and enhance platform independence. Libraries of typical media processing operations, for example, are available on multiple platforms, including those designed to deliver optimal performance on UltraSPARC processor–based platforms. The mediaLib library is a low-level performance library available from Oracle for multimedia applications; functions in mediaLib are accelerated with VIS technology on UltraSPARC-based systems. (These same functions are also supported—although without VIS acceleration—on other SPARC-based and x86-based systems running the Solaris OS.) Media codec functions are also available in libraries such as CodecLib for Solaris and in Java Advanced Imaging APIs, and perform well on the UltraSPARC T2 Plus processor because of its robust floating-point and throughput processing capabilities.

Tuning and Debugging

Administrators and developers alike need to monitor, analyze, and tune applications under real-world conditions. The following tools aid with tuning and debugging:

- **Corestat.** This online tool monitors core utilization of the UltraSPARC T2 Plus processor, providing a more accurate measure of processor and system utilization than tools that only measure the utilization of individual threads. Implemented as a Perl script, corestat aggregates instructions executed by all the threads on a single core, revealing the cycles per instruction of key workloads and indicating where more tuning is needed.
- **Automatic Tuning and Trouble-Shooting System (ATS).** In the interest of automating application tuning, ATS automatically reoptimizes and recompiles binaries with no need for source code. ATS identifies the inadequate optimization and then automatically rebuilds the application with the correct options for optimization. ATS is a plug-in for GCC4SS and Sun Studio software.

Deployment

CoolTools deployment elements provide applications that are already optimized for CoolThreads technology and save critical time in configuring systems for performance and consolidation. Deployment elements include

- **Cool Tuner.** Providing an onsite "virtual" tuning expert, CoolTuner improves system performance by automatically applying current best practices, including both patching and tuning. Depending on administrator experience, CoolTuner can save hours to weeks of effort tuning Sun Fire or Sun Netra servers based on CoolThreads technology.
- **Cool Stack.** Cool Stack represents a collection of the most commonly used free and open source applications, preoptimized for Sun Fire and Sun Netra servers based on CoolThreads technology that run Oracle Solaris. Including such popular applications as Apache, Perl, PHP, Squid, Tomcat, and MySQL, these applications have been recompiled with Sun Studio 12 compilers to deliver a 30 to 200 percent performance improvement over standard binaries compiled with GCC. Cool Stack applications also bring performance benefits to any SPARC processor–based system.
- **Consolidation Tool for Sun Fire Servers.** Powerful Solaris Containers offer myriad consolidation possibilities, and the Consolidation Tool for Sun Fire Servers speeds their deployment. With a wizard-based GUI, this tool simplifies and automates the installation of consolidated applications, enabling even novice administrators to create a fully virtualized and consolidated environment using Oracle Solaris Containers. The result is fast and high-quality consolidated deployments using Sun Fire or Sun Netra servers.

Java Platform, Enterprise Edition (Java EE)

The software industry has traditionally offered point products that solve specific parts of a problem, leaving it to customers to integrate those products into a solution that can support their business applications. Organizations don't purchase their OSs by assembling core components such as drivers, schedulers, command, and administration utilities, and it doesn't make sense for them to assemble and integrate traditional middleware this way either.

Using world-class software, Oracle redefines the software system from the OS up through the J2EE specification layer. Customers can write their business applications to Java software standards, leveraging Java EE network services, and Oracle delivers the end-to-end solution to run them.

Oracle offers the following products built to optimize Java EE environments:

- **Sun Java System Access Manager.** Open, standards-based access control; single sign-on; and federation services that help control costs and minimize the security risks of conducting business more openly.
- **Oracle Communications Enterprise Mobility Server.** A robust, commercial J2EE 4–compliant application server that makes building robust, scalable enterprise applications easier than ever, and is the perfect platform for implementing SOA and Web 2.0 applications.
- **Oracle Directory Server Enterprise Edition.** A secure, highly available, scalable, and easy-to-manage directory infrastructure that effectively manages identities in dynamic environments.
- **Java Studio Creator.** Rapid visual Web application and portlet development.
- **Java Studio Enterprise.** An award-winning IDE for enterprise architects and developers.
- **Oracle Communications Messaging Exchange Server.** A leading business integration messaging server designed to deliver the exceptional scalability, reliability, and advanced security features necessary for large-scale enterprise deployments.
- **Oracle Solaris Cluster.** A multisystem, multisite disaster recovery solution that manages the availability of applications services and data across local, regional, and vastly dispersed data centers.
- **Oracle Solaris Studio.** A high-performance compiler that optimizes C, C++, and Fortran for Oracle Solaris on SPARC, and both Oracle Solaris and Linux on x86 platforms.
- **Oracle iPlanet Web Server.** A leading enterprise Web server, engineered to meet the stringent requirements of organizations that use Web technologies as a competitive advantage.
- **Oracle iPlanet Web Proxy Server.** A powerful system for caching and filtering Web content, boosting network performance, and reducing user wait times.

Sun Netra Data Plane Suite

Combined with UltraSPARC T2 Plus processor technology and Oracle VM Server of SPARC technology, the Sun Netra Data Plane Suite helps simplify the consolidation of control and data plane applications. In this way, network processing elements can leverage Oracle's general-purpose volume market hardware—like the Sun Netra T5440 server—to change the economics of control and data plane solutions to deliver telecommunications applications.

Sun Netra Data Plane Suite is a high-performance packet-processing engine that allows providers to quickly and cost-effectively develop and deploy data plane applications—enabling the aggregation, transport, and routing of voice, video, and data in converged telecommunications networks. Within Sun Netra Data Plane Suite, high-level language tools produce explicitly parallelized American National Standards Institute (ANSI) C code, speeding application development and providing a degree of portability that enhances vendor and platform independence. In addition, a lightweight runtime environment included with Sun Netra Data Plane Suite exploits the UltraSPARC T2 Plus processor architecture to deliver line rate packet processing speeds for data plane applications.

Standards-Based Data Plane Application Development

Sun Netra Data Plane Suite provides developers with standards-based software tools for rapid creation of data plane applications. In addition, the Sun Netra Data Plane Suite -optimized runtime environment exploits multithreaded partitioning firmware for high-speed code execution on UltraSPARC T1, UltraSPARC T2, and UltraSPARC T2 Plus multithreaded processors. The following features of the scalable Sun Netra Data Plane Suite software framework enable fast path network processing:

- Event-driven scheduling with run to completion states
- Explicit parallelization
- Static memory allocation
- Code generation based on hardware description and mapping
- Efficient communication pipes between pipeline states

Using the Sun Netra Data Plane Suite compiler enables programmers to develop scalable, high-performance C applications for embedded multiprocessor target architectures. A system-level view of the application and the following features contribute to the ability of the Sun Netra Data Plane Suite compiler to deliver superior code validation and optimization:

- Inclusion of characteristics of the targeted hardware and software system architecture through execution of a user-supplied architecture specification
- Simultaneous examination of multiple source file sets and target architecture relationships
- System-level contextual recognition of APIs used in application code and software generation

To further simplify and accelerate the creation of packet-processing applications, Sun Netra Data Plane Suite includes a library of reference applications that can be reused and integrated into full-scale network elements. For example, it includes RLP, IP forwarding (ipfwd), user datagram protocol (UDP), and packet classification (PacketClassifier) reference applications. New libraries can also be added over time, based on the work of a community of application developers and Oracle's engineering efforts. Sun Netra Data Plane Suite can help to shorten development time from years to months without compromising performance.

A Lightweight Runtime Environment on General-Purpose Platforms

Applications can achieve excellent performance by leveraging the high-performance, lightweight Sun Netra Data Plane Suite runtime environment and exploiting the virtually unmatched thread density and SoC design of UltraSPARC T2 Plus processors. The NDPS lightweight runtime environment executes directly on UltraSPARC T2 Plus processor threads, enabling packet-processing throughput at 10-Gb line rates and beyond. Avoiding the performance limitations of general-purpose OSs, the lightweight runtime environment has no scheduler or interrupt handler and performs no context switching. Every thread runs to completion without time slicing, making parallel execution extremely scalable and enabling line rate performance and linear scaling. By using the Sun Netra Data Plane Suite lightweight runtime environment on servers with UltraSPARC T2 Plus processors, organizations gain the ability to implement data plane packet processing on a platform with an economic cost structure that enables deployment throughout the network.

Profiling and Redistribution of Data Plane Processing

In many cases, applications are initially coded and debugged on a general-purpose OS. After profiling the software on the OS, forwarding code is moved to Sun Netra Data Plane Suite. The resulting Sun Netra Data Plane Suite code represents embedded packet processing code, and as such it is beneficial to optimize it to leverage hardware threading. Sun Netra Data Plane Suite's multithreaded profiling capabilities can optimize assignment of threads by determining the number of threads, the number of receive-DMA channels to utilize, and the depth of the packet-processing pipeline.

In addition to standard optimizations, the Sun Netra Data Plane Suite profiler enables the collection of critical data during execution. With the ability to profile CPU utilization, instruction counts, I/O wait times, data cache misses, secondary cache misses, memory queue, memory cycles, and more, the Sun Netra Data Plane Suite profiler guides and simplifies performance tuning. The Sun Netra Data Plane Suite profiler also uses the special counters and resources available in the hardware to collect critical information. In fact, profiler API calls can be inserted at desired places to start collecting or updating profiling data.

Tailoring the Sun Netra Data Plane Suite lightweight runtime environment for specific workloads and using interprocess communications between different control and data plane processes can

reduce overall network traffic and lead to greater performance. As a result, organizations can leverage off-the-shelf OSs on Oracle platforms for provisioning, control plane, and exception processing, and only a small amount of code moves to a dedicated packet-processing domain within the same or separate server.

In addition to the variety of other development tools available for the Sun Netra T5440 server, Sun Netra Data Plane Suite makes it extremely easy for telecommunications companies to consolidate control and data plane applications. Because the software automatically takes advantage of multiple threads and domain capabilities in UltraSPARC T2 Plus processors, applications can run safely and securely within the same system while balancing resources to optimize utilization.

Conclusion

Oracle's Sun Netra T5440 servers offer incredible reliability and system density—with robust compute, memory, networking, storage, and I/O capabilities in a compact 4U form factor. Leveraging UltraSPARC T2 Plus processor technology and Oracle's engineering expertise in chassis and systems design, these platforms deliver new levels of performance—and new levels of performance-per-watt—in a rackmountable, NEBS-certified system. Deploying these servers creates an agile and reliable infrastructure that scales easily to deliver new, integrated multimedia application services within a small footprint while delivering mission-critical service levels.

Sun Netra T5440 servers provide expandable, high-capacity resources for demanding Web infrastructure, database, and server consolidation and virtualization initiatives. These carrier-grade systems are ideal for installations where reliability, performance, density, and energy conservation are paramount. Given the high throughput capabilities and energy efficiency of these servers, IT departments can easily consolidate workloads and improve utilization while also preserving their investments in applications.

Sun Netra T5440 Server Architecture
March 2010

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

SOFTWARE. HARDWARE. COMPLETE.