An Oracle White Paper
April 2010

Oracle Sun SPARC Enterprise T5440 Server Architecture

ORACLE

Introduction

Data centers have increasingly become the areas of focus for organizations needing to maximize their IT services' flexibility, performance, and overall return on investment (ROI). Most organizations are also fully aware of their need to minimize their exposure to reliability and serviceability issues. Choosing the right server architecture and deployment scenarios to satisfy these requirements represents an increasingly difficult challenge. To address these issues, Oracle's Sun SPARC Enterprise T5440 server provides a combination of freely available and no-cost virtualization technologies, excellent reliability, availability, and serviceability (RAS) ratings, very high performance hardware, and significant data center efficiency gains.

The Oracle Sun SPARC Enterprise T5440 server comfortably extends its capabilities deep into the data center, hosting core applications such as enterprise resource planning (ERP), customer relationship management (CRM), and other database-intensive applications, as well as the databases themselves—applications once reserved for midrange and high-end enterprise servers where high performance and reliability came at very considerable cost. The performance and capabilities of the Oracle Sun SPARC Enterprise T5440 server allow considerable savings in terms of initial acquisition costs as well as ongoing service and operation costs. At the same time, these servers help organizations maintain or improve performance and reliability when combined with Oracle's clustering technologies, advanced system management, and license and fee-free virtualization technologies.

Employing Oracle's UltraSPARC T2 Plus processors—evolved from the industry's first massively threaded systems-on-a-chip (SoC)—the Oracle Sun SPARC Enterprise T5440 server offers the breakthrough performance and energy efficiency required to address demanding data center IT services challenges. As the industry's first quad-socket system to support Oracle Solaris CoolThreads chip multithreading (CMT) technology, the Oracle

Sun SPARC Enterprise T5440 server supports up to 256 threads in only four rack units (4RU)—providing massive computational density while staying within constrained envelopes of power and cooling. Beyond computational capacity, this server also provides very large memory support (up to 512 GB) and extremely high levels of I/O throughput, with up to four times the I/O capacity of earlier single-socket UltraSPARC T2 implementations.

Oracle's Sun SPARC Enterprise T5440 servers employ very high levels of integration and optimized system design to reduce latency, lower costs, and improve security and reliability. With these advantages and Oracle’s embedded no-cost virtualization software, entire multitier applications and services can be consolidated onto a single Oracle Sun SPARC Enterprise T5440 server. This white paper describes the architecture of the Oracle Sun SPARC Enterprise T5440 server as well as that of the innovative UltraSPARC T2 Plus processors.

.

The Evolution of Chip Multithreading

By any measure, Oracle's first-generation chip multithreading (CMT) processors were an unprecedented success. Delivering upward of five times the throughput in a quarter of the space and using a quarter of the power, systems employing these processors have been rapidly welcomed and accepted. Many organizations have been rewarded with long-term ROI savings and performance benefits as a result of adopting Oracle's CMT technology. Now, CMT technology is evolving rapidly to meet the constantly changing demands of a wide range of IT services and enterprise data center applications.

Business Challenges for Information Technology Services Deployments

In tough economic times, organizations of all kinds are increasingly sensitive to the costs associated with providing IT services. In the past, building in flexibility meant overprovisioning very expensive systems so that organizations could quickly and easily deal with rapid growth or changes in the direction of the business or customer base. Today, that's simply not an option for many data center planners—which is why the ability to grow and change dynamically must be implicit in the server design. Oracle provides the required flexibility to grow and scale using the unique capabilities of Oracle's CMT processors and the servers that are designed to use them.

- **Securing the enterprise at speed.** Organizations are increasingly interested in securing all communications with their customers and partners. Given the risks, end-to-end encryption is essential to inspire confidence in security and confidentiality. Encryption is also increasingly important for storage, helping to secure stored and archived data even as it provides a mechanism to detect tampering and data corruption. Unfortunately, the computational costs of increased encryption can increase the burden on already overtaxed compute resources. Security also needs to take place at line speed, without introducing bottlenecks that can impact the customer experience or slow transactions. Solutions must help ensure security and privacy for clients and bring business compliance for the organization, all without impacting performance or increasing costs.
- **Driving data center virtualization and ecoefficiency.** Coincident with the need to scale services, many data centers are recognizing the advantages of deploying fewer standard platforms to run a mixture of commercial and technical workloads. This process involves consolidating under-utilized and often-sprawling server infrastructure with effective virtualization solutions that enhance business agility, improve disaster recovery, and reduce operating costs. This focus can help reduce energy usage and break through data center capacity constraints by improving the amount of realized performance for each watt of power the data center consumes. Ecoefficiency provides tangible benefits, reducing our carbon footprint to meet legislative and corporate social responsibility goals while improving the economy of the organization paying the electric bill. As systems are consolidated onto more dense and capable computing infrastructure, demand for data center real estate is also reduced.

With careful planning, this approach can also improve service uptime and reliability by reducing hardware failures resulting from excess heat load. Servers with high levels of standard RAS are now considered a requirement.

- **Building out for application scale.** Enterprise applications require stability and performance, but systems must also have the fundamental ability to scale and cope with rapidly increasing scale. In today's economic climate, organizations must accelerate time to market and time to service while delivering scalable high-quality and high-performance applications and services. Many need to start small but be able to scale quickly to accommodate new customers and deliver innovative new services—often implying a doubling of capacity in months rather than years. At the same time, organizations must reduce their environmental impact by working within the power, cooling, and space available in their current data centers. Operational costs, too, are receiving new scrutiny, as are system administrative costs, which can account for up to 40 percent of an IT budget. Simplicity and speed are paramount, giving organizations the ability to respond quickly to dynamic business conditions. Open platforms built around open standards help provide maximum flexibility while reducing both entry and exit costs.

Rule-Changing Chip Multithreading Technology

Addressing these challenges has outstripped the capabilities of traditional processors and systems, requiring a fundamentally new approach.

Moore's Law and the Diminishing Returns of Traditional Processor Design

An often-quoted tenant of Moore's law states that the number of transistors that will fit in a square inch of integrated circuitry will approximately double every two years—a pace that's held firm for more than three decades, driving processor performance to new heights. Processor manufacturers have long exploited these 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 multiple-gigahertz frequencies, they have largely resulted in complex, hot, and power-hungry processors

that are not well-suited to the types of workloads often found in modern data centers. In fact, many data center workloads are simply unable to take advantage of the hard-won ILP provided by these processors. Applications with high shared memory and high simultaneous user or transaction counts are typically more focused on processing a large number of simultaneous threads (thread-level parallelism, or TLP) than running a single thread as quickly as possible (ILP).

Making matters worse, the majority of ILPs in existing applications have already been extracted, and further gains promise to be small. In addition, microprocessor frequency scaling itself has leveled off because of microprocessor power issues. 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 2 GHz to 3 GHz range as a result. Deploying pipelined superscalar processors requires more power, limiting this approach by the fundamental ability to cool the processors.

Chip Multiprocessing with Multicore Processors

To address these issues, many in the microprocessor industry 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 chip multiprocessors, CMP). The individual processor cores introduced by many CMP designs, however, offer no greater performance than previous single-processor chips and, in fact, have been observed to run single-threaded applications more slowly than single-core processor versions. The aggregate chip performance does increase, however, because multiple programs (or multiple threads) can be accommodated in parallel (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, although these designs provide some additional throughput and scalability, they can also consume considerable power and generate significant heat—without a commensurate increase in overall performance.

Chip Multithreading with CoolThreads Technology

Oracle engineers were early to recognize the disparity between processor speeds and memory access rates—specifically that while processor speeds continue to double every two years, memory speeds have typically doubled only every six years. As a result, memory latency now dominates much application performance, erasing even very impressive gains in clock rates. This growing disconnect is the result of memory suppliers focusing on density and cost—rather than speed—as their design center.

Unfortunately, this relative gap between processor and memory speeds leaves ultrafast processors idle as much as 85 percent of the time as they wait for memory transactions to complete.

Ironically, as traditional processor execution pipelines get faster and more complex, the effect of memory latency grows—and fast, expensive processors spend more cycles doing nothing. Worse still, these idle processors continue to draw power and generate heat. Thus, it's easy to see that frequency (gigahertz) is truly a misleading indicator of real performance.

First introduced with the UltraSPARC T1 processor, chip multithreading (CMT) takes advantage of CMP advances but adds a critical capability—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 as other threads stall for memory. Figure 1 illustrates the difference among CMP, fine-grained hardware multithreading (FG-MT), and CMT. The key to this approach is that each core in a CMT processor is designed to switch between multiple threads on each clock cycle. As a result, the processor's execution pipeline remains active, continuing to do real work even as memory operations for stalled threads continue in parallel.

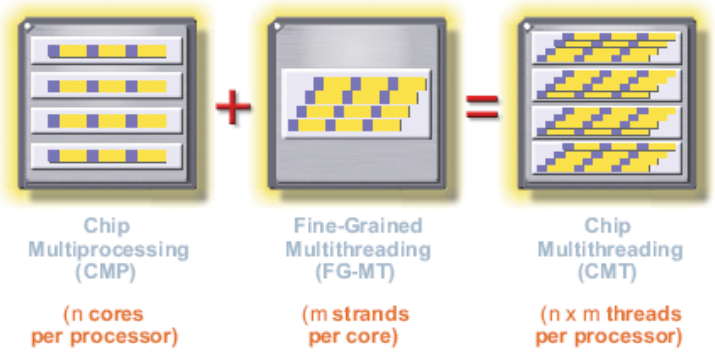

Figure 1. CMT combines CMP and fine-grained hardware multithreading.

CMT provides real value because it increases the ability of the execution pipeline to do actual work on any given clock cycle. Use of the processor pipeline is greatly enhanced because a number of execution threads now share its resources. The negative effects of memory latency are effectively masked, since the processor and memory subsystems remain active in parallel to the processor execution pipeline. Because these individual processor cores implement much-simpler pipelines that focus on scaling with threads rather than frequency (emphasizing TLP over ILP), they are also substantially cooler and require significantly less electrical energy to operate. This innovative approach results in CoolThreads processor technology—multiple physical instruction execution pipelines (one for each core), with multiple active thread contexts per core. UltraSPARC T2 and UltraSPARC T2 Plus processors feature two execution pipelines per core to further boost scalability.

The UltraSPARC T2 Plus Processor

Unlike complex single-threaded processors, CMT processors use the available transistor budget to implement multiple hardware-multithreaded processor cores on a chip die. The UltraSPARC T2 processor took the CMT model to the next level, providing up to eight cores per processor, with each core supporting up to eight threads via two independent pipelines per core—effectively doubling the throughput of the UltraSPARC T1 processor without increasing frequency. Now, UltraSPARC T2 Plus processors extend the proven benefits of the CMT architecture into multisocket implementations to serve mission-critical and online transaction processing (OLTP) database workloads. Unlike many single-threaded processor designs, UltraSPARC T2 Plus processors use the available transistor budget to implement a massively threaded system-on-a-chip (SoC), with a single processor die hosting

- Up to 64 threads per processor (up to eight cores supporting eight threads each)
- On-chip Level 1 (L1) and L2 caches
- Per-core floating-point capabilities
- Per-core cryptographic acceleration
- On-chip PCI Express (PCI-E) interface
- On-chip cache coherency logic and links

Through SoC design, the UltraSPARC T2 Plus processor significantly enhances the general-purpose nature of the CPU—building in eight floating-point units (one per core). Enhanced floating-point capabilities open the UltraSPARC T2 Plus processor to the world of compute-intensive applications as well as the traditionally CMT-friendly data center throughput applications. No-cost security and cryptographic acceleration is provided by the on-chip, per-core cryptographic accelerators. In addition, the ability to move data in and out of the UltraSPARC T2 Plus processor is significantly aided by an integrated on-chip PCI-E interface, memory management units, cache coherency logic, and coherence links that facilitate multisocket system designs.

Oracle Sun SPARC Enterprise T5440 Server

The Oracle Sun SPARC Enterprise T5440 server is designed to leverage the considerable resources of the UltraSPARC T2 Plus processor into a cost-effective, multisocket computing platform. Quad-socket Oracle Sun SPARC Enterprise T5440 servers deliver up to twice the throughput of dual-socket Oracle Sun SPARC Enterprise T5240 servers, while leading competitors in terms of performance, performance per watt, and SWaP performance (as evaluated by the space, watts, and performance metric detailed later in this section). Oracle Sun SPARC Enterprise T5440 servers (Figure 2) share a proven chassis design with the x86-based Sun Fire X4600 server. Enterprise features and data center design focus combine with integrated cryptographic acceleration and on-chip I/O technology.

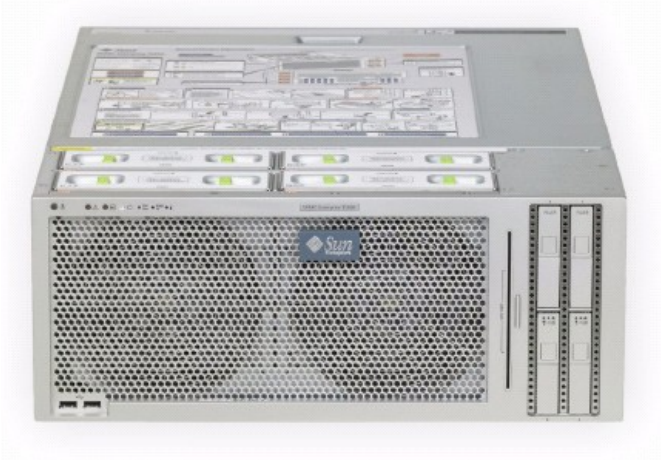

Figure 2. The Oracle Sun SPARC Enterprise T5440 server offers up to four UltraSPARC T2 Plus processors in only four rack units.

Overview

As a consolidation platform, Oracle's Sun SPARC Enterprise T5440 servers excel by coupling scalable and balanced system design with embedded no-cost virtualization technologies such as Oracle VM Server for SPARC (previously called Sun Logical Domains) and Oracle Solaris Containers. With up to four sockets for UltraSPARC T2 Plus processors, large memory support, and significant I/O throughput, the Oracle Sun SPARC Enterprise T5440 server can provide infrastructure that has typically required an entire rack in only 4RU. With per-core floating-point capabilities, the benefits of CMT are also available to technical workloads in addition to multithreaded commercial workloads. Designed to complement Oracle's Sun SPARC Enterprise T5120/T5220 and T5140/T5240 servers, Oracle's Sun SPARC Enterprise T5440 servers bring the capabilities of CMT systems to challenging midrange workloads.

Accelerated Time to Market

Oracle's Sun SPARC Enterprise T5440 servers running the Oracle Solaris operating system (OS) provide full binary compatibility with earlier UltraSPARC systems, preserving investments and speeding time to market. Oracle Solaris Cool Tools help accelerate application selection, profiling, testing, tuning, and debugging, along with deployment of key applications on CMT systems.

Industry-Leading Tools for Virtualization and Consolidation

Oracle's CMT technology is ideal for consolidation, providing low-level multithreading support for virtualization at every layer of the technology stack. Oracle VM Server for SPARC exploit the UltraSPARC T2 Plus processor's 64 threads per socket (up to 256 threads per system) to support multiple guest operating system instances, while Containers provide virtualization within a single

Oracle Solaris instance. The advanced and scalable Oracle Solaris ZFS file system provides storage virtualization for storage and considerable scalability.

System and Data Center Reliability

Reliability is key to keeping applications available and costs down. With greater levels of integration provided by an SoC design, the Oracle Sun SPARC Enterprise T5440 server offers greatly reduced part counts and provides commensurately higher levels of RAS. Lower power consumption and higher performance per watt greatly reduce generated heat loads and the associated issues they cause. Intelligent fan control and vibration monitoring help anticipate issues. Technologies such as Oracle Solaris predictive self-healing are integrated with the hardware and help keep systems available.

Efficient and Predictable Scalability

Each Oracle Sun SPARC Enterprise T5440 server provides support for up to 256 threads, up to 512 GB of memory, and massive I/O bandwidth in just 4RU. As a result, these servers are ideal platforms for consolidation and virtualization. Providing significant computational and I/O throughput, and typically consuming less than 1,800 watts of power, Oracle's Sun SPARC Enterprise T5440 servers can help organizations meet their growing IT and Web infrastructure needs within stringent power, cooling, and space constraints.

Zero-Cost Security

With attempted electronic intrusion and theft at an all-time high, it's never been more important to provide secure communications and data protection. Oracle's Sun SPARC Enterprise T5440 servers allow you to do just that: With up to eight integrated cryptographic accelerators on each UltraSPARC T2 Plus processor, there's simply no need to send plaintext on the network or to store plaintext in storage systems. Oracle's Sun SPARC Enterprise T5440 servers support many more crypto-operations per second than competing systems with dedicated crypto-accelerator cards, and they do so with minimal impact to system overhead.

An Ecoefficiency Tradition

Oracle's Sun Fire/Sun SPARC Enterprise T1000 and Oracle's Sun Fire/Sun SPARC Enterprise T2000 servers were the industry's first ecoresponsible servers. Oracle Sun SPARC Enterprise T5440 servers continue this tradition by extending these benefits to larger-scale midrange workloads. In addition, the UltraSPARC T2 Plus processor incorporates unique power management features at both the core and memory levels of the processor.

Simplified Management

Each Oracle Sun SPARC Enterprise T5440 server provides an Integrated Lights Out Manager (ILOM) service processor, compatible with other Oracle volume servers. ILOM provides a

command-line interface (CLI), a Web-based graphical user interface (GUI), and Intelligent Platform Management Interface (IPMI) functionality to aid with out-of-band monitoring and administration. ILOM on these systems also provides an Advanced Lights Out Management (ALOM) backward-compatibility mode for administrators familiar with Oracle Sun Fire/Sun SPARC Enterprise T1000 servers and Oracle Sun Fire/Sun SPARC Enterprise T2000 servers.

The Industry's Most Open Platform

Oracle Sun SPARC Enterprise servers represent the industry's most open platform, providing the only mainstream processor and hypervisor offered under the GNU General Public License (GPL). These systems offer a choice of operating systems, including Oracle Solaris, Linux, and BSD variants. Oracle Solaris is free and open, offering full binary compatibility and enterprise-class features.

Innovative System Design

Oracle understands that data centers have unique and pressing needs that require the attention of system designers. Density, performance, and scalability are all essential considerations, but systems must also be serviceable and well-adapted to modern data center strategies that consider power, cooling, and serviceability. The underlying design of the Oracle Sun SPARC Enterprise T5440 servers is based on principles that extend across Oracle's volume x64 and SPARC server platforms. These include

- **Common chassis design.** Shared chassis design leverages key system innovations across multiple architectures, provides for common components and subassemblies, and greatly simplifies administration for those deploying multiple processor architectures.
- **Maximum compute density.** Oracle's Sun servers provide leading density in terms of CPU cores, memory, storage, and I/O. In many cases, this density makes it possible for the 4RU Oracle Sun SPARC Enterprise T5440 server to replace an entire rack of legacy or competitive servers.
- **Common, shared management.** The Oracle Sun SPARC Enterprise T5440 server is designed for ease of management and serviceability with service processors shared by other Oracle volume server platforms. Systems and components are designed for easy identification, and hot-swap components facilitate online replacement.
- **Continued investment protection.** Oracle designs for maximum investment protection so that previously developed applications simply run without modification.

Oracle 's Sun SPARC Enterprise T5140, Sun SPARC Enterprise T5240, and Sun SPARC Enterprise T5440 Feature Comparison

Table 1 provides a detailed comparison of the features of the Oracle Sun SPARC Enterprise T5140, Oracle's Sun SPARC Enterprise T5240, and Oracle's Sun SPARC Enterprise T5440 servers.

TABLE 1. ORACLE SUN SPARC ENTERPRISE T5140, T5240, AND T5440 SERVER FEATURES

FEATURE	ORACLE SUN SPARC ENTERPRISE T5140 SERVER	ORACLE SUN SPARC ENTERPRISE T5240 SERVERr-	ORACLE SUN SPARC ENTERPRISE T5440 SERVER
CPU	Dual four-, six-, or eight-core 1.2 GHz or eight-core 1.4 GHz UltraSPARC T2 Plus processors	Dual four-, six-, or eight-core 1.2 GHz or eight-core 1.4 GHz or 1.6 GHz UltraSPARC T2 Plus processors	One to four eight-core 1.2 GHz, 1.4 GHz, or 1.6 GHz UltraSPARC T2 Plus processors
Threads	Up to 128	Up to 128	Up to 256
Memory capacity	Up to 128 GB (2, 4, or 8 GB FB-DIMMs)	Up to 128 GB (2, 4, or 8 GB FB-DIMMs)[1]	Up to 512 GB (2, 4, or 8 GB FB-DIMMs)
Maximum internal disk drives	Up to eight SFF 2.5-inch SAS 73, 146, or 300 GB disk drives, or 32 GB SSDs, RAID 0/1[2]	Up to 16 SFF 2.5-inch SAS 73, 146, or 300 GB disk drives, or 32 GB SSDs, RAID 0/1[3]	Up to four SFF 2.5-inch SAS 73 or 146 GB disk drives, or 32 GB SSDs, RAID 0/1
Removable and pluggable I/O	• Slot loading DVD+/-RW • Four USB 2.0 ports	• Slot loading DVD+/-RW • Four USB 2.0 ports	• Slot loading DVD+/-RW Four USB 2.0 ports
PCI	• One x8 PCI Express slot • Two x8 PCI Express or XAUI combo slots[4]	• Four x8 PCI Express slot • Two x8 PCI Express or XAUI combo slots[4]	• Six x8 PCI Express slots • Two x8 PCI Express or XAUI combo slots[4]
Ethernet	• Four onboard Gigabit Ethernet ports (10/100/1000) • Two 10 Gb Ethernet ports via XAUI combo slots[4]	• Four onboard Gigabit Ethernet ports (10/100/1000) • Two 10 Gb Ethernet ports via XAUI combo slots[5]	• Four onboard Gigabit Ethernet ports (10/100/1000) • Two 10 Gb Ethernet ports via XAUI combo slots[5]
Power supplies	Two hot-swappable AC 720 W or DC 660 W power supply units (N+1 redundancy)	Two hot-swappable AC 1100 W or DC 1200 W power supply units (N+1 redundancy)	Four hot-swappable AC 1120 W power supply units (2+2 redundancy)

Fans	Six hot-swappable fan trays, with two fans per tray, N+1 redundancy	Five hot-swappable fan trays, with two fans per tray, N+1 redundancy	Four hot-swappable fan units, N+1 redundancy
Form factor	One rack unit (1RU)	Two rack units (2RU)	Four rack units (4RU)
Operating system	Oracle Solaris 10 8/07 and 10/08 OS or later	Oracle Solaris 10 8/07 and 10/08 OS or later	Solaris 10 8/07 and 10/08 OS or later

Table Notes:

256 GB maximum memory configuration requires the optional Memory Mezanine Kit.

[2] Using an XAUI adapter card converts one RJ-45 Gb Ethernet port into a 10 Gb Ethernet port. If two XAUI ports are used, only two Gb Ethernet ports are available.

[3] High line power input (200-240V) is required for eight-disk backplane and 1.6 GHz processor support. High-line power input is also required for the 16-disk backplane and either 1.4 GHz or 1.6 GHz processors.

[4] Using an XAUI adapter card converts one RJ-45 Gb Ethernet port into a 10 Gb Ethernet port. If two XAUI ports are used, only two Gb Ethernet ports are available.

[5] High line power input (200-240V) is required for eight-disk backplane and 1.6 GHz processor support. High-line power input is also required for the 16-disk backplane and either 1.4 GHz or 1.6 GHz processors.

Leading Reliability, Availability, and Serviceability

The Oracle Sun SPARC Enterprise T5440 server provides excellent RAS characteristics. Highly reliable parts and a relatively low total component count minimize the opportunity for system errors. Dual PCI-E root complexes and the ability to configure multiple processors add to resiliency. In addition, this server includes core and thread offlining capabilities, processor cache coherency, integrated disk RAID functions, and extensive ECC hardware protection—along with redundant hot-swap disks, power supplies, and fans. The following SPARC Enterprise T5440 server design elements are key to improving the dependability of IT services:

- Processor thread and core offlining and built-in RAID capabilities
- Redundancy and hot-swap components
- Parity protection and error-correction capabilities
- System monitoring
- ILOM service processor
- Superior energy efficiency
- Robust virtualization technology
- Comprehensive fault management

Innovative Support for Solid State Drives

Modern servers are driving throughput levels that can rapidly outpace the capabilities of traditional storage solutions. Although many servers can achieve processing capabilities in excess of 1 million I/O operations per second (IOPS), today's fastest hard disk drives (HDDs) are only capable of about 300 to 400 IOPS. To match throughput more closely to server performance, and to address the challenging demands of data-intensive applications, many data centers implement large pools of high-speed disk drives. In some cases, a large buffer of expensive DRAM is also deployed so that the application's working set can be stored in memory to reduce latency.

Flash technology provides a more-economical alternative that can dramatically enhance application I/O performance while also operating with significantly better energy efficiency than conventional HDDs. Recent advances in the quality of Flash technology have made solid-state drives (SSDs) an effective and reliable solution for enterprise storage. Flash technology contains no moving parts, avoiding the seek times and rotational latencies inherent with traditional HDD technology.

Because SSDs offer low latency and are significantly less expensive than DRAM storage, they balance cost and performance in a manner that can provide significant value for I/O-intensive workloads. SSDs offer a disk drive form factor (Figure 3), and are directly supported by the drive bays of the Oracle Sun SPARC Enterprise T5440 server.

Figure 3. SSDs provide enterprise Flash technology in a standard disk drive form factor.

Space, Watts, and Performance: The SWaP Metric

The Oracle Sun SPARC Enterprise T5120/T5220 and T5140/T5240 servers deliver leading performance across a range of multithreaded workloads and benchmarks. However, with energy and real estate costs and pressures, 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 are valuable as a way of comparing one system to another but limited when it comes to understanding the power and density attributes of the systems being compared.

For this reason, the SWaP—or space, watts, and performance—metric is useful. Designed to provide 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
- *Power* is measured by watts used by the system, taken during actual benchmark runs or from vendor's site planning guides

The UltraSPARC T2 Plus Processor with CoolThreads Technology

The UltraSPARC T2 and UltraSPARC T2 Plus processors are the industry's first systems-on-a-chip (SoC). As such, they supply the most cores and threads of any general-purpose processors, integrating all key system functions.

The World's First Massively Threaded Systems-on-a-Chip

The UltraSPARC T2 Plus processor eliminates the need for expensive custom hardware and software development by integrating computing, security, and I/O on a single chip. Binary compatible with earlier UltraSPARC processors, no other processor delivers so much performance in so little space and with such small power requirements. The result is that organizations are able to rapidly scale the delivery of new network services with maximum efficiency and predictability.

Figure 4. The UltraSPARC T2 Plus processor integrates computing, security, and I/O on a single chip.

UltraSPARC T1, UltraSPARC T2, and UltraSPARC T2 Plus Feature Comparison

Table 2 provides a comparison of the features of the UltraSPARC T1, UltraSPARC T2, and UltraSPARC T2 Plus processors.

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

FEATURE	ULTRASPARC T1 PROCESSOR	ULTRASPARC T2 PROCESSOR	ULTRASPARC T2 PLUS PROCESSOR
Cores per processor	Up to eight	Up to eight	Up to eight
Threads per core	Four	Eight	Eight
Threads per processor	32	64	64
Sockets supported	One	One	Two or four[1]
Hypervisor	Yes	Yes	Yes
Memory	Four memory controllers, 4 DIMMs per controller	Four memory controllers, up to 16 FB-DIMMs	Two memory controllers, up to 16 FB-DIMMs
Caches	16 KB instruction cache, 8 KB data cache, 3 MB L2 cache (four banks, 12-way associative)	16 KB instruction cache, 8 KB data cache, 4 MB L2 cache (eight banks, 16-way associative)	16 KB instruction cache, 8 KB data cache, 4 MB L2 cache (eight banks, 16-way associative)
Technology	Nine-layer Cu metal, CMOS process, 90 nm technology	65 nm technology	65 nm technology
Floating Point	1 FPU per chip	1 FPU per core, 8 FPUs per chip	1 FPU per core, 8 FPUs per chip
Integral resources	Single execution unit per core	Two integer execution units per core	Two integer execution units per core
Cryptography	Accelerated modular arithmetic operations (RSA)	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 Gb Ethernet interfaces, PCI Express interface (x8)	PCI Express interface (x8), Coherency logic and links (6.4 GB/second)

Table Notes:

An external coherency hub is used for four-socket implementations.

Taking Chip Multithreaded Design to the Next Level

Oracle's next-generation CMT processors were designed with the following goals in mind:

- Increase computational capabilities to meet the growing demand from Web applications by providing twice the throughput of the UltraSPARC T1 processor
- Support larger and more-diverse workloads with greater floating-point performance
- Power faster networking to serve new network-intensive content
- Provide end-to-end data center encryption
- Increase service levels and reduce downtime
- Improve data center capacity while reducing costs

Allowing different modular combinations of processors, cores, and integrated components, CMT architecture is extremely flexible. The above considerations drove an engineering effort that compared different approaches to improving on the successful UltraSPARC T1 architecture. For example, simply increasing the number of cores would have provided throughput gains but consumed extra die area, leaving no room for integrated components such as floating-point processors.

The final UltraSPARC T2 Plus processor design recognizes that memory latency is the true bottleneck to improved performance. Thus, by increasing the number of threads supported by each core and further increasing network bandwidth, these processors are able to provide approximately twice the throughput of the UltraSPARC T1 processor. Figure 5 provides a simplified high-level illustration of the thread model supported by an eight-core UltraSPARC T2 Plus processor.

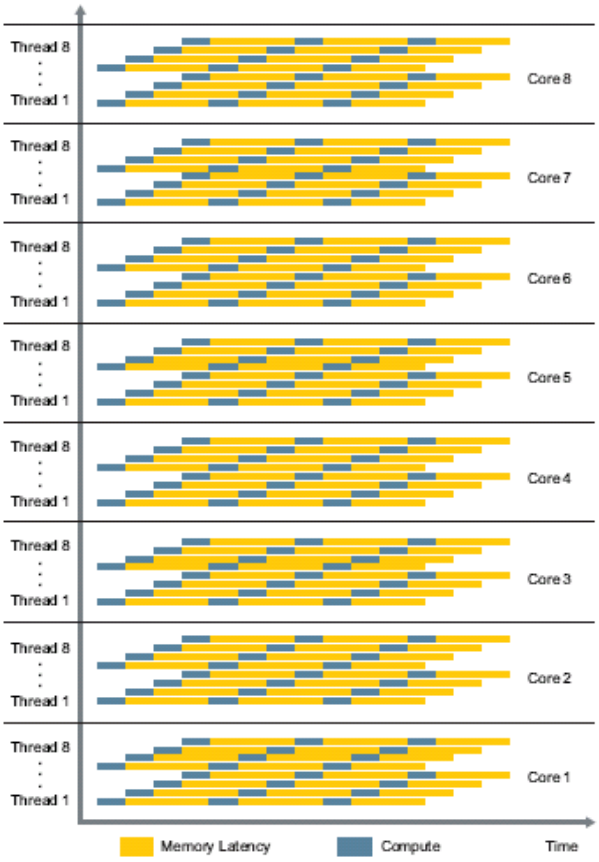

Figure 5. A single eight-core UltraSPARC T2 Plus processor supports up to 64 threads, with up to two threads running in each core simultaneously.

Each UltraSPARC T2 and UltraSPARC T2 Plus processor provides up to eight cores, with each core able to switch between up to eight threads (64 threads per processor). In addition, each core provides two integer execution units, so that a single UltraSPARC core is capable of executing two threads at a time.

UltraSPARC T2 Plus Processor Architecture

The UltraSPARC T2 Plus processor extends Oracle's CMT initiative with an elegant and robust architecture that delivers real performance to applications through multisocket implementations. A high-level block diagram of the UltraSPARC T2 Plus processor is shown in Figure 6.

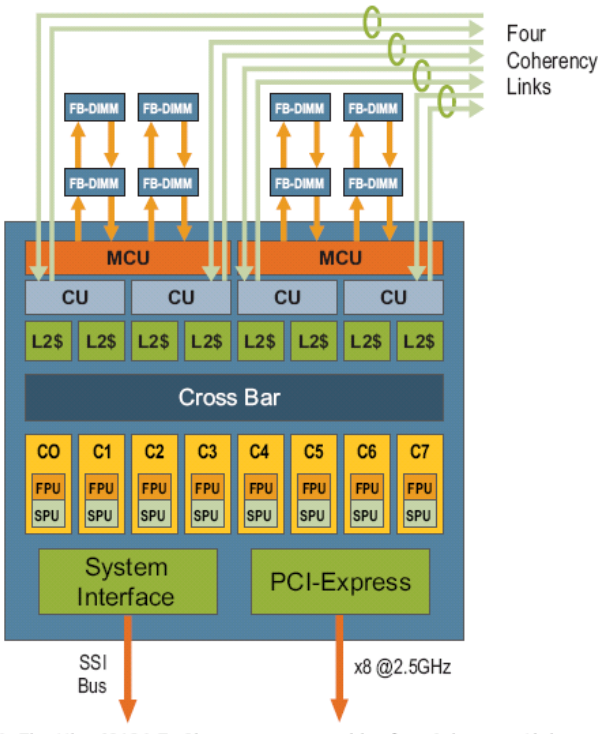

Figure 6. The UltraSPARC T2 Plus processor provides four coherence links to connect to up to four other processors.

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 approximately 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.

Each core provides its own fully pipelined floating point and graphics unit (FPU or FGU), as well as a stream-processing unit (SPU). The FGUs greatly enhance floating-point performance over that of the UltraSPARC T1, while the SPUs provide wire-speed cryptographic acceleration with more than 10 popular ciphers supported, including DES, 3DES, AES, RC4, SHA-1, SHA-256, MD5, RSA to 2048 key, ECC, and CRC32. Embedding hardware cryptographic acceleration for these ciphers allows end-to-end encryption, with no penalty in either performance or cost.

The UltraSPARC T2 Plus architecture also provides four coherence units to support multisocket system implementations. Four coherence channels (or coherence links) are provided—one associated with each coherence unit. These links run a cache coherence (snoopy) protocol over an FB-DIMM interface to provide up to 4.8 gigatransfers per port, providing 204 Gb/sec in each

direction. The memory link speed of the UltraSPARC T2 Plus processor was increased to 4.8 Gb/sec over the 4.0 Gb/sec of the UltraSPARC T2 processor.

The UltraSPARC T2 Plus processor can support both two- and four-socket implementations. Dual-socket UltraSPARC T2 Plus implementations—such as those provided by Oracle's Sun SPARC Enterprise T5140 and Sun SPARC Enterprise T5240 servers—interconnect the processors' four coherence links, and require no additional circuitry. Four-socket implementations require additional logic in the form of four external coherence hubs. A typical four-socket implementation is shown in Figure 7, with more information on the details of the external coherence hub provided later in this white paper.

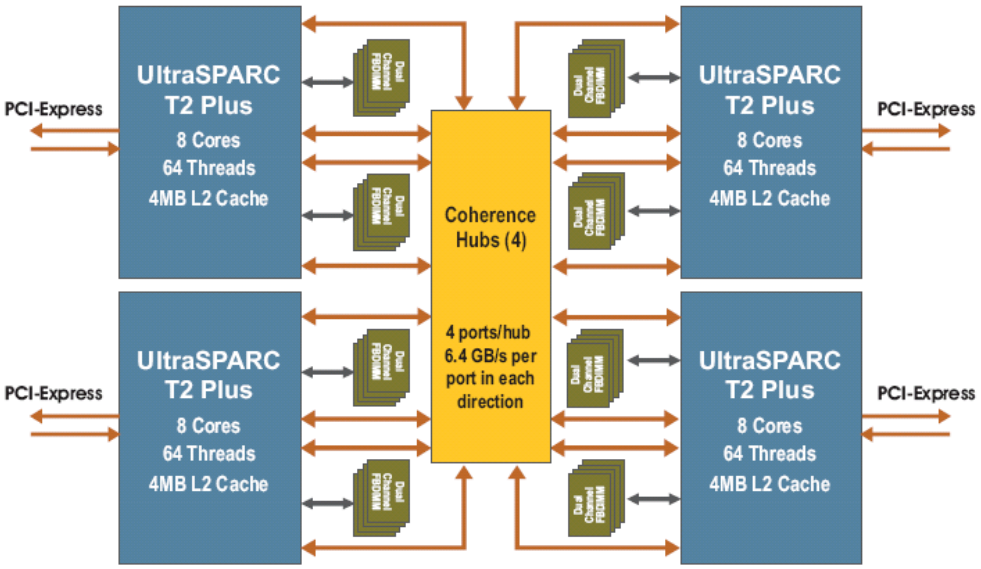

Figure 7. This illustrates the quad-socket UltraSPARC T2 Plus configuration in the Oracle Sun SPARC Enterprise T5440 server.

Core Architecture and Pipelines

Both the UltraSPARC T2 and UltraSPARC T2 Plus processors share the same core design. Figure 8 provides a block-level diagram representing a single UltraSPARC core on the UltraSPARC T2 processor (up to eight cores are supported per processor).

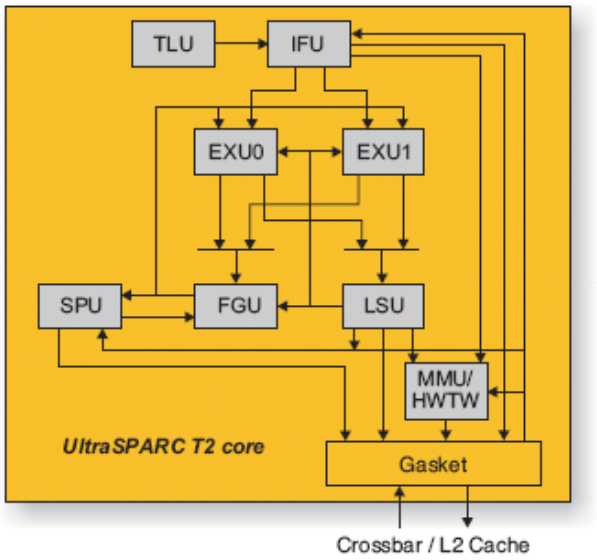

Figure 8. This block-level diagram represents a single UltraSPARC core on the UltraSPARC T2 processor.

Components implemented in each core include

- **Trap logic unit.** The trap logic unit (TLU) updates the machine state and handles exceptions and interrupts.
- **Instruction fetch unit.** The instruction fetch unit (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.** Dual integer execution units (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 (FGU) is provided within each core, and it is shared by all eight threads assigned to the core. Thirty-two floating-point register file entries are provided per thread.
- **Stream-processing unit.** Each core contains an SPU that provides cryptographic coprocessing.
- **Memory management unit.** The memory management unit (MMU) provides a hardware table walk (HWTW) and supports 8 KB, 64 KB, 4 MB, and 256 MB pages.

Each UltraSPARC T2 and UltraSPARC T2 Plus processor core provides an eight-stage integer pipeline and a 12-stage floating-point pipeline (Figure 9). A new "pick" pipeline stage has been added to choose two threads (out of the eight possible per core) to execute each cycle.

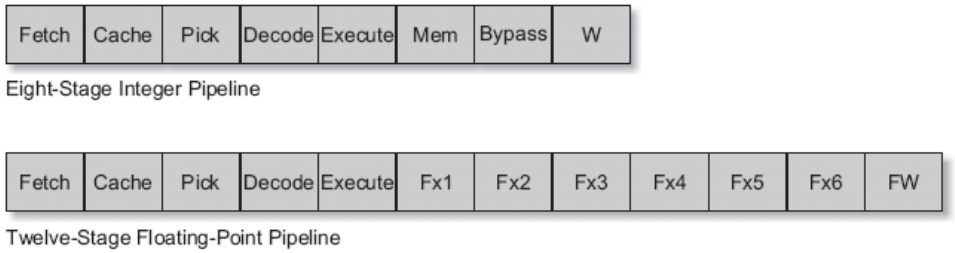

Figure 9. The UltraSPARC T2 Plus processor provides integer and floating-point pipelines.

To illustrate how the dual pipelines function, Figure 10 depicts the integer pipeline with the load store unit (LSU). The instruction cache is shared by all eight threads within the core, and 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.

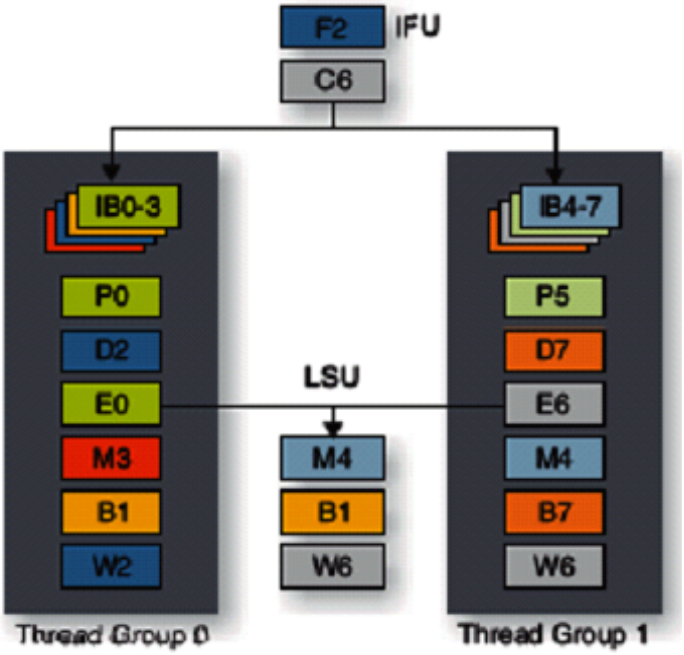

Figure 10. Threads are interleaved between pipeline stages with very few restrictions (integer pipeline shown; letters depict pipeline stages and numbers depict different scheduled threads).

The pick stage chooses one thread per cycle within each thread group. Picking within each thread group is independent of the other, 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 Figure 10, 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. Load/store and FPUs are shared among all eight threads. Only one thread from either thread group can be scheduled on such a shared unit.

Stream-Processing Unit

The SPU on each UltraSPARC T2 Plus processor core runs in parallel with the core at the same frequency. The SPU is designed to achieve wire-speed encryption and decryption on both of the processor's 10 Gigabit Ethernet ports.

Integral PCI Express Support

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

Power Management

Beyond the inherent efficiencies of CMT design, the UltraSPARC T2 and UltraSPARC T2 Plus processors are the first processors to incorporate unique power management features at both the core and memory levels of the processor. These features include reduced instruction rates, parking of idle threads and cores, and the ability to turn off clocks in both cores and memory to reduce power consumption. Substantial innovation is present in the areas of

- 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

Server Architecture

The Oracle Sun SPARC Enterprise T5440 server is designed to provide breakthrough performance and scalability while building on the capabilities of the Oracle Sun SPARC Enterprise T5140 and Oracle Sun SPARC Enterprise T5240 servers. Oracle's Sun SPARC Enterprise T5440 servers expand on the strengths of the UltraSPARC T2 Plus processor while providing innovative chassis design features. The result is a highly scalable UltraSPARC T2 Plus-based server with considerable CPU, memory, and I/O expansion capabilities.

System-Level Architecture

The focus of the design effort for the Oracle Sun SPARC Enterprise T5440 server was on providing cache coherency for up to four UltraSPARC T2 Plus processors with minimal latencies. The proven chassis design was chosen to provide computational density in a predictable thermal and power envelope. See Figure 11 for a high-level block diagram of the Oracle Sun SPARC Enterprise server motherboard.

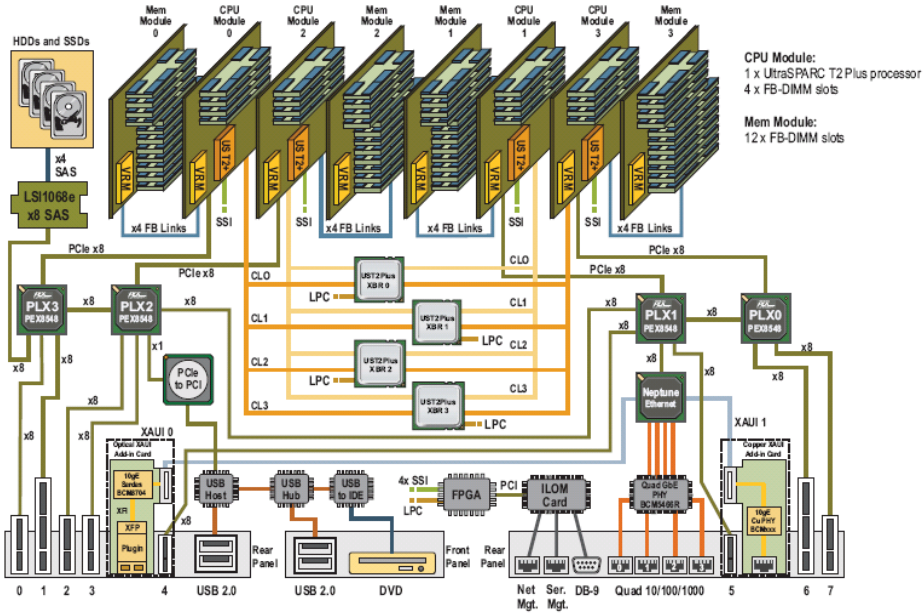

Figure 11. The Oracle Sun SPARC Enterprise T5440 server's motherboard supports all system capabilities.

The Oracle Sun SPARC Enterprise T5440 motherboard is a printed circuit board that supports all of the system capabilities, including processors and memory (on CPU and memory modules, respectively), disk controller, and I/O subsystems. An ILOM service processor is provided on a daughter card that attaches to the motherboard. I/O options include USB, DVD control, quad Gigabit Ethernet, and four independent PCI-E buses providing sockets for a wide variety of third-party PCI-E expansion.

Key features of the Oracle Sun SPARC Enterprise T5440 motherboard include

- **Four sockets for CPU modules.** Each module contains an UltraSPARC T2 Plus processor and four FB-DIMM sockets. Each CPU module contains the minimum memory configuration to support the UltraSPARC T2 Plus processor.
- **Four sockets for memory modules.** Each memory module provides expansion memory for the corresponding CPU module through use of up to 12 FB-DIMM sockets. Each memory module connects to a CPU module through a physical FB-DIMM interface via the motherboard.

- **Four external coherence hub ASICs.** Each external coherence hub is connected to the coherence links on each UltraSPARC T2 Plus processor.
- **Four PCI Express expansion hubs.** Each hub is connected to the PCI-E interface on each UltraSPARC T2 Plus processor through the CPU module socket. The PCI-E expansion hubs support a variety of peripheral system devices as well as eight PCI-E slots at the rear of the chassis, and are interconnected to help ensure connectivity from any CPU module to any peripheral device.
- **Oracle's Neptune chip.** The Neptune chip facilitates 10 Gigabit Ethernet functionality as well as four standard Gigabit Ethernet ports (10/100/1000-BaseT).

The motherboard interconnect for the Oracle Sun SPARC Enterprise T5440 server is greatly simplified compared with previous-generation systems. A pair of metal bus bars connected to a power distribution board (PDB) distributes 12-volt power to the motherboard, and a single flex-circuit connector routes all critical power control, DVD drive, and SAS disk signaling over to the PDB and on from the motherboard to the disk drive backplane, providing data access to the system hard drives.

External Coherence Hub (UltraSPARC T2 Plus Crossbar)

Dual-socket systems such as Oracle's Sun SPARC Enterprise T5140 and Oracle's Sun SPARC Enterprise T5240 servers can be designed to connect directly to the coherence links of two UltraSPARC T2 Plus processors. For larger systems such as the Oracle Sun SPARC Enterprise T5440 server, an external coherence hub is required. The external coherence hub extends the cache hierarchy of a single UltraSPARC T2 Plus processor across up to four sockets and processors (nodes).

The external coherence hub is a four-port arbiter/switch implemented in a custom ASIC that interfaces to the coherency control portion of the UltraSPARC T2 Plus processor. (See Figure 12 for a high-level block diagram of the external coherence hub.) The external coherence hub performs a number of functions, including

- Serializing requests to the same address for loads, stores, and write-backs
- Broadcasting snoops and aggregating responses
- Providing flow-control read and write requests to a processor
- Providing ECC or parity protection

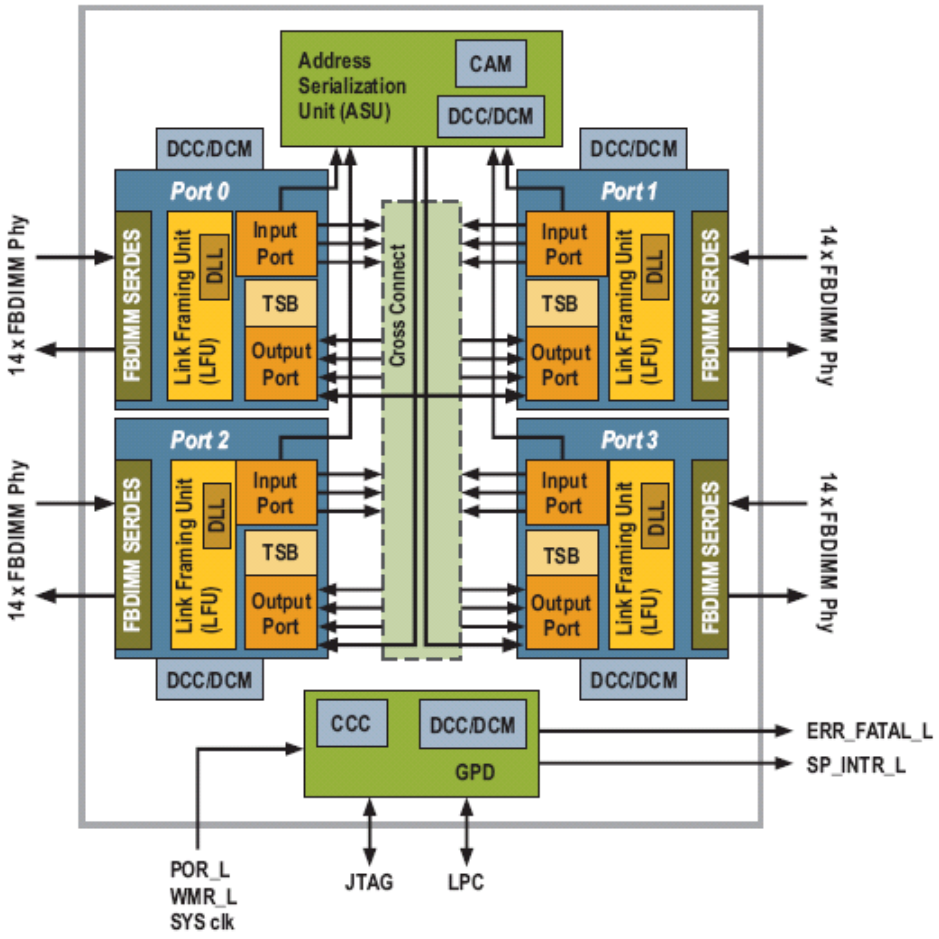

Figure 12. Pictured in this block-level diagram, the external coherence hub is a four-port arbiter/switch implemented in a custom ASIC that interfaces to the coherency control portion of the UltraSPARC T2 Plus processor.

When memory requests fail to hit a processor's L2 cache, they are forwarded to the external coherence hub. The hub fulfills these requests through a series of messages to multiple nodes in a system. The hub also functions as the single point of serialization for addresses within a coherency plane, and it enforces ordering rules necessary to comply with the Total Store Ordering (TSO) consistency model. The external coherence hub is also responsible for forwarding a node's noncacheable space and locations mapped to I/O devices on the PCI-E I/O fabric attached to the destination node.

Each Oracle Sun SPARC Enterprise T5440 server provides four external coherence hubs, each of which is connected to all four CPU module sockets. Components of the external coherence hub include

- **FB-DIMM interface.** The FB-DIMM interface block contains the SerDes cores and symbol alignment logic, and is virtually identical to that used for an FB-DIMM northbound interface.

- **Link framing unit.** The link framing unit (LFU) on each external coherence hub implements a protocol on top of the FB-DIMM frame/cell link layer that includes packetization, CRC checking, and retransmission of packets that had errors detected at the receiver.
- **I/O ports.** The output port receives forwarded messages from all input ports while processing some messages. A Transaction Score Board (TSB) tracks the state of requests, replies, and data, respectively, for coherent read- and write-back.
- **Address serialization unit.** The address serialization unit (ASU) is responsible for serializing requests to the same memory address, and for maintaining the list of requests pending on a particular address. The ASU handles requests from all four ports.
- **Cross-connect.** The cross-connect consists of the set of interconnect wires that create the various port-to-port networks.

Memory Subsystem

In the Oracle Sun SPARC Enterprise T5440 server, the UltraSPARC T2 Plus processor provides on-chip memory controllers that communicate directly with FB-DIMM memory through high-speed serial links. Two dual-channel FB-DIMM memory controller units (MCUs) are provided on each UltraSPARC T2 Plus processor in addition to the four coherence units. Each MCU on the UltraSPARC T2 Plus processor can transfer data at an aggregate rate of 4.8 Gb/sec.

In addition to the UltraSPARC T2 Plus processor, each CPU module provides four populated FB-DIMM slots. These four slots make up the primary bank of FB-DIMMs and are the minimum required to support a single processor. The CPU module is always accompanied by its corresponding memory module, which is connected through an FB-DIMM interface via the motherboard connectors. By default, memory modules are populated with four FB-DIMMs, allowing sufficient DIMM space to double memory capacity for each processor. Memory modules with no memory can be installed, if desired, and are provided with DIMM fillers to maintain correct airflow.

Each memory module provides an additional 12 FB-DIMM sockets, yielding a total of 16 memory socket locations per processor. Note that the FB-DIMMs on each of the memory modules are physically connected to only one CPU module and one processor. With up to 16 667 MHz FB-DIMMs associated with each processor, the maximum system memory capacity is 128 GB using 2 GB FB-DIMMs, 256 GB using 4 GB FB-DIMMs, and 512 GB using 8 GB FB-DIMMs.

I/O Subsystem

Each UltraSPARC T2 Plus processor incorporates a single eight-lane (x8) PCI-E port capable of operating at 4 GB/sec bidirectionally. In the Oracle Sun SPARC Enterprise T5440 server, this port provides a native interface to the I/O devices through four PLX-technology PCI-E expander chips (bridge chips). All of the PCI-E expander chips are themselves interconnected

with x8 PCI-E interfaces, helping ensure that any processor can access any I/O device. The PCI-E expander chips connect to either PCI-E card slots or to bridge devices that interface with PCI-E, such as those listed below:

- **Disk controller.** Disk control is managed by an LSI Logic SAS1068E SAS controller chip that interfaces to a four-lane (x4) PCI-E port. RAID levels 0 and 1 are provided as standard.
- **Modular disk backplanes.** A four-disk backplane is attached to the LSI disk controller by an x4 SAS link.
- **Solid-state drives.** Solid-state drives (SSDs) can be substituted for disk drives in the Oracle Sun SPARC Enterprise T5440 server. Up to four 32 GB SSDs can be installed, with remaining slots occupied with conventional SAS HDDs.
- **Quad Gigabit Ethernet.** On the Oracle Sun SPARC Enterprise T5440 server, the Neptune Ethernet chip provides two 10/100/1000-BaseT ports and two 10/100/1000/10000-BaseT interfaces, exposed as four RJ-45 connectors on the rear panel.
- **Dual 10 Gigabit Ethernet.** The Oracle Sun SPARC Enterprise T5440 server provides dual 10 Gb XAUI connections, expressed through shared XAUI/PCI-E slots. These 10 Gigabit Ethernet interfaces are provided by the Neptune Ethernet chip. When the 10 Gigabit Ethernet ports are connected, the corresponding number of Gigabit Ethernet ports becomes unavailable for use. The same XAUI expansion cards are supported on Oracle's Sun SPARC Enterprise T5120, Oracle's Sun SPARC Enterprise T5220, Oracle's Sun SPARC Enterprise T5140, Oracle's Sun SPARC Enterprise T5240, and Oracle's Sun SPARC Enterprise T5440 servers.
- **USB and DVD.** A single-lane PCI-E port connects to a PCI bridge device. A second bridge chip converts the 32-bit 33 MHz PCI bus into multiple USB 2.0 ports. The system's USB interconnect is driven from those ports. In addition, the DVD drive is driven from a further bridge chip that interfaces one of the USB ports to IDE format.

Chassis Design Innovations

Oracle's Sun SPARC Enterprise T5120/T5220, Oracle's Sun SPARC Enterprise T5140/T5240, and Oracle's Sun SPARC Enterprise T5440 servers all share basic chassis design innovations. This approach not only provides a consistent look and feel across the product line, but it simplifies administration through consistent component placement and shared components. Beyond mere consistency, this approach provides a data center design focus that places key technology where it can make a difference for the data center.

Enhanced System and Component Serviceability

Finding and identifying servers and components in a modern data center can be challenging. The Oracle Sun SPARC Enterprise T5440 server is 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. Fault Remind features identify failed components.

Consistent connector layouts for power, networking, and management make moving between Oracle's x64 and SPARC systems straightforward. All hot-plug components are tool-less and easily available for serviceability. For instance, an integral hinged lid provides access to dual fan modules so that fans can be serviced without exposing sensitive components or causing unnecessary downtime.

Robust Chassis, Component, and Subassembly Design

Sun volume servers share chassis that are carefully designed to provide reliability and cool operation. Even features such as the hexagonal chassis ventilation holes are designed to provide the best compromise for high strength, maximum airflow, and electronic noise.

Fan assemblies direct airflow to efficiently cool the system. Fan modules are isolated from chassis to avoid transfer of rotational vibration to other system components. In addition, fans are monitored for vibration and actively controlled to avoid the effects of sympathetic vibrations being passed on to disk drives. Active monitoring and control of fan speeds is employed to reduce aural noise and minimize fan power draw from the system while still maintaining sufficient system cooling.

Despite their computational, I/O, and storage density, Oracle's Sun servers are able to maintain adequate temperatures using conventional technologies. Minimized DC-to-DC power conversions also contribute to overall system efficiency. By providing 12-volt power to the motherboard, multiple power conversion stages can be eliminated. This approach reduces generated heat and introduces further efficiencies to the system.

Minimized Cabling for Maximized Airflow

To minimize cabling and increase reliability, a variety of smaller boards and connectors are employed, appropriate to each chassis. These infrastructure boards serve various functions in the Oracle Sun SPARC Enterprise T5440 server.

- A power distribution board distributes system power from the dual power supplies to the motherboard and to the disk backplane (via a connector board).
- Connector boards eliminate the need for many discrete cables, providing a direct card plug-in interconnect to distribute control and most data signals to the disk backplane, fan boards, and the PDB.
- PCI-E cards plug directly into the motherboard.

- Two XAUI slots provide access to the 10 Gigabit Ethernet interfaces on the Neptune Ethernet chip. Alternatively, these slots can provide access to PCI-E interfaces. Each slot can either accept an optical/copper XAUI card or an industry-standard low-profile PCI-E card with up to an x8 form factor edge connector.
- The disk backplane mounts to the disk cages in the chassis, delivering disk data through a four-channel discrete miniSAS cable from the motherboard. A four-disk backplane is offered for the Oracle Sun SPARC Enterprise T5440 server.
- Also provided by the disk backplane, a DVD-RW and two USB connections route to the front of the system.

Oracle Sun SPARC Enterprise T5440 Server Overview

The Oracle Sun SPARC Enterprise T5440 server provides breakthrough computational power and scalability in a space-efficient 4RU rackmount package. By extending the proven benefits of Oracle's CMT architecture to the most mission-critical and OLTP database workloads supporting Web services, these systems redefine midrange computing. The server is also designed to address the challenges of modern data centers with greatly reduced power consumption and a small physical footprint. Depending on the model selected, the Oracle Sun SPARC Enterprise T5440 server features one to four eight-core UltraSPARC T2 Plus processors and up to 512 GB of memory.

System Motherboard and Chassis Perspective

Figure 13 provides a top-down perspective of the Oracle Sun SPARC Enterprise T5440 chassis with the top cover removed and a full complement of CPU modules and memory modules installed. As discussed, each CPU module is paired with a corresponding memory module; however, CPU modules are not required to have their corresponding memory module installed, nor are all CPU modules required for the system to operate. If any CPU or memory module is not installed, a filler module is installed to help ensure proper airflow and front-to-back cooling.

All eight PCI-E slots are low profile and are wired to x8 PCI-E interfaces. Two slots provide x16 physical connectors to support x16 PCI-E cards such as low-profile graphics accelerators. Two PCI-E slots share the rear opening with optional XAUI cards that plug in behind the PCI-E connectors. When an XAUI card is installed, no PCI-E card can be used in the corresponding PCI-E slot. Also, when an XAUI card is installed, one of the Gigabit Ethernet ports attached to the Neptune chip becomes unavailable.

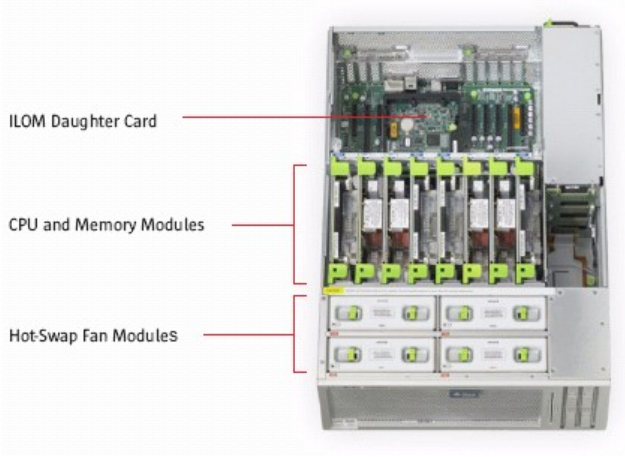

Figure 13. This top-down view of the Oracle Sun SPARC Enterprise T5440 server shows the server's CPU and memory modules.

Four system fans insert from the top of the chassis, and four power supplies insert from the rear of the chassis. The four power supplies provide 2+2 redundancy. The system can continue to operate at full capacity with any combination of two of the four power supplies, thereby covering a failure of an individual power supply or an entire circuit within the data center.

Enclosure

The 4RU Oracle Sun SPARC Enterprise T5440 server enclosure is designed for use in a standard 19-inch rack (Table 3).

TABLE 3. DIMENSIONS AND WEIGHT OF THE ORACLE SUN SPARC ENTERPRISE T5540 SERVER

DIMENSION	UNITED STATES	INTERNATIONAL
Height	6.92 inches (4 RU)	176 millimeters
Width	17.5 inches	445 millimeters
Depth	14.9 inches	633 millimeters
Weight (approximate maximum, without PCI cards, rackmounts)	88 pounds	40 kilograms

The Oracle Sun SPARC Enterprise T5440 server includes the following major components:

- One to four UltraSPARC T2 Plus processors with eight cores at speeds of 1.2 GHz, 1.4 GHz, or 1.6 GHz

- Up to 512 GB of memory in 16 FB-DIMM slots per processor (2 GB, 4 GB, and 8 GB FB-DIMMs supported)
- Four onboard 10/100/1000 Mb/sec Ethernet ports
- Six dedicated low-profile PCI-E slots (x8 electrically, two x 16 physical slots)
- Two combination XAUI or low-profile PCI-E x8 slots
- Four USB 2.0 ports (two forward, two rear facing)
- Four available disk drive slots supporting either SAS disk drives or SSDs
- ILOM system controller
- Four hot-swappable, high-efficiency 1,120-watt power supply units, providing 2+2 redundancy when distributed over two independent power circuits
- Four (N+1) cooling fans under environmental monitoring and control, accessed directly from the top of the chassis

Front and Rear Perspectives

Figure 14 illustrates the front and rear panels of the Oracle Sun SPARC Enterprise T5440 server.

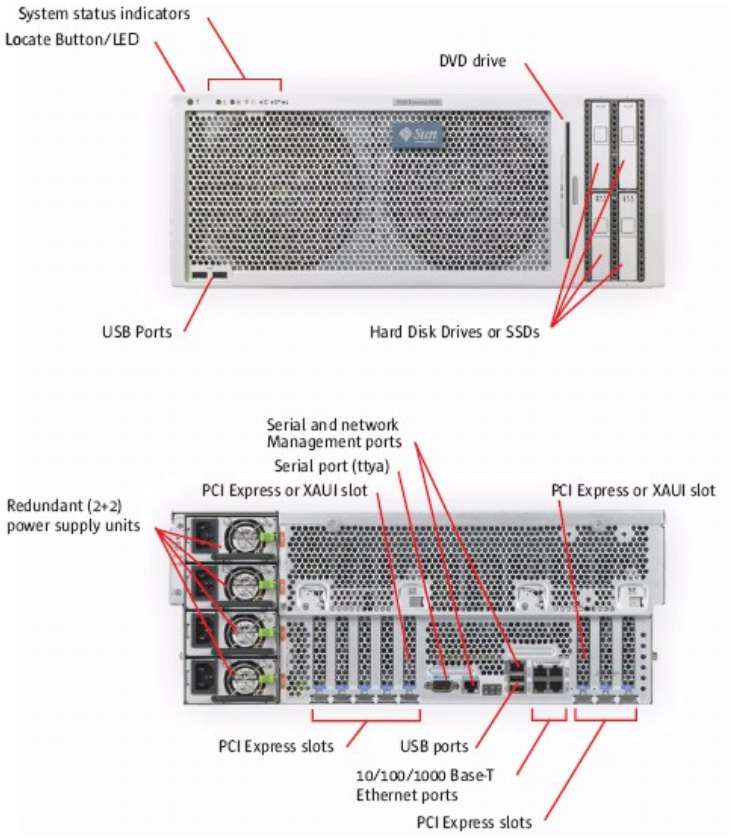

Figure 14. This provides front- and rear-panel perspectives of the Oracle Sun SPARC Enterprise T5440 server.

External features of the Oracle Sun SPARC Enterprise T5440 server include

- Front and rear system- and component-status indicator lights that provide locator (white), service required (amber), and activity status (green) for the system
- Four hot-plug SAS disk drives that insert through the front panel of the system
- One slim-line, slot-accessible DVD-RW accessible through the front panel
- Four USB 2.0 ports—two on the front panel and two on the rear
- Four hot-plug/hot-swap (2+2) power supplies with integral fans that can be inserted from the rear
- Rear power-supply indicator lights that convey the status of each power supply
- A single AC plug on each hot-plug/hot-swap power supply
- Four 10/100/1000Base-T autosensing Ethernet ports

- A DB-9 TTYA serial port for serial devices (not connected to the ILOM system controller serial port)
- A total of eight PCI-E card slots, two of which can alternatively support XAUI cards connected to the Neptune 10 Gigabit Ethernet chip
- Two management ports for use with the ILOM system controller; RJ-45 serial management port provides default connection to the ILOM controller

PCI Express Expansion Unit

With four sockets for UltraSPARC T2 Plus processors, Oracle Sun SPARC Enterprise T5440 servers are ideally suited for mission-critical applications and databases, applications that often require fast access to considerable near-line storage. With each of up to four UltraSPARC T2 Plus processors providing a PCI-E root complex, the system also has considerable I/O capacity and bandwidth available for external expansion. To scale I/O expansion beyond the constraints of the 4RU system chassis, Oracle Sun SPARC Enterprise T5440 servers support the attachment of up to two optional Sun External I/O Expansion Units to provide additional I/O connectivity. As a result, a maximally configured Oracle Sun SPARC Enterprise T5440 server can provide up to 28 PCI-E slots.[1]

The Sun External I/O Expansion Unit is a 4RU rackmountable device that accommodates up to two 12 additional PCI-E slots—connected to, and managed by, the Oracle Sun SPARC Enterprise T5440 server. By using cassettes, the external I/O chassis supports active replacement of hot-plug cards. An I/O link card mounted in the host provides connectivity to the Sun External I/O Expansion Unit and enables host management control via sideband signals. Available as a low-height copper card, the I/O link card includes a single eight-lane PCI-E bus with 4 GB/sec bandwidth. The architecture of the Sun External I/O Expansion Unit provides high-throughput I/O performance, supporting maximum data rates for many types of PCI-E cards and bursty traffic from additional PCI-E cards. Front and rear perspectives of the Sun External I/O Expansion Unit are shown in Figure 15.

[1] One I/O boat per installed CPU module represents the maximum configuration.

Figure 15. This provides front and rear perspectives of the Sun External I/O Expansion Unit with one I/O boat installed.

Each Sun External I/O Expansion Unit contains either one or two "I/O boats," with each boat providing six external x8 PCI-E slots. Individual I/O boats connect to the Oracle Sun SPARC Enterprise T5440 server via a link card that is installed in one of the system's PCI-E slots. Each installed I/O boat requires an installed CPU module and one I/O boat per installed CPU module is the maximum configuration. The Sun External I/O Expansion Unit includes several key technologies, including

- Link card side-band communication technology along with I/O manager capabilities built into the system software, allowing seamless remote management and integration with the host server
- Redundancy and hot-plug capabilities for power supply units, fans, I/O boats, and I/O cards
- Thermal monitoring and remote diagnostic capabilities

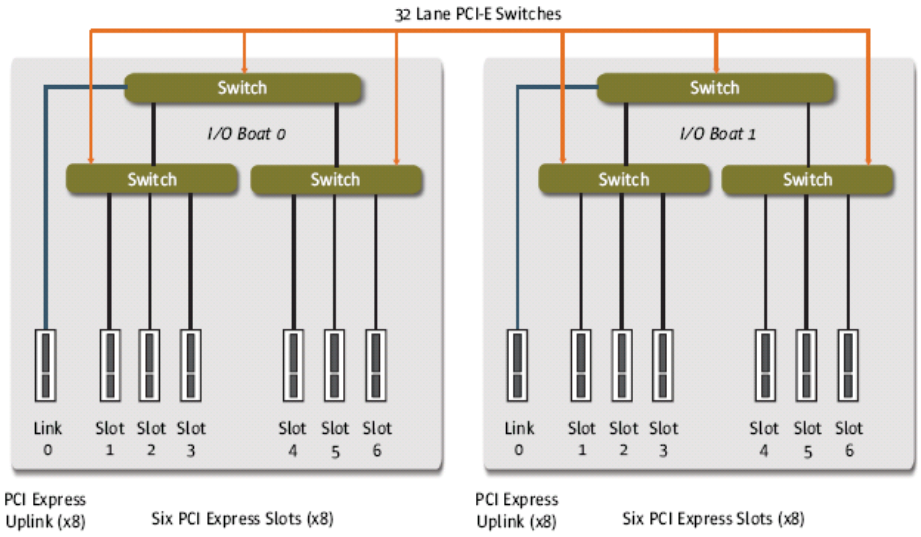

Figure 16. The Sun External I/O Expansion Unit's architecture is revealed in this block-level diagram.

The Sun External I/O Expansion Unit consists of an advanced network of switches, bridges, and controllers to allow I/O data communication along with system control information via the PCI-E link card connection to the host server. The x8 PCI-E connection from the host is forwarded over the link card cable to a switch in each I/O boat. Two additional switches then connect to the six x8 PCI-E slots in the I/O boat. The switches support the side-band management function of the link card by providing the following functionality:

- Gathering and communicating I2C diagnostic and environmental information to the host server's ILOM service processor
- Executing system instructions to modify fan speeds or selectively power down components within the unit
- Performing locate/warning information via the front-panel LEDs
- Providing support for the online maintenance and replacement of power supply units, I/O boats, or I/O cards during system operation; also providing support for the addition and deletion of active I/O cards

Enterprise-Class Management and Software

While new technology often requires time for tools and applications to arrive, delivering agile and highly available services that take advantage of available resources requires stable development tools, operating systems, middleware, and management software. Fortunately, in spite of the breakthrough UltraSPARC T2 Plus processor technology, Oracle Sun SPARC Enterprise T5440 servers provide full binary compatibility with earlier SPARC systems and are delivered ready to

run with preloaded tools and the solid foundation of the Solaris OS. Moreover, these systems are provided with a wealth of sophisticated tools that let organizations develop and tune applications as they consolidate and manage workloads, while effectively using the resources of UltraSPARC T2 Plus processors.

System Management Technology

As the number of systems in any organization grows, 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 key administrative tasks.

Integrated Lights Out Manager System Controller

The Oracle Integrated Lights Out Manager service processor acts as a system controller, facilitating remote management and administration of Oracle Sun SPARC Enterprise T5440 servers. The full-featured service processor is similar in implementation to that used in other Oracle modular and rackmount servers. As a result, these servers integrate easily with existing management infrastructure. In the Oracle Sun SPARC Enterprise T5440 server, the service processor is provided on a daughter card that connects to the system motherboard.

Critical to effective system management, the ILOM service processor

- Implements an IPMI 2.0-compliant services processor, providing IPMI management functions to the server's firmware, OS, and applications, and to IPMI-based management tools accessing the service processor via the ILOM Ethernet management interface, providing visibility to the environmental sensors (both on the server module and elsewhere in the chassis)
- Manages inventory and environmental controls for the server (including CPUs, DIMMs, and power supplies), and provides HTTPS/CLI/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 also allows the administrator to remotely manage the server, independent of the operating system running on the platform and without interfering with any system activity. Oracle Integrated Lights Out Manager can also send e-mail alerts of hardware failures and warnings, as well as other events related to each server. The ILOM circuitry runs independently from the server, using the server's standby power. As a result, ILOM firmware and software continue to function when the server operating system goes offline as well as when the server is powered off. ILOM monitors the following Oracle Sun SPARC Enterprise T5440 server conditions:

- CPU temperature

- Hard drive presence
- Enclosure thermal conditions
- Fan speed and status
- Power supply status
- Voltage
- Oracle Solaris watchdog, boot time-outs, and automatic server restart events

Oracle Enterprise Manager Ops Center

Beyond affording local and remote management capabilities, data center infrastructures need to be agile and flexible, allowing not only fast deployment but also streamlined redeployment of resources as required. Oracle Enterprise Manager Ops Center technology provides an IT infrastructure management platform for integrating and automating the management of thousands of heterogeneous systems. To improve lifecycle and change management, Oracle Enterprise Manager Ops Center supports the management of applications and the servers on which they run, including Oracle Sun SPARC Enterprise T5440 servers.

Oracle Enterprise Manager Ops Center takes a step-by-step approach to unraveling the challenges of getting systems operational quickly, providing functionality that includes

- **Discovery.** 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.
- **Grouping.** Given the number of systems IT organizations must manage and constantly repurpose, it's critical to find ways to group such resources. Oracle Enterprise Manager Ops Center provides such a means, letting users logically group systems and then perform actions across groups as easily as they would on a single system. Systems can be grouped by function (for example, Web servers versus cluster servers), administrative responsibility, or other categorization based on organizational needs.
- **Provisioning.** Oracle Enterprise Manager Ops Center remotely installs operating systems onto select systems. Administrators can use this functionality to provision operating systems 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 select systems.
- **Monitoring.** When systems are up and running, administrators can use Oracle Enterprise Manager Ops Center to monitor system health, helping to 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 threshold levels and set preferred notification methods (including e-mail, pager, or Simple Network Management Protocol traps) for each monitored component, as business needs demand.

- **Management.** Businesses require that infrastructure lifecycle management extend beyond just deploying and monitoring systems. 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 their IT infrastructure from remote locations. 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 the GUI and CLI 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

The Oracle Solaris 10 is designed to deliver the considerable resources of UltraSPARC T2 Plus processor-based systems, delivering functionality that facilitates virtualization, optimal use, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. Oracle Solaris 10 runs on a broad range of SPARC and x86-based systems.

One of the most attractive features of systems based on UltraSPARC T2 Plus processors is that they appear as familiar symmetric multiprocessing (SMP) systems to Oracle Solaris and the applications it supports. In addition, Oracle Solaris 10 provides a variety of features to improve application performance on CMT architectures, including the following:

- **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 physical processor as 64 logical processors, Oracle Solaris 10 understands the correlation between cores and the threads they support, and provides a fast and efficient thread implementation.
- **Fine-granularity manageability.** For the UltraSPARC T2 Plus processor, Oracle Solaris 10 has the ability to enable or disable 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.
- **Binding interfaces.** Oracle Solaris allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set, if required or desired.
- **Support for virtualized networking and 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 PCI-E interface and

cryptographic processors. As part of a high-performance network architecture, CMT-aware device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices. Accelerated cryptography is supported through the cryptographic framework in Oracle Solaris.

- **Non-Uniform Memory Access optimization in Oracle Solaris.** With memory managed by each UltraSPARC T2 Plus processor on Oracle Sun SPARC Enterprise T5440 servers, the implementation represents a Non-Uniform Memory Access (NUMA) architecture. In NUMA architectures, the speed needed for a processor to access its own memory is slightly different than that required to access memory managed by another processor. Oracle Solaris 10 provides technology that helps applications improve performance on NUMA architectures.
  - Memory placement optimization—Oracle Solaris 10 uses memory placement optimization (MPO) to improve the memory placement across the physical memory of a server, resulting in increased performance. Through MPO, Oracle Solaris 10 helps ensure that memory is as close as possible to the processors that access it, while still maintaining a sufficient balance within the system. As a result, many database and high-performance computing applications are able to run considerably faster with MPO.
  - Hierarchical lgroup support—Hierarchical lgroup support (HLS) improves the MPO feature in the Solaris OS. HLS helps Oracle Solaris optimize performance for systems with more complex memory latency hierarchies. HLS lets Oracle Solaris distinguish between the degrees of memory remoteness, allocating resources with the lowest-possible latency for applications. If local resources are not available by default for a given application, HLS helps Oracle Solaris allocate the nearest remote resources.
- **Oracle Solaris ZFS.** Oracle Solaris ZFS offers a dramatic advance in data management, automating and consolidating complicated storage administration concepts and providing unlimited scalability with the world's first 128-bit file system. Oracle Solaris ZFS is based on a transactional object model that removes most of the traditional constraints of I/O issue order, resulting in dramatic performance gains. Oracle Solaris ZFS also provides data integrity, protecting all data with 64-bit checksums that detect and correct silent data corruption.
- **A secure and robust enterprise-class environment.** Best of all, Oracle Solaris doesn't require arbitrary sacrifices. Certified multilevel security protects Oracle Solaris environments from intrusion. Oracle's comprehensive fault management architecture means that elements such as Oracle Solaris predictive self-healing can communicate directly with the hardware to help reduce both planned and unplanned downtime. Effective tools such as DTrace help organizations tune their applications to get the most from their systems' resources.

End-to-End Virtualization Technology

As organizations struggle to consolidate disparate workloads onto fewer, more-powerful systems while increasing use, virtualization technology is becoming increasingly popular. The Oracle Sun

SPARC Enterprise T5440 server is designed for virtualization, providing a very fine-grained division of multiple resources—from processing to virtualized networking and I/O. Most important, Oracle's virtualization technology is provided as part of the system, not as an expensive add-on.

A Multithreaded Hypervisor

Like the UltraSPARC T1 processor and UltraSPARC T2 processors, the UltraSPARC T2 Plus processor offers a multithreaded hypervisor—a small firmware layer that provides a stable virtual-machine architecture that's tightly integrated with the processor. Multithreading is crucial, because the hypervisor interacts directly with the underlying CMT processor. This architecture is able to context-switch between multiple threads in a single core, a task that requires additional software and considerable overhead in competing architectures.

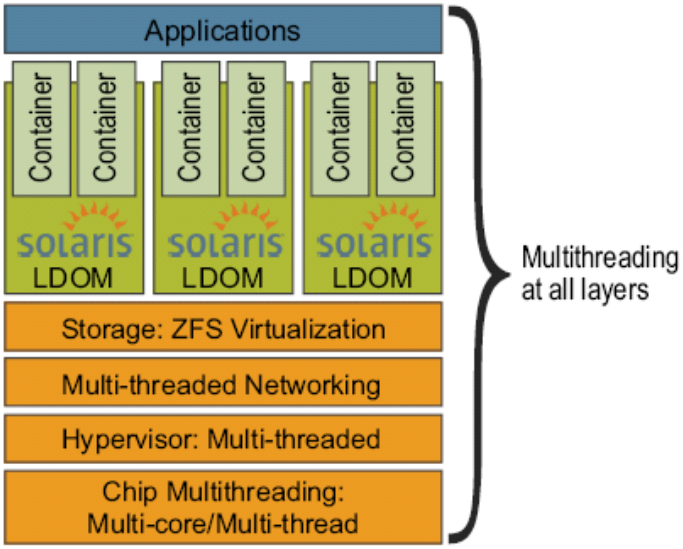

Figure 17. Oracle provides parallelization and virtualization at every level of the technology stack.

As shown in Figure 17, corresponding layers of virtualization technology are built on top of the hypervisor. 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. This considerable experience has informed technology decisions at other levels, 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 fully multithreaded Oracle Solaris ZFS file system. With multithreading, Oracle VM Server for SPARC, Oracle Solaris Containers, and applications are able to receive exactly the resources they need.

Domains

Supported in all Oracle servers that use CMT technology, Oracle VM Server for SPARC provides full virtual machines that run an independent operating system instance, and contain virtualized CPU, memory, I/O, storage, console, and cryptographic devices. Within the Oracle VM Server for SPARC architecture, operating systems such as Oracle Solaris 10 are written to the hypervisor, which provides a stable, idealized, and virtualizable representation of the underlying server hardware to the operating system in each logical domain. Each domain is completely isolated, and the maximum number of virtual machines created on a single platform relies on the capabilities of the hypervisor rather than the number of physical hardware devices installed in the system. For example, the Oracle Sun SPARC Enterprise T5440 server supports up to 128 domains, and each individual domain can run a unique OS instance.[2]

By taking advantage of domains, organizations are able to deploy multiple operating systems 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. Oracle Solaris VM Server for SPARC can also host Oracle Solaris Containers to capture the isolation, flexibility, and manageability features of both technologies. Deeply integrating logical domains with the UltraSPARC T2 Plus processor and Oracle Solaris 10 increases flexibility, isolates workload processing, and improves the potential for maximum 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 operating system and the server hardware. Operating system instances controlled and supported by the hypervisor are called *guest domains*. The *control domain* handles communication to the hypervisor, hardware platform, and other domains to create and control guest domains. Guest domains are granted virtual device access via a *service domain,* which controls both the system and hypervisor, and also assigns I/O.

To support virtualized networking, Oracle VM Server for SPARC implement a virtual L2 switch (vswitch), to which guest domains can be connected. Each guest domain can be connected to multiple vswitches, and multiple guest domains can connect 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 within the same server. This approach also gives guest domains a direct communication channel to the network (Figure 18). Each guest domain believes 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.

[2] Although supported, running 128 domains is not generally recommended.

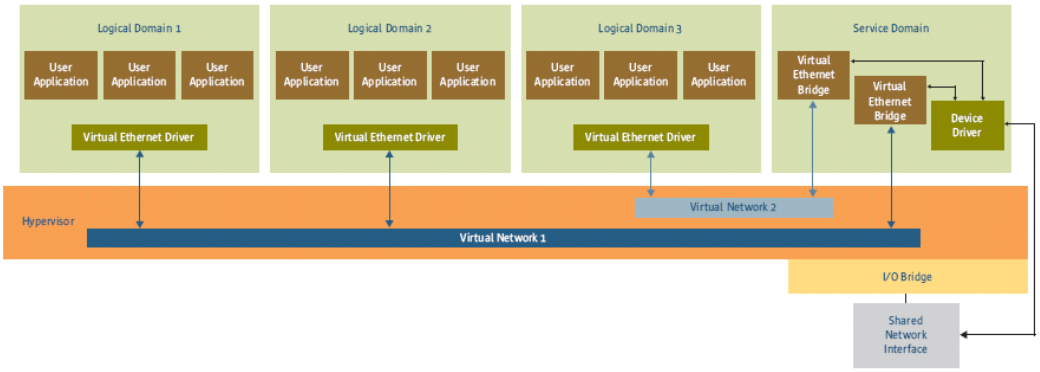

Figure 18. Data moves directly between a domain and a virtualized device.

Oracle Solaris Containers

A Container is a virtualized operating system environment created within a single instance of Oracle Solaris. Containers can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability because processes in one Container are prevented from interfering with processes running in another.

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 Container. Resource pools provide the capability to separate workloads so that CPU resource consumption doesn't overlap. They also provide a persistent configuration mechanism for processor sets and scheduling-class assignment. In addition, the dynamic features of resource pools enable administrators to adjust system resources in response to changing workload demands.

Fault Management and Predictive Self-Healing

Oracle Solaris 10 introduced a new architecture for building and deploying systems and services capable of fault management and predictive self-healing. 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 software misconfiguration problems.

- **Oracle Solaris fault manager.** The fault manager facility collects data relating to hardware and software errors, automatically and silently detecting and diagnosing the underlying problem and using an extensible set of agents to automatically take the faulty component offline. Easy-to-understand diagnostic messages link to articles in Oracle's knowledgebase to clearly guide administrators through corrective tasks that require human intervention. The open design of the fault manager facility also permits administrators and field personnel to

observe the activities of the diagnostic system. With Oracle Solaris fault manager, the overall time from fault condition to automated diagnosis and human intervention is greatly reduced. The result: increased application uptime.

- **Oracle Solaris service manager.** The service manager facility in Oracle Solaris 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, the service manager 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. The service manager feature is also simple to deploy. Developers can convert most existing applications to take full advantage of Oracle Solaris service manager features by simply adding a simple XML file to each application.

Predictive self-healing and fault management provide the following capabilities on Oracle Sun SPARC Enterprise T5440 servers:

- **CPU offlining.** Takes any core or thread offline that's been deemed faulty. Offlined CPUs are stored in the resource cache and stay 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 that have been marked as faulty. Pages are stored in the resource cache and stay retired on reboot, unless the offending 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.

Oracle Solaris Cool Tools for SPARC: Performance and Rapid Time to Market

No matter how compelling new hardware or OS platforms might be, organizations must be assured that the costs and risks of adoption are in line with the rewards. In particular, organizations want to be able to continue to leverage the considerable advantages of popular commercial and open source software. Developers don't want to have to switch compilers and basic development tools. Administrators can scarcely afford a more-complex support matrix or more time spent getting applications to run effectively in a new environment. Oracle's Cool Tools Program is designed to take the cost and risk out of deploying Web-tier environments.

Application Selection

Application selection helps identify those applications that stand to benefit from CoolThreads technology. The CoolThreads Selection Tool (coolst) helps determine application suitability for the UltraSPARC T1, UltraSPARC T2, and UltraSPARC T2 Plus architectures, accelerating the

understanding of application execution and helping to take the risk out of investment decisions and other tasks. The tool measures 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 productivity with their chosen tools.

- **GCC for SPARC Systems.** Specifically tuned and optimized for SPARC systems, GCC for SPARC Systems (GCC4SS) 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, GCC4SS 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 performance, optimizing C, C++, and Fortran compilers for Oracle Solaris on SPARC and x86 platforms. Command-line tools and a NetBeans-based IDE are provided for application performance analysis and debugging of mixed source language applications. In addition to providing multiplatform support, Oracle Solaris Studio 12 compilers are compatible with GCC, Visual C++, C99, OpenMP, and Fortran 2003.
- **Binary Improvement Tool and Simple Performance Optimization Tool.** Used for code coverage analysis, Binary Improvement Tool (BIT) provides instruction and call-count data at runtime, improving developer productivity and application performance. BIT does not require source code and works with both executables and libraries. Simple Performance Optimization Tool (SPOT) also helps improve developer productivity by automating the gathering and reporting of code data.
- **Sun Memory Error Discovery Tool (Discover).** Memory access errors can be one of the hardest types of errors to detect, because symptoms of the error typically appear arbitrarily far from the point where the error occurred. The Sun Memory Error Discovery Tool (Discover) is designed to detect and report common memory access errors. Reported errors include accessing uninitialized memory, writing past the end of an array, or accessing memory after it has been freed.

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.** Serving as an online monitoring tool for core use of the UltraSPARC T2 Plus processor, corestat provides a more-accurate measure of processor and system use than do tools that only measure the use of individual threads. Implemented as a Perl script and updated for UltraSPARC T2 Plus processor awareness, corestat aggregates instructions executed by all

of the threads on a single core, revealing the cycles per instruction of key workloads and indicating where more tuning is needed.

- **Automatic Tuning and Troubleshooting System.** In the interest of automating application tuning, Automatic Tuning and Troubleshooting System (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 12.

Deployment

Cool Tools 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.** Serving as an onsite "virtual" tuning expert, Cool Tuner delivers system performance improvements by automatically applying current best practices, including both patching and tuning. Depending on administrator experience, Cool Tuner can save hours to weeks of effort tuning servers based on CoolThreads technology.
- **Cool Stack.** This collection of the most commonly used free and open-source applications has been preoptimized for servers based on CoolThreads technology running Oracle Solaris. Including such popular applications as Apache, Perl, PHP, Squid, Tomcat, and MySQL database, these applications have been recompiled with Sun Studio 12 compilers to deliver a 30 percent to 200 percent performance improvement over standard binaries compiled with GCC. Cool Stack applications also bring performance benefits to any SPARC system.
- **Consolidation Tool for Oracle's Sun Fire Servers.** Powerful Solaris Containers offer myriad consolidation possibilities, and the Consolidation Tool for Oracle's 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. As a result, you achieve fast, high-quality, consolidated deployments.

Conclusion

Delivering on the demands of IT services applications and virtualized, ecoefficient data centers requires a comprehensive approach that includes innovative processors, system platforms, and operating systems, along with leading application, middleware, and management technology. With its strong technology positions and R&D investments in all of these areas, Oracle is in a unique position to deliver on this vision. Far from futuristic, Oracle has effective solutions today that can help organizations cope with the need for performance and capacity, while effectively managing space, power, and heat.

Building on the successful UltraSPARC T1 and UltraSPARC T2 processors from Oracle, the UltraSPARC T2 Plus processor extends the proven benefits of Oracle's CMT architecture into the most mission-critical applications and OLTP database workloads supporting IT services. With up to 256 threads, up to 512 GB of memory, and massive I/O bandwidth, these systems represent ideal consolidation and virtualization platforms for essential applications and services. Providing 64 threads per processor, on-chip memory management, a PCI-E interface, on-chip cryptographic acceleration, and cache coherency for multisocket support, the UltraSPARC T2 Plus processor fundamentally redefines the capabilities of a modern processor. The Oracle Sun SPARC Enterprise T5440 servers take advantage of these strengths to provide a powerful and highly scalable server platform, while delivering new levels of performance and performance-per-watt in a compact 4RU rackmount chassis. The result is a data center infrastructure that can take the place of racks of existing equipment—in the process lowering power and space requirements.

Oracle's Sun SPARC Enterprise T5440 server provides the computational, networking, and I/O resources needed by the most demanding enterprise applications, databases, and workloads—facilitating highly effective consolidation efforts. With end-to-end support for multithreading and virtualization, these systems can consolidate workloads and effectively use system resources even as they preserve investments in SPARC and Oracle Solaris technology, and provide tools for open-source software environments. With innovations such as Oracle VM Server for SPARC, Oracle Solaris Containers, and Java technology, organizations can adopt these radical new systems for their most important projects—acting responsibly toward both the environment and the bottom line.

ORACLE ise T5440
Server Architecture
October 2009

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

Oracle is committed to developing practices and products that help protect the environment

SOFTWARE. HARDWARE. COMPLETE.