SUN FIRE™ T1000 and T2000 SERVER ARCHITECTURE

Unleashing the UltraSPARC® T1 Processor with CoolThreads™ Technology

White Paper
December 2005

Table of Contents

Executive Summary

Marked by the prevalence of web services and service-oriented architecture (SOA), the emerging *Participation Age* promises the ability to deliver rich new content and high-bandwidth services to larger numbers of users than ever before. Through this transition, organizations across many industries hope to address larger markets, reduce costs, and gain better insights into their customers. At the same time, an increasingly broad array of wired and wireless client devices are bringing network computing into the everyday lives of millions of people, redefining data center scalability and capacity requirements.

While these developments offer considerable potential benefits, many organizations face the seemingly contradictory challenges of scaling their services and improving utilization within a fixed envelope of space, power, and cooling. Even as they struggle to add performance and capacity, real hard limits of available real estate and power are abundantly clear. In fact, IDC now predicts a 50 percent increase in the installed base of servers in the United States alone by 2009, further estimating the annual power bill for that installed base at five billion dollars[1] (excluding cooling). Gartner now estimates that more than 80 percent of all data centers are already constrained by electrical power, physical space, or cooling capacity[2]. Simply adding more hot and power-hungry systems is clearly no solution, only contributing to new levels of consumption, complexity, and waste. Efficient power and cooling are critical, indicated by data center managers who recently reported a 10 percent or more increase in power requirements over the past year, with fully 41 percent reporting that they anticipated upgrading their power and cooling systems over the next three years.[3]

Consistent with Sun's Throughput Computing initiative and eco-responsibility commitment, Sun Fire™ T1000 and T2000 servers with CoolThreads™ technology provide a welcome new approach. These systems are powered by the breakthrough innovation provided by the UltraSPARC® T1 processor and they completely change the equation on space, power, and cooling in the data center. The UltraSPARC T1 processor delivers up to seven times the performance of its competitors[4], effectively condensing the resources of a large symmetric multiprocessing (SMP) system onto a single chip. Supporting up to 32 active threads, the UltraSPARC T1 processor offers massive improvements in application scalability and processor utilization—presenting opportunities for considerable application consolidation while requiring only about as much energy to operate as a common light bulb.

Best of all, Sun is bringing this cutting-edge technology to market in the familiar and proven SPARC® architecture and the compact and reliable enterprise-grade Sun Fire T1000 and T2000 servers. Thanks to Sun's Binary Compatibility Guarantee and the robust Solaris™ Operating System (OS), the increased resources and increased utilization from these systems are immediately available to existing applications and environments. With its strong research and development investments and extensive intellectual property in processor and system design, operating systems, middleware and management software, Sun is uniquely positioned to rapidly bring the benefits of Throughput Computing to its customers, without requiring so much as a recompile.

1. "Server Power Consumption Reemerges as a Critical Cost Factor in Datacenters", Vernon Turner, IDC, August 2005
2. Informal poll from Gartner's 2004 Datacenter Conference, as reported in "Dealing with a Less Than Perfect Data Center Location", by Johanna Ambrosio, Data Center Futures, TechTarget mailing, January 21, 2005
3. Joint study published by AFCOM and InterUnity Group in April 2005, and reported on at http://searchdatacenter.techtarget.com/originalContent/0,289142,sid80_gci1081270,00.html
4. RSA and DSA sign operation @1024-bit, based on cryptography performance tests

Chapter 1
Throughput Computing Realized

Businesses today are increasingly defined by their applications, and now more than ever, an organization's prospects for success are increasingly fixed to its ability to deploy technology in an agile and effective fashion. The risks are extreme. In today's competitive and highly-regulated business environment, the cost of technology failure can be rapid and severe. Even small lapses in IT competence can result in wide-spread damage and loss.

Business Requirements

Increasing the pressure, an endless variety of new networked devices and users are demanding ever-higher levels of performance, capacity, availability, and security from the applications and services that serve them. Real estate concerns along with very real and rising energy costs for both power and cooling are now significant factors that discourage merely adding endless racks of traditional servers. The cost and complexity of managing very large numbers of systems is another pressing concern, especially when coupled with the very low levels of utilization typically found in traditional infrastructure.

To respond to these myriad challenges, business must:

- Grow infrastructure and increase application throughput, capacity, and performance to address pressing business needs as well as capture new customers and opportunities
- Reduce power, cooling, and real estate costs both to save money and to facilitate necessary growth and scalability for the future
- Push for effective consolidation to counter high complexity and management costs, resulting in fewer systems that accomplish more work
- Increase virtualization of resources to aid with business agility and improved resource utilization
- Maintain application compatibility and enhance security across the organization to preserve investments and limit risks to the firm and its clientele

Beyond mere packaging, these issues drive to the very technology used to design processors, systems, and applications. Processor design in particular can have enormous ramifications for business-level issues and solutions. Unfortunately, traditional high-frequency, single-threaded processors are increasingly yielding diminishing returns. Even with ever-higher clock rates, these processors are producing only small improvements in real-world application performance. At the same time, these high-frequency processors generate escalating costs in the form of higher levels of power consumption, and significantly higher levels of heat load that must be addressed by multiple large and expensive HVAC systems. With economic and competitive pressures at an all-time high, most understand that significant change is needed.

While optimistic marketing statements constantly call attention to presumably impressive multiple-gigahertz frequencies and high levels of cache for new generations of processors, corresponding small gains in real-world system performance and productivity continue to frustrate IT professionals. Throughput Computing, along with Sun's focus on optimizing real workload performance is designed to help resolve these divergent trends. This approach provides higher levels of *delivered* performance and computational throughput while greatly simplifying the data center. Understanding the importance of throughput computing requires a look at how both processors and systems have been designed in the past, and the trends that are defining better ways forward.

The Diminishing Returns of Traditional Processor Design

The oft-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. For over three decades the pace of Moore's law has held, driving processor performance to new heights. Processor manufacturers have long exploited these chip real estate gains 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 pre-fetches

While these techniques have produced faster processors with impressive-sounding multiple-gigahertz frequencies, they have largely resulted in complex, hot, and power-hungry processors that don't serve many modern applications, or the constraints of today's data centers. In fact, many of today's data center workloads are simply unable to take advantage of the hard-won ILP provided in these processors. As shown in Table 1, applications with high shared memory and data requirements are typically more focused on processing a large number of simultaneous threads (thread-level parallelism, TLP) rather than running a single thread as quickly as possible (ILP).

Table 1. Attributes of common commercial workloads favor thread-level parallelism over instruction-level parallelism

Workload Attributes	**Web-Centric**	**Application-Centric**		**Data-Centric**		
Benchmark	**Web (SPECweb99)**	**Application (SPECjAppServer2002)**	**SAP-SD 2Tier**	**Data (TPC-C)**	**SAP-SD 3Tier (DB)**	**DSS (TPC-H)**
Application category	Web server	Server, Java™	ERP	OLTP	ERP	DSS
Instruction-level parallelism	Low	Low	Medium	Low	Low	High
Thread-level parallelism	High	High	High	High	High	High
Instruction/Data working set	Large	Large	Medium	Large	Large	Large
Data sharing	Low	Medium	Medium	High	High	Medium

(SPEC, SPECweb, SPECjAppServer, and SPECjbb are registered trademarks of the Standard Performance Evaluation Corporation (SPEC). SAP is a registered trademark of SAP AG in Germany and other countries. TPC-C and TPC-H are trademarks of the Transaction Processing Performance Council (TPC). For more information see www.spec.org, www.sap.com/benchmark, and www.tpc.org.)

Complicating matters, the disparity between processor speeds and memory access speeds means that memory latency dominates application performance, erasing even very impressive gains in clock rates. While processor speeds continue to double every two years, memory speeds have typically doubled only every six years. This growing disconnect is the result of memory suppliers focusing on density and cost as their design center, rather than speed. Unfortunately, this relative gap between processor and memory speeds leaves ultra-fast processors idle as much as 85 percent of the time, waiting for memory to return required data. Ironically, as traditional processor execution pipelines get faster and more complex, the effect of memory latency grows—fast, expensive processors spend more cycles doing nothing. Worse still, idle processors continue to draw power and generate heat. Its easy to see that frequency (gigahertz) is truly a misleading indicator of real performance.

Figure 1 illustrates how even doubling processor performance (frequency) often provides only a small relative increase in application performance. In this example, though the compute time is reduced by half, only a small overall improvement in execution time results, due to the constant and dominant influence of memory latency.

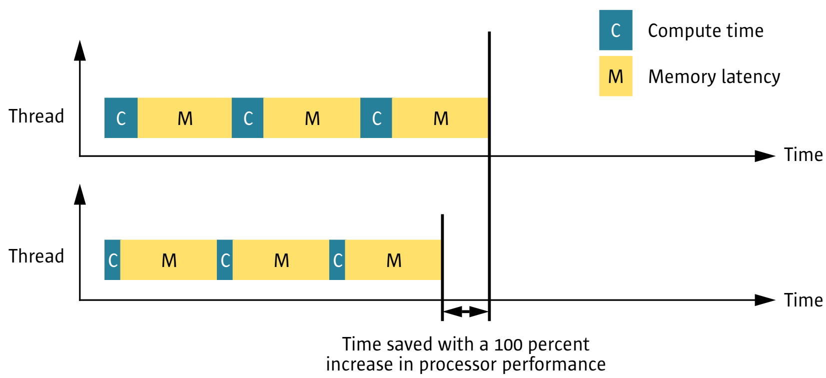


Figure 1. Increasing single-threaded processor performance by 100 percent (a 50-percent reduction in compute-time) provides only a small relative gain in application performance due to memory latency

While some vendors have seemingly awakened to the inherent limitations of traditional, frequency-based processor designs, they are now attempting to graft power-saving technologies and multiple cores onto old, once-discarded architectures. Unfortunately, these efforts represent stop-gap measures at best. Effective approaches can only be realized with fundamentally new processor designs that deliver truly compelling benefits.

Throughput Computing from Sun: A new Approach to Scalability and Eco-Responsibility

In the face of these challenges, Sun's Throughput Computing strategy provides an effective combination of processor, platform, OS, and application technology. The result is a systemic focus on the performance of key workloads, not just the frequency of the processor, or the performance of a single thread of execution:

- With up to seven times the performance of other modern processors, UltraSPARC T1 processors with CoolThreads technology provide massive amounts of thread level parallelism and increased application throughput through chip multithreading (CMT).
- Sun Fire T1000 and T2000 servers deliver the considerable resources of UltraSPARC T1 processors in reliable enterprise-class 1U and 2U platforms while significantly improving power consumption and heat generation.
- The innovative and proven Solaris OS facilitates consolidation through fine-granularity partitioning and virtualization, security, and very high levels of utilization while staunchly upholding Sun's binary compatibility guarantee.
- Compilers, development tools, middleware, and an end-to-end systems strategies leverage the resources of innovative CMT processors, resulting in real-world application performance improvements.

Figure 2. The UltraSPARC T1 Processor with CoolThreads technology

Sun Fire™ Servers Based on the UltraSPARC T1 Processor with CoolThreads™ Technology

Sun Fire T1000 and T2000 servers represent clean-sheet designs that leverage the considerable resources of the UltraSPARC T1 processor (Figure 2) to address the challenges of modern data centers. Optimized for network workloads, these systems deliver up to three times the throughput of competitive systems, while yielding up to four times better performance per watt[1], and up to four times better SWaP (as evaluated by the Space, Watts, and Power metric described later in this document).

With large memory support and the breakthrough performance of the UltraSPARC T1 processor, these systems are ideally suited for compute, data, and transaction-intensive applications. With a six- or eight-core UltraSPARC T1 processor and dense 1U packaging, the Sun Fire T1000 server (Figure 3) is an ideal platform for web, portal, network or security servers. With considerable expansion capabilities and features such as hot-swappable redundant fans, power supplies, and disk drives, the 2U Sun Fire T2000 server is ideal for OLTP, CRM, ERP, database, and collaboration. Both servers excel in the delivery of Java based applications and services, as proven by a suite of world record benchmarks at the product introduction in December 2006 (http://www.sun.com/servers/sparc_benchmarks/).

Sun Fire T1000 Server

Sun Fire T2000 Server

Figure 3. Sun Fire T1000 and T2000 servers with CoolThreads Technology

1.Disclosure: SPEC and SPECweb are registered trademarks of the Standard Performance Evaluation Corporation. Sun Fire T2000 server (8 cores, 1 chip) 14001 SPECWeb2005. IBM eServer Xseries x346 (2 cores, 2 chips) 4348 SPECWeb2005. Sun Fire T2000 server results submitted to SPEC and other results from www.spec.org as of 12/06/05. IBM x346 specifications from product brochure, 09/05/05: http://www-03.ibm.com/servers/eserver/xseries/x346.html. IBM x346 power rating estimated by calculating 70% of the power supply data reported in the product brochure. Sun Fire T2000 server power consumption taken from measurements made during the benchmark run.

Designed to complement each other, the Sun Fire T1000 and T2000 servers share capabilities that are essential to addressing the needs of the modern data center:

- ***Breakthrough performance in less space:***
  With support for up to 32 compute threads in a single processor, Sun Fire T1000 and T2000 servers provide very high compute density in a very small footprint, easing the limitations that many data centers are currently experiencing. In fact, according to Sun's analysis, companies using UltraSPARC T1 processor-based systems can experience a 10:1 reduction in the number of servers they already have deployed[1].
- ***Lower power costs with CoolThreads technology:***
  CoolThreads technology in the UltraSPARC T1 processor requires significantly less power than competitive processors, making them substantially less expensive to operate and facilitating greater levels of capacity in existing data center facilities. The UltraSPARC T1 processor draws as little as a typical 72 watts per processor, compared with competitive processors that require about the same amount of power *per thread*. This crucial difference means that Sun Fire T1000 and T200 servers provide up to five times the performance per watt of competing servers[2]. Power cost reductions alone from deploying Sun Fire T1000 and T2000 servers can save organizations millions of dollars annually.
- ***Increased reliability through drastic simplification:***
  Consolidation of business-critical applications requires platforms that can deliver very high levels of reliability, availability, and serviceability (RAS). Because the UltraSPARC T1 processor replaces many processors and interconnects found in traditional SMP systems, it can deliver very high levels of RAS efficiency. In addition, Sun Fire T1000 and T2000 servers use far fewer components than competing systems, greatly improving on service intervals. The UltraSPARC T1 processor coupled with the Solaris OS also offer a wealth of technologies to improve reliability, availability and serviceability (detailed later in this document).
- ***Greater application efficiency, security, and investment protection:***
  The Sun Fire T1000 an T2000 servers come installed with the Solaris 10 OS, providing an efficient and secure application environment while further increasing performance and utilization. The Solaris 10 OS multithreaded design takes full advantage of these server's chip multithreaded (CMT) architecture, boosting throughput and efficiency. Solaris Containers consist of a group of technologies that work together to efficiently manage the considerable resources of the UltraSPARC T1 processor, virtualize the environment, and provide a complete, isolated, and secure runtime environment for multiple applications on a single server. The Solaris 10 OS protects against both malicious external attacks and data access violations from the inside, and of course UltraSPARC T1 based systems are fully binary compatible with other SPARC based Solaris systems, and source-code compatible with Solaris based systems running on any architecture.

1.Results achieved with AMT-C tool, a Sun consolidation ROI tool. The comparison was based on 200 older Dell PE 2650 servers with 2x2.4GHz Intel Xeon processors. The consolidation was based on 20 Sun Fire T1000 servers running the Solaris 10 OS and Solaris Containers. The Dell system average processor utilization was 25% compared to an 80% utilization for the consolidated Sun Fire T1000 servers.

2.Disclosure: SPEC and the benchmark name SPECjAppServer are trademarks of the Standard Performance Evaluation Corporation. Comparison between results based on the SPECjAppServer2004 benchmark. The Sun Fire T2000 server (8 cores, 1 chip) achieved 615.64 JOPS@Standard while the HP rx4600 server (4 cores, 4 chips) achived 471.28 JOPS@Standard. Sun Fire T2000 results submitted to SPEC as of 12/06/05. Other results from www.spec.org as of 12/06/05. HP rx4640 server specifications taken on 10/19/05 from http://www.hp.com/products1/servers/integrity/engry_level/rx4640/index.html. HP rx4640 power rating of 1,303 watts taken from HP Enterprise Configurator, 10/19/05. HP rx4640 configured with redundant power, 4x1.6 GHz Itanium processors, 8x2GB DIMMs, no PCI cards, and 2x73GB HDDs. Sun Fire T2000 server power consumption of 320 watts taken from actual measurements made during the benchmark run.

Table 2 compares the features of Sun Fire T1000 and T2000 servers.

Table 2. –Comparison of Sun Fire T1000 and T2000 servers

Feature	Sun Fire T1000 Server	Sun Fire T2000 Server
CPUs	Six- or eight-core 1.0 GHz UltraSPARC T1 processor	Four-, six-, or eight-core 1.0 or 1.2 GHz UltraSPARC T1 processor
Active threads	Up to 32	Up to 32
Maximum memory	16 GB	32 GB
Maximum internal disk drives	One 80 GB SATA disk	Four 73 GB SFF SAS disks
Removable/plugable I/O	N/A	Slimline DVD-R/CD-RW Four USB 1.1 ports
PCI	One PCI-E slot, low profile	Three PCI-E slots, low profile Two PCI-X slots, low profile[a]
Ethernet	Four on-board Gigabit Ethernet ports (10/100/1000)	Four on-board Gigabit Ethernet ports (10/100/1000)
Power supplies	One 300W power supply	Two redundant (N+1) hot-swap 550W power supplies[b]
Fans	Single fan tray assembly	3 redundant, hot-swappable cooling fans
Form factor	1 rack unit (1U)	2 rack units (2U)

a. One PCI-X slot occupied with a disk controller
b. 450 watt power supplies will be available in early 2006

Space, Watts, and Power: Introducing the SWaP Metric

Sun Fire T1000 and T2000 servers deliver leading performance across a range of web and application tier 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 are limited when it comes to understanding the power and density attributes of the systems being compared.

For this reason, Sun has developed the SWaP metric, standing for *Space, Watts, and Performance*. 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 audited benchmarks such as those sponsored by the Systems Performance Evaluation Corporation (SPEC)
- *Space* refers to the height of the server in rack units (RUs)
- *Power* is measured by watts used by the system, taken during actual benchmark runs or from vendor's site planning guides

Table 3 provides an example comparison made using the SWaP metric, comparing the Sun Fire T2000 server to an IBM eSeries Xseries x346 server running the SPECWeb2005 benchmark[1]. As measured by the SWaP metric, the Sun Fire T2000 server exceeds this competitor by over four times.

Table 3. Comparison of a Sun Fire T2000 server with an IBM eServer Xseries x346 server using the SWaP metric

Feature	Sun Fire T2000 Server	IBM eServer x346	Sun Fire T2000 Server Advantage
Space (RU)	2	2	None
Watts	330	438	25%
Performance (composite)	14,001	4,348	3.22 times
Performance/watt	42.427	9.927	4.27 times
SWaP	21.2	5.0	4.24 times

1.Disclosure: SPEC and SPECWeb are registered trademarks of the Standard Performance Evaluation Corporation. Sun Fire T2000 server (8 cores, 1 chip) 14001 SPECweb2005. IBM eServer Xseries x346 (2 cores, 2 chips) 4348 SPECWeb2005. Sun Fire T2000 server results submitted to SPEC and other results from www.spec.org as of 12/06/05. IBM x346 specifications from product brochure, 09/05/05: http://www-03.ibm.com/servers/eserver/xseries/x346.html. IBM x346 power rating estimated by calculating 70% of the power supply data reported in the product brochure. Sun Fire T2000 server power consumption taken from measurements made during the benchmark run.

Chapter 2

The UltraSPARC T1 Processor with CoolThreads Technology

Dissatisfied with the performance characteristics of traditional single-threaded processors, Sun's extensive in-house design team—one of the largest microprocessor design engineering teams in the world—has taken a bold new approach to processor design. Sun understands that the network-computing environments found in most modern data centers are inherently multithreaded, where the execution speed of an individual thread is typically less important than overall application throughput. For this reason, Sun is focusing on processors and architectures that maximize throughput for commercial network-computing workloads. These efforts are resulting in new chip multithreaded (CMT) processor technology that leverages the additional gains delivered by Moore's Law to provide thread-level parallelism rather than instruction-level parallelism.

Chip Multithreaded (CMT) Design

Recently, a number of processor vendors have started making multi-core processors available that place multiple processors on a single die (so-called chip multi-processing). Unfortunately, many of these efforts too have resulted in complex, hot, high-power processors that don't solve the real issues driving processing and productivity in the data center. The most significant potential benefits of multi-core technology derive from a fundamental simplification and rethinking of how processors are designed and built.

Hardware Multithreading Processor Cores

Unlike traditional single-threaded processors and even most current multi-core processors, hardware multithreaded processor cores allow rapid switching between active threads as other threads stall for memory. Figure 4 illustrates a multithreaded processor core such as those found on the UltraSPARC T1 processor. The key to this approach is that each core is designed to switch between up to four threads on each clock cycle. As a result, the processor's execution pipeline remains active doing real useful work, even as memory operations for stalled threads continue in parallel.

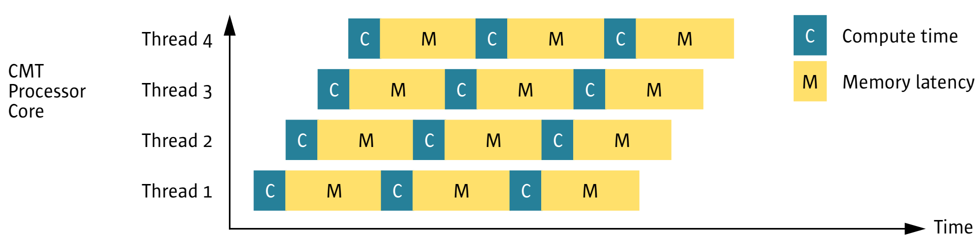


Figure 4. A hardware multithreaded processor core switches between a number of active threads, doing useful work even while threads stall to perform memory-related operations

Hardware multithreading provides real value since it increases the ability of the execution pipeline to do actual work. Utilization of the processor pipeline is greatly enhanced since a number of execution threads now share its resources. The negative effects of memory latency are masked since the processor and memory subsystems remain active in parallel to the processor execution pipeline.

Chip Multithreading with CoolThreads Technology

Like single-threaded processors, multi-processing (multi-core) technology can also be used to scale and multiply the benefits of hardware multithreading. Sun calls the result chip multithreading. Unlike complex single-threaded processors, CMT processors utilize the available transistor budget to implement multiple hardware multithreaded processor cores on a single silicon wafer or chip. Because these individual processor cores implement much simpler pipelines (emphasizing TLP over ILP), they are also substantially cooler and require significantly less electrical energy to operate. This innovative approach results in CoolThreads processor technology—multiple physical instruction execution pipelines (one for each core), with several active thread contexts per pipeline or core.

Application throughput is greatly improved with CoolThreads processors, as is utilization of pipeline resources. Thread-rich applications common in commercial workloads benefit greatly from this model, whether comprised of larger multithreaded applications, or of large numbers of smaller single-threaded applications. The number of simultaneous threads that can be accommodated is quite large, and a wide range of processor designs are possible. Figure 5 illustrates a four-core UltraSPARC T1 processor in which four hardware multithreaded cores are combined into a single CMT processor, supporting up to sixteen active execution threads.

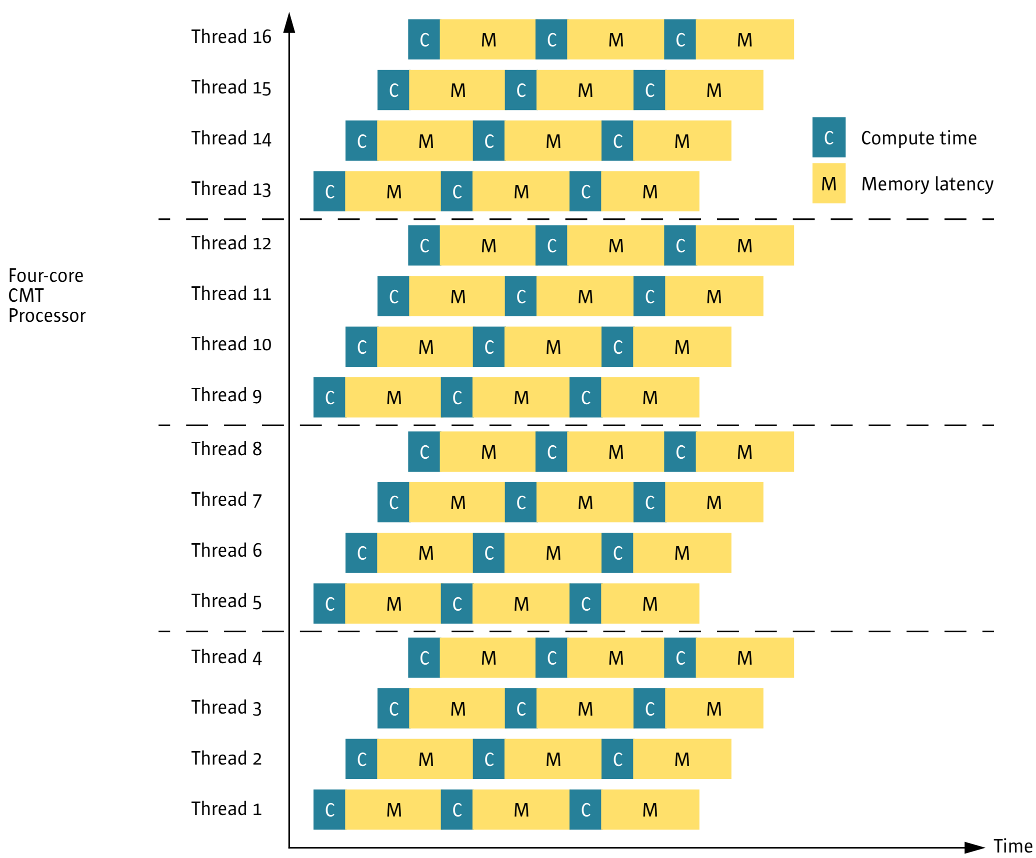


Figure 5. A four-core UltraSPARC T1 processor supporting four multithreaded cores and 16 active thread contexts (an eight-core UltraSPARC T1 processor supports 32 active thread contexts)

UltraSPARC T1 Processor Architecture

The UltraSPARC T1 multi-core, multithreaded processor is the first chip that fully implements Sun's Throughput Computing initiative. Each UltraSPARC T1 processor has either four, six, or eight cores, or individual execution pipelines, all on the same chip—essentially an "SMP system on a chip". Each core, in turn, supports up to four hardware thread contexts, a set of registers that represent the thread's state. The processor is able to switch threads on every clock cycle in a round robin ordered fashion, and skip threads that are stalled (e.g. those threads waiting for a memory access). In spite of its innovative new technology, the UltraSPARC T1 processor is fully SPARC v7, v8, and v9 compatible and binary compatible with earlier SPARC processors.

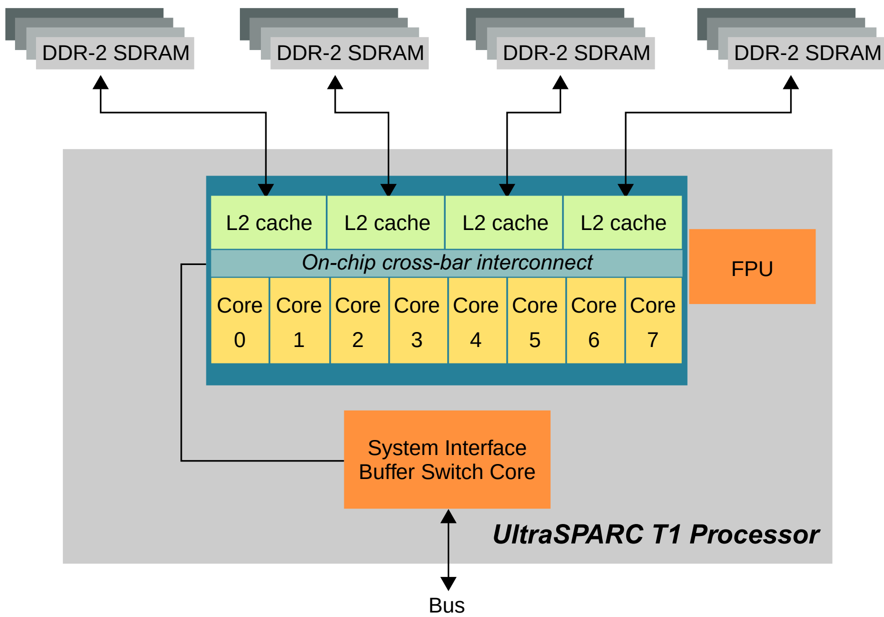


Figure 6. Block-level diagram of an eight-core UltraSPARC T1 processor

As shown in the Figure 6, the individual processor cores are connected by a high-speed, low-latency crossbar interconnect implemented on the silicon itself. The UltraSPARC T1 processor includes very fast interconnects between the processor, cores, memory, and system resources, including:

- A 134 GB/second crossbar switch that connects all cores
- A JBus interface with a 3.1 GB/second peak effective bandwidth
- Four DDR2 channels (25.6 GB/second total) for faster access to memory

The memory subsystem of the UltraSPARC T1 processor is implemented as follows:

- Each core has an Instruction cache, a Data cache, an Instruction TLB, and a Data TLB, shared by the four thread contexts. Each UltraSPARC T1 processor has a twelve-way associative unified Level 2 (L2) on-chip cache, and each hardware thread context shares the entire L2 cache.
- This design results in unified memory latency from all cores (Unified Memory Access, UMA, not Non-Uniform Memory Access, NUMA).
- Memory is located close to processor resources and four memory controllers provide very high bandwidth to memory, with a theoretical maximum of 25GB per second.
- Extensive built-in RAS features include ECC protection of register files, Chipkill, memory sparing, soft error rates and rate detection, and extensive parity/retry protection of caches.

Each core has a Modular Arithmetic Unit (MAU) that supports modular multiplication and exponentiation to help accelerate Secure Sockets Layer (SSL) processing. There is a single Floating Point Unit (FPU) shared by all cores, thus the UltraSPARC T1 processor is not an optimal choice for applications with floating point intensive requirements.

Unsurpassed Reliability, Availability, and Serviceability

By its very nature, any sort of consolidation of business-critical applications or services must be coupled with extremely high levels of reliability, availability, and serviceability (RAS). The innovative UltraSPARC T1 processor together with the Sun Fire T1000 and T2000 servers and Solaris technology provide very high levels of reliability. These capabilities stem not only from processor and system-level features, but from the very way that the processor is designed.

UltraSPARC T1 Processor Technology

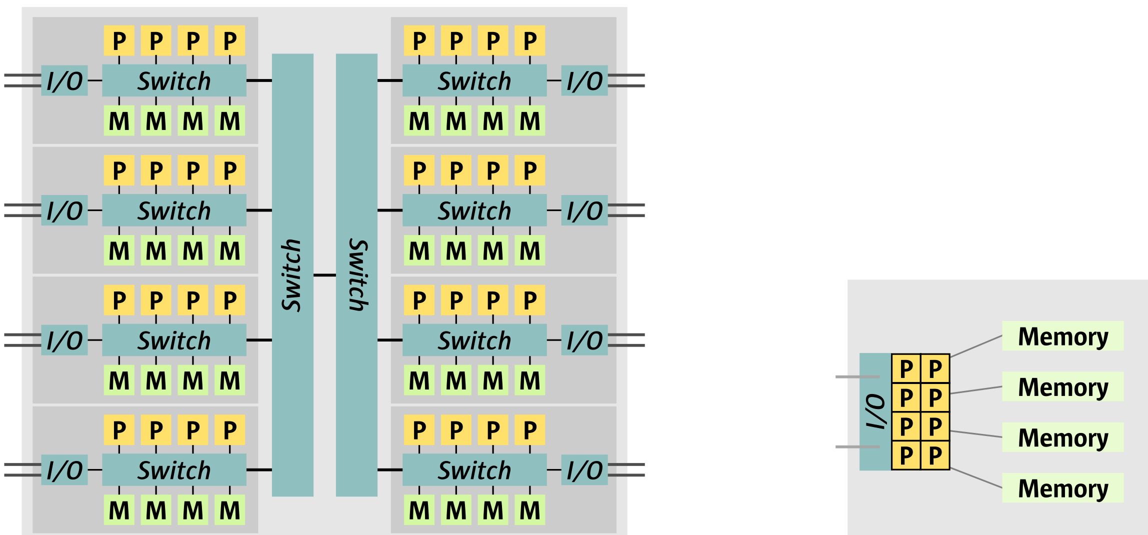
The UltraSPARC T1 processor itself provides a number of built-in features that contribute directly to greater reliability and availability. These features include:

- Extensive error correcting code (ECC) logic on the UltraSPARC T1 processor
- Memory Chipkill support
- Memory DRAM sparing/reconfiguration
- Memory address parity protection
- A memory hardware scrubber and an L2 cache scrubber
- Soft error rates and soft error rate detection
- DRAM channel deconfiguration
- Redundancy (cache RAM row/column sparing, core sparing, and eFUSE)

The UltraSPARC T1 processor provides parity protection on its internal cache memories, including tag parity and data parity on the D-cache and I-cache. The internal 3 MB L2 cache has parity protection on the tags and ECC protection on the data. DIMMs employ ECC as well to help ensure high levels of data integrity. The system reports and logs correctable ECC errors, and errors are corrected as soon as they are detected.

Increased Reliability from CMT technology

Beyond individual features in the UltraSPARC T1 processor, CMT processor design itself can dramatically increase reliability and availability by condensing many physical interfaces and components onto a single processor. Since one UltraSPARC T1 processor essentially replaces many individual processors and interconnects in an equivalent SMP system, the reliability effects can be profound. As shown in Figure 7, the associated consolidation of individual processors and inter-chip communication is greatly simplified in a system based on the UltraSPARC T1 processor.



Traditional 32-thread SMP server: One cabinet, eight processor/memory boards, many switch ASICS and connections

Sun Fire Server with CoolThreads Technology: One motherboard, no switch ASICS, no board connections

Figure 7. By compressing 32 threads onto a single processor, the UltraSPARC T1 processor helps enable system designs that are less susceptible to failure, with fewer parts and connections

The traditional SMP system with 32 single-threaded processors represented on the left of the figure is housed in a refrigerator-sized cabinet and consists of eight processor/memory boards, each with four processors, memory, and an I/O interface. Each board, in turn, has switch ASICS to connect the on-board components, and the cabinet has another set of switch ASICS to connect one group of four boards to the other. In contrast, the Sun Fire server based on the UltraSPARC T1 processor offers a much more integrated and tightly-coupled solution.

With the UltraSPARC T1 processor, a single chip now houses processors and interconnects, and without the need for switch ASICS, the entire system now fits on a single motherboard and has many fewer parts and pins to fail. Moreover, maintenance and servicing is reduced to a single board. In fact, Sun Fire T1000 and T2000 servers are built using many fewer components than even competitive rack-mounted servers. Table 4 lists the component counts of Sun Fire T1000 and T2000 servers as compared to several competitors. Smaller component counts translate directly to larger numbers of years between service events.

Table 4. Sun Fire T1000 and T2000 servers feature very low component counts when compared to competitors

System	Total number of parts
Sun Fire T1000 Server	2357
Sun Fire T2000 Server	3411
IBM p520	4064
HP DL585	5176
Dell 2850	5566
Dell 6850	7086

Protection of On-chip memories

As semiconductor technology continues to trend toward increasing chip densities, processors have inevitably become more susceptible to soft error rates. These soft errors are typically caused by charged particles or radiation and are transient and random in nature—usually doing no damage to the device. Many systems can tolerate some level of soft errors. With these risks in mind, Sun systematically designed the UltraSPARC T1 processor with the appropriate level of protection of its on-chip memories. In general, the UltraSPARC T1 processor protects memory arrays with either single error correction/double error detection (SEC/DED) or parity protection. Redundant arrays are protected with parity while non-redundant arrays are protected with ECC. Table 5 lists the UltraSPARC T1 processor's on-chip memories and their corresponding protection mechanism.

Table 5. Protection mechanisms for UltraSPARC T1 on-chip memories

Memory Array	Protection
Integer Register File	ECC
Floating Point Register File	ECC
L1 Instruction Cache - Data	Parity/retry
L1 Instruction Cache - Tag	Parity/retry
Instruction TLB	Parity/retry
Data TLB	Parity/retry
L1 Data Cache - Data	Parity/retry
L1 Data Cache - Tag	Parity/retry
L2 Cache - Data	ECC
L2 Cache - Tag	ECC

A notable feature of this scheme is the ECC protection of the integer and floating point register files, an extensive level of protection only matched by mainframe-class processors. While processor designs have mainly focused on protecting the data path, caches, and main memories, the register file has largely been neglected. Because the register file is accessed very frequently, the probability of errors is increased, and protecting the register file is critically important. In addition, protecting the register file with ECC prevents errors from quickly spreading to different parts of the system, and can help prevent application crashes or silent data corruption.

Main memory reliability and availability

Standard ECC memory is a proven industry standard technology that has had a considerably positive impact on server reliability. ECC memory is able to detect and correct single bit memory errors, comprising the vast majority of memory errors. However, increased memory capacity, the density of memory on a single DIMM, and the higher speed of memory subsystems have significantly increased the risks of multi-bit memory errors that cannot be corrected by standard ECC memory and result in system hangs. To address this issue, designers implemented Chipkill technology to correct multi-bit memory errors, increasing system availability considerably.

The UltraSPARC T1 processor protects main memory using several mechanisms. Chipkill technology is used to withstand multi-bit memory errors within a DRAM device, including a failure that causes incorrect data on all data bits of the device. The Chipkill mechanism in UltraSPARC T1 processors uses Galois Field instead of Hamming in its Chipkill implementation. The Galois Field algorithm provides higher bandwidth than Hamming Chipkill (21.33 vs. 10.66 GB/second) and it can correct any error contained within a single memory nibble (4 bits), while detecting any uncorrectable errors contained within any two nibbles. When writing data to the DIMM, data is written in the form of a checksum appended to the data. If a single-nibble memory error occurs, then the data is immediately recovered by recalculating the data from the checksum information. This procedure allows the system to correct not only the single-bit errors that standard ECC memory can correct but also 2, 3, and 4-bit errors and even a whole DRAM chip failure. Even assuming larger size memories that increase the likelihood of systems to crash due to a memory error, Chipkill-equipped servers fail at a rate of about two orders of magnitude less than those with an ECC-protected subsystem.

In conjunction with Chipkill, DRAM sparing is implemented in the UltraSPARC T1 processor to improve main memory availability. Where Chipkill detects a failed DRAM chip, DRAM sparing reconfigures a DRAM channel to map out the failed chip, effectively replacing it with a corrected DRAM chip. This technique restores the capability of correcting any random single-nibble error and allows the system to run with minor impaired memory error protection until the DIMM can be replaced.

System-level RAS features

In addition to RAS features in the UltraSPARC T1 processor, Sun Fire T1000 and T2000 servers implement a variety of technologies to improve reliability.

- ***Hot-swappable and hot-pluggable components***

  Sun Fire T2000 server hardware is designed to support hot-swapping of both the fan units and power supplies. By using the proper software commands, these components can be installed or removed while the system is running. Hot-swap technology significantly increases the system's serviceability and availability by providing the ability to replace fan units and power supplies without service disruption.

  The Sun Fire T2000 server features two hot-swappable power supplies so the system can continue to operate if one of the power supplies were to fail or if one power source were to fail. The server also has a single hot-swappable blower unit that works in conjunction with the power supply fans to provide cooling for the internal disk drives. If the blower unit fails, the power supply unit fans provide sufficient cooling to keep the disk subsystem running until the blower can be replaced. The Sun Fire T2000 server also features three hot-swappable system fans. The fans enable the system to continue operating with adequate cooling in the event that one of the fans fails. Four hot-pluggable disk drives are also supported in the Sun Fire T2000 server.

- ***Environmental Monitoring***
  Sun Fire T1000 and T2000 servers feature an environmental monitoring subsystem designed to protect the server and its component against extreme temperatures, lack of adequate airflow, power supply failures, and hardware faults. Temperature sensors are located throughout the system to monitor the ambient temperature of the system and internal components. The server software and hardware help ensure that the temperatures within the enclosure do not exceed safe operational ranges.

  All error and warning messages are sent to the system controller system console and are logged in the Advanced Lights Out Monitor (ALOM) console log file. Required LEDs remain lit after an automatic system shutdown occurs, to aid in problem diagnosis. The power subsystem is handled in a similar fashion by monitoring power supplies and reporting any fault through the front and rear panel LEDs.

- ***Fault Management and Predictive Self-Healing***
  Sun Fire T1000 and T2000 servers feature the latest fault management technologies, including Fault Management and Predictive Self-Healing introduced with the Solaris 10 OS. Self-healing technology enables Sun systems to accurately predict component failures and to mitigate many serious problems before they actually occur. This technology is incorporated into both the hardware and software of the Sun Fire T1000 and T2000 servers. More details on Solaris Fault Management and Predictive Self Healing are provided later in this document.

- ***Automatic System Recovery***
  Sun Fire T1000 and T2000 servers provide for automatic system recovery (ASR) from failures in UltraSPARC T1 processor cores, memory modules, or PCI cards. The ASR functionality allows the system to resume operation after experiencing certain nonfatal hardware faults or failures. When ASR is enabled, the system's firmware diagnostics automatically detect failed hardware components. An auto-configuring capability designed into the system firmware enables the system to unconfigure failed components and to restore system operation. As long as the system is capable of operating without the failed component, the ASR features allows the system to reboot automatically, without operator intervention.

Chapter 3
Sun Fire T1000 and T2000 Server Architecture

Both the Sun Fire T1000 server and the Sun Fire 2000 server have been designed to provide breakthrough performance while maximizing reliability and minimizing power consumption and complexity. This section details the physical and architectural aspects of these systems.

Sun Fire T1000 Server Architecture Overview

The compact Sun Fire T1000 server is an ideal platform for the delivery of horizontally-scaled transaction and web services. The server is designed to address the challenges of today's data center by delivering an excellent price-to-performance ratio with greatly reduced power consumption and a small physical footprint. Depending on the model selected, the Sun Fire T1000 server features a single six- or eight-core UltraSPARC T1 processor.

Enclosure

The 1U Sun Fire T1000 server enclosure is designed for use in a standard 19-inch rack (Table 6). The chassis consists of a base unit, a removable top cover, and a front bezel.

Table 6. Dimensions and weight of the Sun Fire T1000 server

Dimension	U.S.	International
Height	1.75 inches (1 RU)	4.3 centimeters
Width	16.8 inches	42.5 centimeters
Depth	19.0 inches	48.3 centimeters
Weight (with side rails)	24 pounds	10.9 kilograms

The Sun Fire T1000 server includes the following major components:

- An UltraSPARC T1 processor with six or eight cores
- Up to 16 GB of DDR2 SDRAM memory in 8 available memory slots
- Four on-board Gigabit Ethernet ports
- One PCI Express (PCI-E) slots
- One SATA disk drive
- Advance Lights out Management (ALOM) system controller

System Board

The Sun Fire T1000 server system board contains all of the logic components for the system (Figure 8). These components include the UltraSPARC T1 processor and all DIMMS, the I/O bridge ASIC, the I/O subsystem, and the service processor subsystem. The system board also contains the rear-panel I/O connectors for the host system and the service processor as well as a PCI Express (PCI-E) expansion card connector. These components are described below.

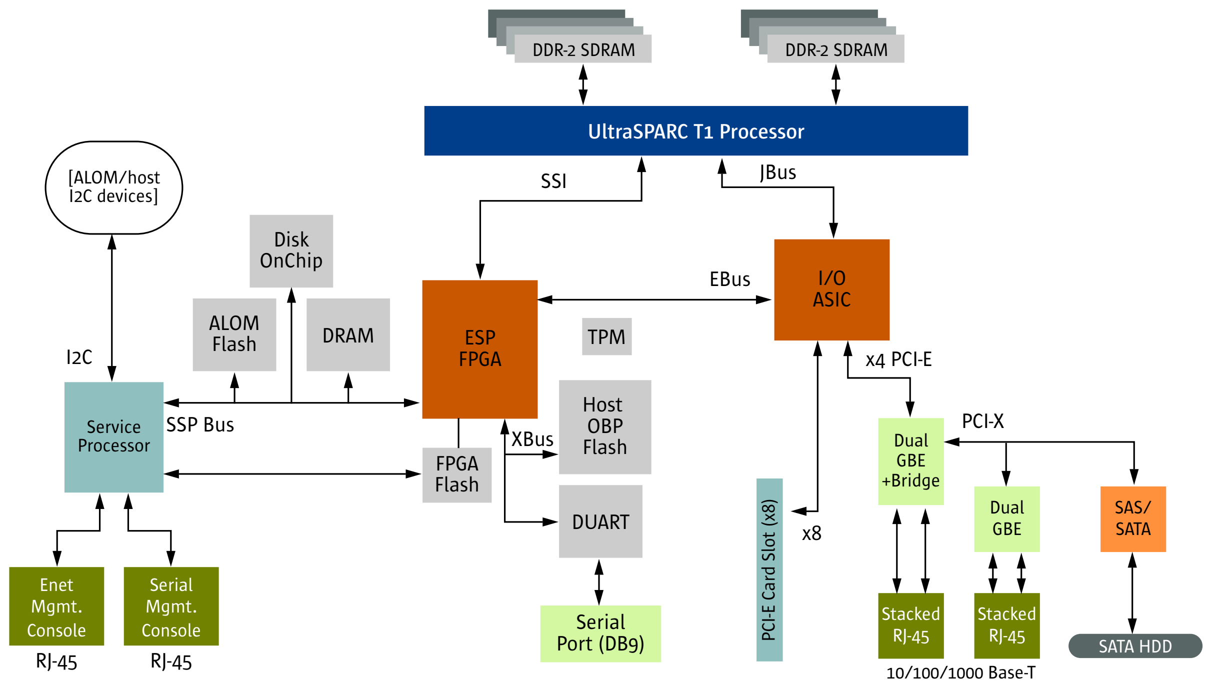


Figure 8. Sun Fire T1000 server block level diagram

- ***Memory Subsystem***
  Each UltraSPARC T1 processor features four UltraSPARC DDR2 memory controllers, each with a 16-byte wide (128 data bits plus 16 ECC check bits) data interface. Two basic memory configurations are supported, with either two or four DIMMs per controller. All DIMMs must have identical capacity. The Sun Fire T1000 server uses only two of the four UltraSPARC T1 DDR2 memory controllers, with a maximum memory configuration of eight DIMMs. DIMM sizes of 512 MB, 1 GB, and 2 GB are supported for a maximum memory configuration of 16 GB. DIMMs must be installed four at a time (with one pair on each controller), yielding a four-DIMM configuration and an eight-DIMM configuration.

- ***I/O Subsystem***
  The UltraSPARC T1 processor contains a JBus controller and an I/O ASIC is employed in the Sun Fire T1000 server to translate between JBus and an industry-standard I/O bus. The I/O ASIC that performs this function in the Sun Fire T1000 server contains two PCI Express (PCI-E) root complexes, each providing one PCI-E link. Each PCI-E link supports a maximum port width or eight lanes. The two PCI-E interfaces operate independently of each other and each supports link widths of one, two, four, or eight lanes. The JBus is clocked at 200 MHz.

  The first PCI-E interface from the I/O ASIC connects directly to the system's PCI-E expansion slot. Full link speed and link width (x8) are supported. The second PCI-E interface connects to an on-board combination of an PCI-E to PCI-X bridge and Ethernet controller, acting as the root of all integrated I/O within the system. The bridge/ Ethernet device, and the PCI-E link operate with an x4 link width.

- *PCI-Express slot*
  The Sun Fire T1000 server provides one PCI express (PCI-E) slot for low profile cards. This slot supports 1x, 4x, and 8x link width cards.
- *Network interfaces*
  The Sun Fire T1000 server includes four 10/100/1000Base-T auto-negotiating Ethernet ports, implemented with two dual-ported controllers. Each of the four Ethernet RJ-45 connectors includes two LEDs; a green link indicator, and an amber activity indicator. A fifth 10/100Base-T port is available on the back panel (Figure 9) for a remote management connection as part of the ALOM system controller.

Storage
Internal hard drive storage is supported by a serial attached SCSI (SAS)/serial ATA (SATA) controller. The combined SATA/SAS controller allows the Sun Fire T1000 server to provide a low-cost SATA configuration. The Sun Fire T1000 server supports a single internal hard drive. A diskless configuration is also supported. An expansion card is necessary for the provision of external storage connectivity on the Sun Fire T1000 server.

Power Supply
The server includes a single 300W AC power supply. To provide for the best (i.e. lowest) possible power rating and to minimize excess heat generation, high efficiency is a key aspect of the power supply design. The power supply contains its own cooling fan that also provides cooling for the system's hard disk drive. The power supply operates from a wide range of 100 to 240V AC input.

Front and Back Panels
Figure 9 illustrates the front and rear panels of the Sun Fire T1000 server.

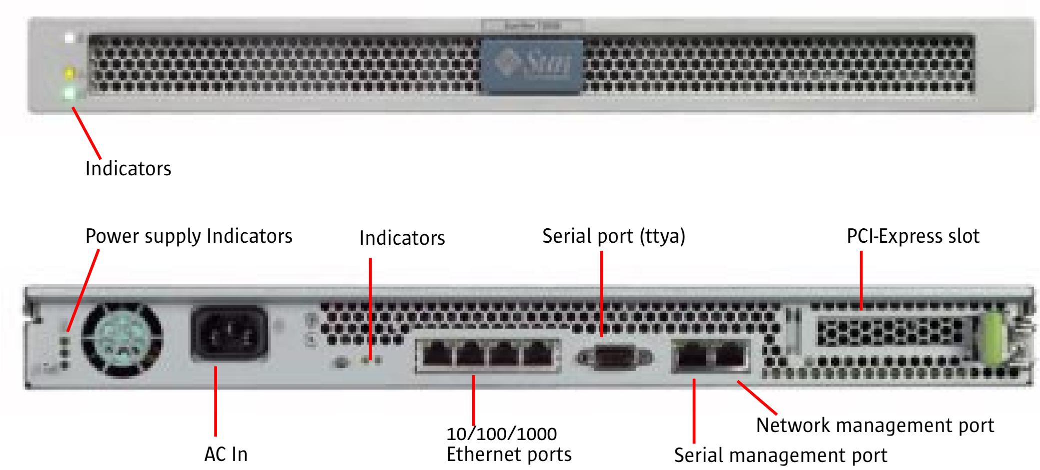


Figure 9. Sun Fire T1000 server, front and rear panels

External features of the Sun Fire T1000 server include:

- Front and rear indicator lights provide locator (white), service required (amber), and activity status (green) for the system.
- Rear power-supply indicator lights convey the status of the power supply
- A single AC plug is provided on the rear of the system
- Four 10/100/1000Base-T autosensing Ethernet ports are provided
- A DB-9 TTYA serial port is provided for serial devices (not connect to the ALOM system controller serial port)
- Two management ports are provided for use with the ALOM system controller. The RJ-45 serial management port provides the default connection to the ALOM controller. The network management port supports an optional RJ-45/10/100Base-T connection to the ALOM system controller. The serial management port is always available. The network management port is configured through the serial management port.

Sun Fire T2000 Server Architecture

The expandable Sun Fire T2000 Server is optimized to deliver transaction and web services, including J2EE application services, enterprise application services (ERP, CRM, and SCM) and distributed databases. The Sun Fire T2000 server is also an ideal platform for consolidated tier-1 workloads

Enclosure

The Sun Fire T2000 server features a compact, yet expandable 2U rack-mountable chassis (Table 7). The Sun Fire T2000 server gives customers the flexibility to scale their processing and I/O needs without wasting precious space.

Table 7. Dimensions and weight of the Sun Fire T2000 servers

Server/Dimension	U.S.	International
Height	3.5 inches (2 RU)	8.9 centimeters
Width	17.3 inches	44.0 centimeters
Depth	24.3 inches	61.7 centimeters
Weight (without PCI cards or rack mounts)	37 pounds	17 kilograms

The Sun Fire T2000 server includes the following major components:

- An UltraSPARC T1 processor with four, six, or eight cores
- Up to 32 GB of DDR2 SDRAM memory in 16 available memory slots
- Four on-board Gigabit Ethernet ports
- Three PCI Express (PCI-E) and two PCI-X slots (one occupied by a disk controller)
- Up to four hot-pluggable SAS disk drives
- Advance Lights out Management (ALOM) system controller

System Board

Like the Sun Fire T1000 server, the Sun Fire T2000 server system board contains all of the logic components for the system, including the UltraSPARC T1 processor and all DIMMs, the I/O bridge ASIC, the I/O subsystem, and the service processor subsystem. The system board also contains the rear-panel I/O connectors for the host system and the service processor. A logical block-level diagram of the Sun Fire T2000 system board is shown in Figure 10.

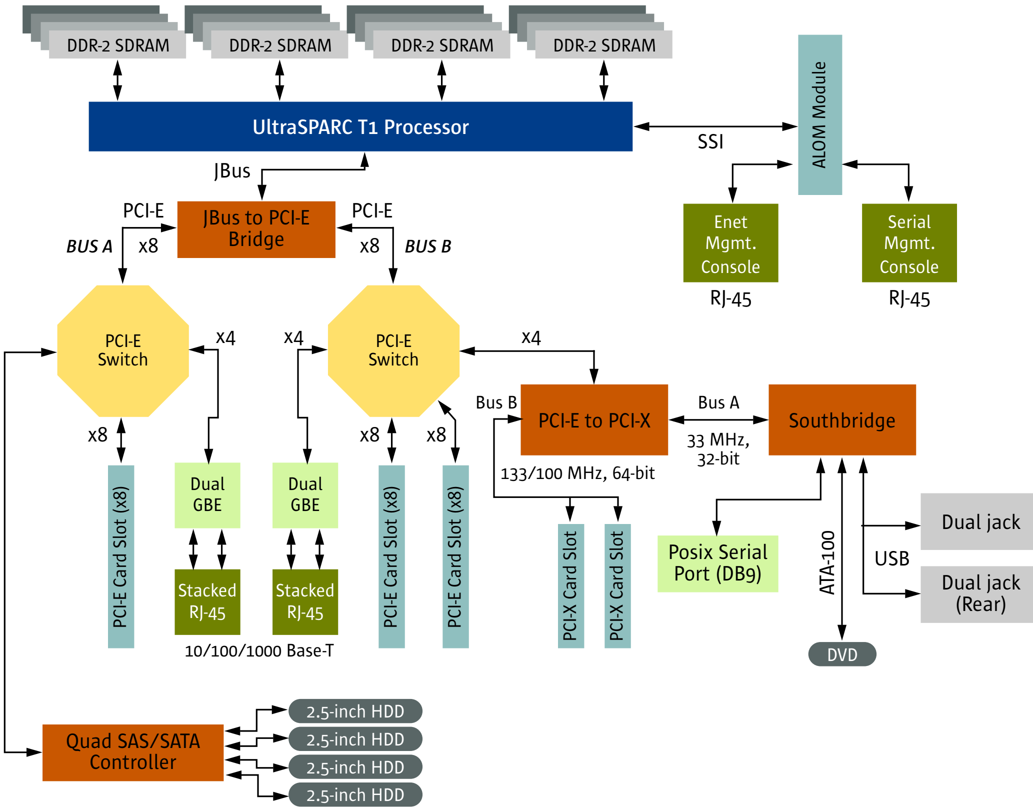


Figure 10. Sun Fire T2000 server block level diagram

- ***Memory***

  Unlike the Sun Fire T1000 server, the Sun Fire T2000 server uses all four of the available UltraSPARC T1 DDR2 memory controllers for a maximum memory configuration of 16 DIMMs. All DIMMs must have identical capacity. DIMMs must be installed eight at a time (with one pair on each controller) yielding an 8-DIMM configuration or a 16-DIMM configuration. DIMM sizes of 512 MB, 1 GB, and 2 GB are supported, resulting in a maximum capacity of 32 GB for the system.

- ***I/O Subsystem***

  The Sun Fire 2000 server utilizes an I/O bridge ASIC that translates between the JBus interface on the UltraSPARC T1 processor and an industry-standard I/O bus. On the Sun Fire T2000 server, two PCI-E root complexes are provided, with each link having a maximum port width of eight lanes (x8). The two PCI-E interfaces operate independently, and each supports link widths of one, two, four, or eight lanes. The JBus runs at 200 MHz while the PCI-E links run at 2.5 GHz.

  The three external x8 PCI-E expansion slots are connected to port A or B of the I/O ASIC through a pair of PCI-E switches. Full link speed and link width are supported in each slot. The two PCI-X slots are connected to port B of the I/O ASIC through a PCI-E switch and a PCI-E to PCI-X bridge. Port A devices share bandwidth with on-board SAS/SATA and dual Gigabit Ethernet devices. Port B slots share bandwidth with an on-board dual Gigabit Ethernet device.

- *Expansion slots*
  The Sun Fire T2000 server is well equipped with both legacy PCI-X slots, and the more current PCI-E high-speed slots. These slots can accommodate low-profile PCI cards that adhere to the MD2 physical specifications. Three x8 PCI-E slots are provided in the Sun Fire T2000 server, and these also support x1 and x4 cards. Most cards, including InfiniBand network cards, currently do not require more than an x4 slot for full-bandwidth operation. Two PCI-X slots are provided on the Sun Fire T1000 server to support older PCI or PCI-X I/O cards. The PCI-X slots are both 133 MHz, 64 bits wide, and 3.3V, providing the highest speed available in PCI-X configurations.
- *Network interfaces*
  Like the Sun Fire T1000 server, the Sun Fire T2000 server includes four 10/100/1000Base-T auto-negotiating Ethernet ports, implemented with two dual-ported controllers. Each of the four Ethernet RJ45 connectors includes two LEDs, a green link indicator, and a yellow activity indicator. A fifth 10/100Base-T port is available on the back panel (Figure 11) for remote management connection as part of the Advance Lights Out Management (ALOM) controller.
- *USB Ports*
  The Sun Fire T2000 server provides four USB 1.1 ports. Two of the ports are located on the left of the front panel of the server and are associated with one USB controller. The second two USB ports are located on the right side of the rear of the panel and are associated with a second USB controller.

Storage

The Sun Fire T2000 server supports up to four hot-pluggable hard disk drives, controlled by an LSI SAS1064 controller, providing four ports of SAS connectivity to the disk drives at bandwidths of 3 Gb/second full duplex for each disk. The disks provided with the Sun Fire T2000 server are 73 GB, 10000-rpm SAS disks, 2.5-inch small form factor server grade, and are certified for 24x7 operation. Due to the small physical size of these drives, a the high spindle speeds, access times are extremely good. A side benefit of using small disks is that they allow designers to maximize the air intake area at the front of the server to improve airflow, further increasing environmental margins and server reliability.

An optional slimline DVD-R/CD-RW is also provided.

Power supplies

The Sun Fire T2000 server is equipped with a dual redundant hot-swappable power-supply system. One power supply is sufficient to run a full-populated server. However, for maximum protection against power supply failures, Sun recommends that both power supplies be installed in the system at all times. Both power supplies at First Customer Ship are rated at 550 watts each, these will be replaced by 450 watt power supplies during early 2006. In normal operation, the power supplies share the power demands of the system equally between the pair. The power supplies can also be plugged into different electrical sources if dictated by availability requirements.

Front and Back Panels

Figure 11 illustrates the front and back panels of the Sun Fire T2000 server.

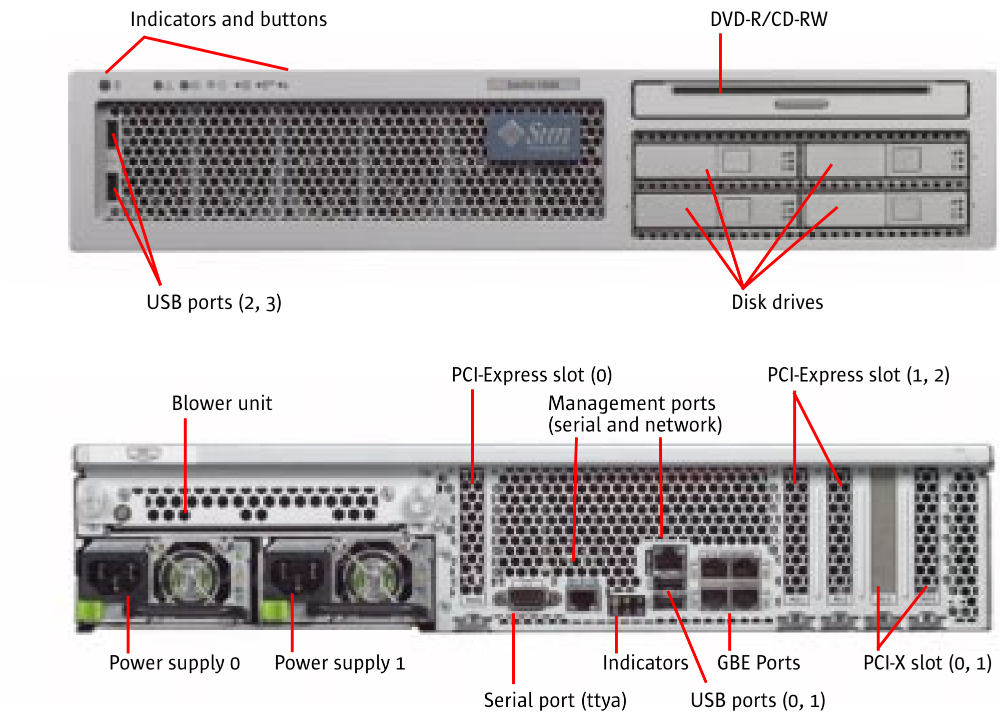


Figure 11. Sun Fire T2000 server, front and rear panels

External features of the Sun Fire T2000 server include:

- Four USB ports are provided, two on the front panel, and two on the rear.
- One slimline DVD-R/CD-RW is accessed through the front panel.
- Up to four hot-pluggable 73 GB SFF SAS disks are supported
- Front and rear indicator lights provide locator (white), service required (amber), and activity status (green) for the system.
- Rear indicator lights on each power supply convey the status of the power supply.
- A single AC plug is provided on each hot-swappable power supply
- Three hot-swappable fans (not shown) can be accessed through a panel on the top of the enclosure. The hot-swappable blower unit is accessed through the back panel, while the system is running.
- Four 10/100/1000Base-T autosensing Ethernet ports are provided
- A DB-9 TTYA serial port is provided for serial devices (not connected to the ALOM system controller serial port)
- Two management ports are provided for use with the ALOM system controller. The RJ-45 serial management port provides the default connection to the ALOM controller. The network management port supports an optional RJ-45/10/100Base-T connection to the ALOM system controller. The serial management port is always available. The network management port is configured through the serial management port.
- Three PCI-E slots and two PCI-X slots are accessed through the rear panel.

Chapter 4

Getting the Most from Throughput Computing: A Complete and Cohesive Software and Management Environment

As the industry has witnessed many times, having a fast processor or even well-designed system is often not sufficient for success. New technology often requires time for tools and applications to arrive, and 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 T1 processor technology, Sun Fire T1000 and T2000 servers provide full binary compatibility with earlier SPARC systems and are delivered ready to run with pre-loaded tools and the solid foundation of the Solaris OS.

The Solaris™ Operating System: Scalability and Support for CoolThreads Technology

The Solaris 10 Operating System is specifically designed to deliver the considerable resources of UltraSPARC T1 processor based systems such as the Sun Fire T1000 and T2000 servers. In fact, the Solaris 10 OS provides new functionality for optimal utilization, relentless availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. The Solaris 10 OS runs on a broad range of SPARC and x86-based systems and compatibility with existing applications is guaranteed.

The Solaris 10 OS provides specific features that facilitate Throughput Computing. One of the most attractive features of systems based on the UltraSPARC T1 processor is that they appear as a familiar SMP system to the Solaris OS and the applications it supports. The Solaris 10 OS has incorporated many features to improve application performance on CMT architectures:

- ***CMT awareness***
  The Solaris 10 OS is aware of the UltraSPARC T1 processor hierarchy so that the scheduler can effectively balance the load across all the available pipelines. Even though it exposes the UltraSPARC T1 processor as 32 logical processors, the Solaris OS understands the correlation between cores and the threads they support.
- ***Fine-granularity manageability***
  The Solaris 10 OS has the ability to enable or disable individual processors. In the case of the UltraSPARC T1 processor, this ability extends to enabling or disabling individual cores and logical processors. In addition, standard Solaris OS features such as processor sets provide the ability to define a group of logical processors and schedule processes or threads on them.
- ***Binding interfaces***
  The Solaris OS allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set, if required or desired.
- ***Solaris Containers***
  Comprised of several key technologies, Solaris Containers provide fine-grained partitioning, virtualization, and allocation of resources within a given Solaris instance. For example, the resources of a single UltraSPARC T1 processor in a Sun Fire T1000 or T2000 server can be easily partitioned into multiple containers, with each securely supporting a separate web or application server.

Solaris Containers Technology

Solaris Containers consist of a group of technologies that work together to efficiently manage system resources, virtualize the environment, and provide a complete, isolated, and secure runtime environment for applications. Solaris containers include important technologies that work together with the fair-share scheduler: Solaris Zones partitioning technology and resource management tools. Solaris Zones enable an administrator to create separate environments for applications on a single system, while the resource management framework allows for the allocation, management, and accounting of system resources such as CPU and memory.

- ***Solaris Zones***
  New to the Solaris 10 Operating System is a unique partitioning technology called Solaris Zones that can be used to create an isolated and secure environment for running applications. A zone is a virtualized operating system environment created within a single instance of the Solaris Operating System. Zones can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability since processes in one zone are prevented from interfering with processes running in another zone.
- ***Resource Management***
  Resource management tools provided with the Solaris Operating System help enable system resources such as CPU resources to be dedicated to specific applications. CPUs in a multiprocessor system can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to a Solaris zone. Resource pools provide the capability to separate workloads so that consumption of CPU resources do not overlap, and 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

With the Solaris 10 OS, Sun 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 the Solaris 10 OS 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 mis-configuration problems.

- ***Solaris Fault Manager***
  The Solaris Fault Manager facility collects data relating to hardware and software errors. It automatically and silently detects and diagnoses the underlying problem, with an extensible set of agents that automatically respond by taking the faulty component offline. Easy-to-understand diagnostic messages link to articles in Sun's knowledge base to clearly guide administrators through corrective tasks that require human intervention. The open design of the Solaris Fault Manager facility also permits administrators and field personnel to observe the activities of the diagnostic system. With Solaris Fault Manager, the overall time from a fault condition, to automated diagnosis, to any necessary human intervention is greatly reduced, increasing application uptime.

- ***Solaris Service Manager***
  The Solaris Service Manager facility creates a standardized control mechanism for application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services can then be automatically restarted if they are accidentally terminated by an administrator, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem. In addition, the Solaris Service Manager software 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 Solaris Service Manager is also simple to deploy; developers can convert most existing applications to take full advantage of Solaris Service Manager features by simply adding a simple XML file to each application.

Predictive self healing and fault management provide the following specific capabilities on Sun Fire T1000 and T2000 servers:

- ***CPU Offlining*** takes a core offline that has 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 (Note that CPU offlining on the Sun Fire T1000 and T2000 server will be delivered after their initial release).
- ***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.

The Java Enterprise System (JES)

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

Suns Java Enterprise System (JES) provides a complete set of infrastructure software that is integrated to work as a whole, and that offers shared components, common technologies, a consistent architecture and user experience. Using world-class software, Sun redefines the software system from the operating system up through the J2EE specification layer. Customers can write their business applications to Java software standards, leverage Java Enterprise System network services, and Sun delivers the end-to-end solution to run them.

Bundled with the Sun Fire T1000 and T2000 servers, and now available at no cost, the Java Enterprise System 2005Q1 release includes the components listed in Table 8.

Table 8. Java Enterprise System categories and components

Category	Components
Network Identity Services	Sun Java System Directory Proxy Server 5 2005Q1 Sun Java System Directory Server 5 2005Q1 Sun Java System Access Manager 6 2005Q1
Web and Application Services	Sun Java System Application Server Enterprise Edition 8.1 2005Q1 Sun Java System Message Queue 3 Enterprise Edition 2005Q1 Sun Java System Web Server 6.1 SP4 2005Q1
Portal Services	Sun Java System Portal Server 6 2005Q1 Sun Java System Portal Server Mobile Access 6 2005Q1 Sun Java System Portal Server Secure Remote Access 6 2005Q1
Communication and Collaboration Services	Sun Java System Calendar Server 6 2005Q1 Sun Java System Instant Messaging 7 2005Q1 Sun Java System Messaging Server 6 2005Q1
Availability Services[a]	Sun Cluster 3.1 9/04 Sun Cluster Agents for JES Services (Web, application, directory, messaging, and calendar servers and message queues)

a. As of this writing, Sun Cluster software is not supported on the Sun Fire T1000 and T2000 servers

Sun Studio 11 Compilers and Analysis Tools

Sun Studio 11 software is provided on each Sun Fire T1000 and T2000 server, offering the latest release of record-setting optimizing compilers and tools for the C, C++, and FORTRAN developer. These compilers deliver the highest optimizations and the best performance in the development of 32-bit and 64-bit applications on Sun's newest hardware platforms, including Sun Fire T1000 and T2000 servers as well as the latest multi-core x64 and x86 platforms. Sun Studio now removes the price barrier and is available at no cost.

Sun Studio software compilers allow developers to leverage the latest in parallel programming and maximize throughput on multi-core and multi-threaded systems. In addition, even single-threaded applications gain as the compilers can identify parallelization opportunities—generating back-end code that automatically parallelizes execution without source-code changes. Integrated into a fully-featured IDE, Sun Studio 11 software includes an advanced graphical debugger and a suite of performance analysis tools that simplify the development process for even the most sophisticated programming needs.

- ***Acceleration for multi-core, multi-processor, and multithreaded technology***
  Sun Studio software can help achieve higher system throughput with multithreaded applications. Applications can be built using the OpenMP v2.5 application interface (API) with C, C++, and FORTRAN code, along with an improved debugger and performance analysis tools. The OpenMP implementation supports nested parallelism, where nested regions can be executed by a team of two or more threads. Paired together with multi-core and CMT optimizations, applications can maximize performance by taking advantage of the latest UltraSPARC T1, x86, and x64 based systems.
- ***Highest performance on Sun platforms***
  Optimizing compilers produce record-setting runtime performance that consistently exceeds performance from open-source alternatives and prior Sun Studio software releases.

- ***Simple debugging***
  An intuitive graphical user interface makes it easy to access advanced debugging features. Breakpoints can be set, variables examined, and the call stack navigated—all via the debugger's convenient menus and buttons. The unique *Fix and Continue* feature lets developers slash turnaround time for fixes and achieve greater debugging productivity. Developers can even debug multi-threaded OpenMP code as well as seamlessly debug code in mixed languages, including C, C++, FORTRAN, and Java.
- ***Sophisticated performance analysis tools***
  Developers can assess the performance of their applications, from algorithm changes to hardware system counters, with one tool that correlates source code to actual machine execution. Also, new to this release, dataspace profiling is available on UltraSPARC systems to provide unique views into the performance costs associated with application memory references.
- ***Simple migration path***
  Source- and object-level compatibility with prior Sun Studio releases as well as GNU C/C++ compatibility features simplify upgrade and adoption.
- ***Free license model***
  Sun Studio 11 introduces a new, no-charge licensing model with for-fee Sun Developer Support Plans. Standard S4 support is available at launch with new Sun Developer Support Plans being introduced soon.

System Management

As the number of systems grow in any organization, the complexities of managing the infrastructure through its lifecycle becomes increasingly difficult. Effective system management requires both integrated hardware that can sense and modify the behavior of key system elements as well as advanced tools that can automate key administrative tasks.

Advanced Lights-Out Management (ALOM) System Controller

The integral Advanced Lights Out Manager (ALOM) system controller allows Sun Fire T1000 and T2000 servers to be remotely managed and administered. The ALOM software comes pre-installed on the Sun Fire T1000 server so ALOM works as soon the administrator installs and applies power to the server. ALOM can then be customized to work with a particular installation.

ALOM allows the administrator to monitor and control a server, either over a network or by using a dedicated serial port for connection to a terminal or terminal server. ALOM provides a command-line interface that can be used to remotely administer geographically-distributed or physically-inaccessible machines. In addition, ALOM allows administrators to run diagnostics remotely (such as power-on self-test) that would otherwise require physical proximity to the server serial port. ALOM can also be configured to send email alerts of hardware failures, hardware warnings, and other events related to the server or to ALOM.

The ALOM circuitry runs independently of the server, using the server's standby power. As a result, ALOM firmware and software continue to function when the server operating system goes offline or when the server is powered off. ALOM monitors disk drives, fans, CPUs, power supplies, system enclosure temperature, voltages, and the server front panel, so that the administrator does not have to.

ALOM specifically monitors the following Sun Fire T1000 and T2000 server components:

- CPU temperature conditions
- Enclosure thermal conditions
- Fan speed and status
- Power supply status
- Voltage thresholds

Sun Management Center Software

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

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

Sun Management Center software version 3.6 supports the Sun Fire T1000 and T2000 servers.

Sun N1™ System Manager

The Sun N1™ System Manager is infrastructure lifecycle management software for deploying, monitoring, patching, and managing large and small installations of Sun systems. Sun N1 System Manager takes a step-by-step approach to unraveling the challenges of getting systems operational quickly:

- ***Discover***
  As systems are added to the management network, administrators can use Sun N1 System Manager to discover bare metal systems based on a given subnet address or IP range.
- ***Group***
  Given the number of systems to manage and the constant re-purposing of systems, it is critical for IT organizations to find ways to group resources together. Sun N1 System Manager enables users to logically group systems together and perform actions across a group of systems as easily as performing actions on a single system. Systems can be grouped by function (web servers versus grid computing), administrative responsibility, or other categorization based on organizational needs.
- ***Provision***
  Sun N1 System Manager remotely installs operating systems (Solaris OS, RedHat, or SuSE Linux) onto selected systems. Administrators can use this functionality to provision operating systems onto bare metal systems or

reprovision existing systems. As the infrastructure life cycle continues, Sun N1 System Manager can update firmware and provision software packages and patches to selected systems.

- ***Monitor***
  When systems are up and running, administrators can use Sun N1 System Manager to monitor system health, helping to ensure that everything is running at the optimal levels. The software provides detailed hardware monitoring for attributes such as fans, temperature, disk, and voltage usage, including bare metal systems. Sun N1 System Manager also monitors OS attributes such as swap space, CPU, memory, and file systems. Administrators can define specific threshold levels and set preferred notification methods, including e-mail, pager, or Simple Network Management Protocol (SNMP) traps, for each monitored component as business needs demand.
- ***Manage***
  Businesses require that infrastructure life cycle management extend beyond just deploying and monitoring systems. Sun N1 System Manager 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 Sun N1 System Manager software's Role-Based Access Control (RBAC) feature, organizations can grant permissions to specific users to perform specific management tasks.
- ***Hybrid User Interface***
  Sun N1 System Manager offers users a hybrid user interface (UI), accessible from the Web, that integrates both the GUI and CLI into one console. With this hybrid UI, operations performed in the GUI are simultaneously reflected in the CLI, and vice versa.

The Sun N1 System Manager, version 1.2 provides support for the Sun Fire T1000 and T2000 servers.

Chapter 5
Conclusion

Delivering the benefits of Throughput Computing in the Participation Age will require a comprehensive approach that includes innovative processors, system platforms, operating systems, along with application, middleware, and management technology. With its strong technology positions and R&D investments in all of these areas, Sun is in a unique position to deliver on this vision. Far from futuristic, Sun has effective solutions today that can help organizations cope with the need for performance and capacity while effectively managing space, power and heat.

Powered by the breakthrough UltraSPARC T1 processor, Sun Fire servers with CoolThreads technology represent a powerful new approach that delivers virtually unprecedented levels of performance while fundamentally changing the equation on power and cooling. The result is data center infrastructure that can scale to meet new challenges, even as it allows the organization to act responsibly toward the environment and the bottom line.

As the first systems to fully implement Sun's Throughput Computing vision, Sun Fire T1000 and T2000 servers provide the resources for considerable web, application, and database consolidation while preserving customers investments in SPARC/Solaris technology. With innovations such as Solaris Containers and Java technology, customers can move to adopt this radical new technology without disrupting their ongoing operations. Innovative Java Enterprise System middleware and Sun N1 system management software round out these offerings, providing a complete and compelling solution.