An Oracle White Paper
February 2014

Oracle's SPARC T5-2, SPARC T5-4, SPARC T5-8, and SPARC T5-1B Server Architecture

Introduction

Employing Oracle's new SPARC T5 processor, Oracle's SPARC T5-2, SPARC T5-4, SPARC T5-8 servers and SPARC T5-1B server module offer breakthrough performance and energy efficiency to help simplify data center infrastructures and address other demanding challenges. New levels of performance and scalability across a variety of workloads mean that these versatile systems can deliver a virtualized infrastructure for the entire enterprise while also enabling IT managers to deploy and manage fewer types of platforms and fewer servers.

The SPARC T5 processor takes the industry's very successful SPARC T4 processor to the next level by doubling the core count and providing the improved power management necessary for cloud infrastructures. Sixth-generation multicore, multithreading technology supports up to 128 threads in as little as two rack units (2RU), providing increased computational density while staying within constrained envelopes for power and cooling. Very high levels of integration help reduce latency, lower costs, and improve security and reliability. The optimized system design provides support for all enterprise services and application types. Uniformity of management interfaces and adoption of standards also help reduce administrative costs, while an innovative chassis design shared across Oracle's T-Series servers provides density, efficiency, and economy for modern data centers.

Figure 1. SPARC T5-1B, SPARC T5-2, SPARC T5-4, and SPARC T5-8 servers.

Comparison of SPARC T5–Based Server Features

Table 1 compares the SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module.

TABLE 1. SPARC SERVER FEATURES

FEATURE	SPARC T5-2 SERVER	SPARC T5-4 SERVER	SPARC T5-8 SERVER	SPARC T5-1B SERVER MODULE
CPUS	• 16-core 3.6 GHz SPARC T5 processor (single or dual)	• 16-core 3.6 GHz SPARC T5 processor (quad)	• 16-core 3.6 GHz SPARC T5 processor (eight)	• 16-core 3.6 GHz SPARC T5 processor
THREADS	• 128 (single) • 256 (dual)	• 512 (quad)	• 1024 (eight)	• 128
MEMORY CAPACITY	• 128 GB (single) or 256 GB (dual) (8 GB DDR3 dual inline memory modules [DIMMs]) • 256 GB (single) or 512 GB (dual) (16 GB DDR3 DIMMs) • 512 GB (single) or 1 TB (dual) (32 GB DDR3 DIMMs)	• 1 TB (16 GB DDR3 DIMMs) • 2 TB (32 GB DDR3 DIMMs)	• 2 TB (16 GB DDR3 DIMMs) • 4 TB (32 GB DDR3 DIMMs)	• 128 GB (8 GB DDR3 DIMMs) • 256 GB (16 GB DDR3 DIMMs) • 512 GB (32 GB DDR3 DIMMs)
MAXIMUM INTERNAL DISK DRIVES	• Up to 6 HDDs (2.5-inch SAS3 or SSD 300/600 GB disk drives) • RAID 0/1/1E	• Up to 8 HDDs (2.5-inch SAS3 or SSD 300/600 GB disk drives) • RAID 0/1/1E	• Up to 8 HDDs (2.5-inch SAS3 or SSD 300/600 GB disk drives) • RAID 0/1/1E	• Up to 2 HDDs (2.5-inch SAS3 or SSD 300/600 GB disk drives) • RAID 0/1
VIDEO	• One HD-15 VGA port	• One HD-15 VGA port	• One HD-15 VGA port	• One HD-15 VGA port (dongle)
REMOVABLE, PLUGGABLE I/O	• Slimline DVD+R/-W • Five USB 2.0 ports	• Slimline DVD+R/-W • Five USB 2.0 ports	• No DVD (accessed via rKVMS) • Four USB 2.0 ports	• No DVD (accessed via rKVMS) • Three USB 2.0 ports
PCI	• Eight x8 PCIe Gen3 slots	• 16 hot-plug x8 PCIe Gen3 slots	• 16 hot-plug x8 PCIe Gen3 slots	• Optional Fabric Expansion Module[1] • Two EM x8 PCIe Gen2 slots

ETHERNET	• Four onboard 10 GbE ports	• Four onboard 10 GbE ports	• Four onboard 10 GbE ports	• Two onboard GbE ports (10/100/1000) [3]
POWER SUPPLIES	• Two hot-swappable AC 2000 W power supplies • N+1 redundancy	• Two hot-swappable AC 3000 W power supplies • N+1 redundancy	• Four hot-swappable AC 3000 W power supplies • N+N redundancy	• Contained within Oracle's Sun Blade 6000 chassis
FANS	• Six hot-swappable fan modules, with counterrotating fans per module • N+1 redundancy	• Six hot-swappable fan trays, with counterrotating fans per module • N+1 redundancy	• Five hot-swappable fan modules, with counterrotating fans per module • N+1 redundancy	• Contained within Sun Blade 6000 chassis
OPERATING SYSTEM	• Oracle Solaris 11.1, Oracle Solaris 10 1/13, Oracle Solaris 10 8/11 + Oracle Solaris 10 1/13 Patch Bundle, Oracle Solaris 10 9/10 + Oracle Solaris 10 1/13 Patch Bundle[4]	• Oracle Solaris 11.1, Oracle Solaris 10 1/13, Oracle Solaris 10 8/11 + Oracle Solaris 10 1/13 Patch Bundle, Oracle Solaris 10 9/10 + Oracle Solaris 10 1/13 Patch Bundle[4]	• Oracle Solaris 11.1, Oracle Solaris 10 1/13, Oracle Solaris 8/11 + Oracle Solaris 10 1/13 Patch Bundle, Oracle Solaris 10 9/10 + Oracle Solaris 10 1/13 Patch Bundle[4]	• Oracle Solaris 11.1, Oracle Solaris 10 1/13, Oracle Solaris 10 8/11 + Oracle Solaris 10 1/13 Patch Bundle, Oracle Solaris 10 9/10 + Oracle Solaris 10 1/13 Patch Bundle[4]

[2] To connect to optional Network Expansion Modules

[3] And appropriate Network Expansion Module

[4] Solaris 11.1 preloaded at factory

SPARC T5 Processor

The SPARC T5 processor is the industry's most highly integrated system-on-a-chip, supplying the most high-performance threads of any multicore processor available and integrating all key system functions. It is the first sixteen-core threaded SPARC system-on-a-chip to achieve high single-threaded performance.

The SPARC T5 processor, shown in Figure 2, eliminates the need for expensive custom hardware and software development by integrating computing, security, and I/O onto a single chip. Achieving binary compatibility with earlier SPARC processors, no other processor delivers so much performance in so little space and with such low power requirements. This processor enables organizations to rapidly scale the delivery of new network services with maximum efficiency and predictability.

Figure 2. The SPARC T5 processor allows organizations to rapidly scale the delivery of new network services and compute-intensive workloads with maximum efficiency and predictability.

Table 2 provides a comparison between the SPARC T5, SPARC T4, SPARC T3 processors. The SPARC T5 leverages many of the elements from the SPARC T4 processor.

TABLE 2. SPARC T5, SPARC T4, AND SPARC T3 PROCESSOR FEATURE COMPARISON

FEATURE	SPARC T5 PROCESSOR	SPARC T4 PROCESSOR	SPARC T3 PROCESSOR
CPU FREQUENCY	3.6 GHz	• 2.85/3.0 GHz	• 1.65 GHz
OUT-OF-ORDER EXEC. DUAL INSTR. ISSUE DATA/INSTR. PREFETCH LEVEL 3 CACHE	• Yes • Yes • Yes • Yes	• Yes • Yes • Yes • Yes	• No • No • No • No
THREADS/CORE CORES/PROCESSOR THREADS/PROCESSOR HYPERVISOR	• 8 • 16 • 128 • Yes	• 8 • 8 • 64 • Yes	• 8 • Up to 16 • 128 • Yes
SOCKETS SUPPORTED	• One, two, four, or eight	• One, two, or four	One, two, or four
MEMORY	• Four memory controllers • Up to 16 DDR3 DIMMs	• Two memory controllers • Up to 16 DDR3 DIMMs	• Two memory controllers • Up to 16 DDR3 DIMMs
CACHES	• 16-KB instruction cache • 16-KB data cache • 128-KB level 2 (L2) cache • 8-MB level 3 (L3) cache (8 banks, 16-way, set associative)	• 16-KB instruction cache • 16-KB data cache • 128-KB L2 cache • 4-MB L3 cache (8 banks, 16-way, set associative)	• 16-KB instruction cache • 8-KB data cache • 6-MB L2 cache (16 banks, 24-way associative)
TECHNOLOGY	• 28 nm technology	• 40 nm technology	• 40 nm technology

FLOATING POINT	• 1 FPU with Mul/Add per core • 16 FPUs per chip	• 1 FPU with Mul/Add per core • 8 FPUs per chip	• 1 FPU with Mul/Add per core • 16 FPUs per chip*
INTEGER RESOURCES	• Two integer execution units/core	• Two integer execution units/core	• Two integer execution units/core
CRYPTOGRAPHY	• Stream processing unit/core, integrated within pipeline • 14 most-popular ciphers	• Stream processing unit/core, integrated within pipeline • 14 most-popular ciphers	• Stream processing unit/core • 12 most-popular ciphers
ADDITIONAL ON-CHIP RESOURCES	• Dual PCIe Gen3 interface (x8) • Coherency switch (7 x 153.6 Gb/sec)	• Dual PCIe Gen2 interface (x8) • Dual 10 GbE XAUI interfaces (x8) • Coherency logic and links (6 x 9.6 Gb/sec)	• Dual 10 GbE Gen2 interfaces (x8) • Dual 10 GbE PCIe XAUI interface (x8) • Coherency logic and links (6 x 9.6 Gb/sec)

* Two-socket implementation represents SPARC T5-2 server, whereas SPARC T5-4 server represents a four-socket implementation.

Taking Oracle's Multicore/Multithreaded Design to the Next Level

When designing the next-generation of Oracle's multicore/multithreaded processors, the in-house design team started with the following key goals in mind:

- Improve the single-threaded computational capabilities over that of the SPARC T4 processor for workloads that require this level of performance.
- Maintain the computational capabilities to meet the growing demand from Web applications by providing double the throughput of the SPARC T4 processor.
- Support larger and more-diverse workloads with greater floating-point performance.
- Provide networking performance equivalent to the SPARC T4 CPU to serve network-intensive workloads.
- Provide end-to-end data center encryption with significantly higher performance as well as adding new ciphers implemented within hardware.
- Increase service levels and reduce planned and unplanned downtime.
- Improve data center capacities while reducing costs.

The SPARC T5 processor design recognizes that memory latency is truly the bottleneck to improving performance. By redesigning the cores within each processor, designing a new floating-point pipeline, and further increasing network bandwidth, this processor is able to provide approximately 30 percent higher single-threaded throughput than the SPARC T4 processor.

Each SPARC T5 processor provides sixteen cores, with each core able to switch between up to eight threads (128 threads per processor) using a modified LRU (Least Recently Used) algorithm for thread choice. In addition, each core provides two integer execution pipelines, so a single SPARC core is capable of executing two threads at a time. Unlike the SPARC T3 processor, the SPARC T5 processor fetches one of eight threads for instruction propagation through stages of the pipeline to present to the select stage by the fetch3 stage. Thread instructions are grouped into two-instruction decode groups and proceed through the decode, rename, and pick stages before proceeding to the issue stage, after which they are sent to one of four subsequent execution pipelines, depending upon the type of instruction to be performed.

Up to this point, each instruction from any thread has proceeded through the pipeline independent of the type of instruction. Two instructions are issued for execution by the issue stage per cycle, unlike with the SPARC T3 processor, which issued only one instruction per cycle

Figure 3 provides a simplified high-level illustration of the thread model supported by a 16-core SPARC T5 processor.

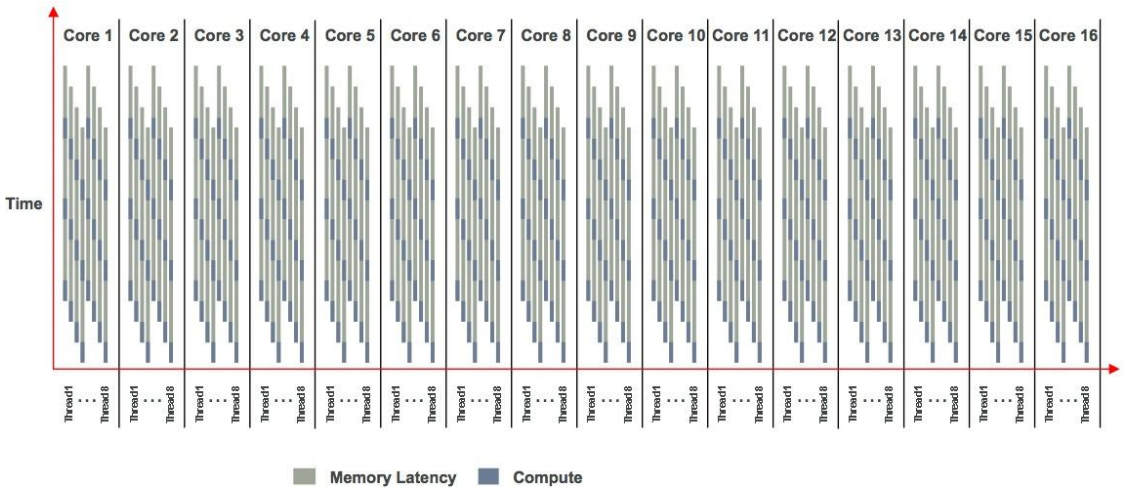

Figure 3. A single 16-core SPARC T5 processor supports up to 128 threads, with up to two threads running in each core simultaneously.

SPARC T5 Processor Architecture

The SPARC T5 processor extends Oracle's multicore/multithreaded initiative with an elegant and robust architecture that delivers real performance to a wide range of applications. From OLTP transactional workloads, to data warehousing, as well as single-thread sensitive batch operations. Figure 4 provides a block-level diagram of the SPARC T5 processor.

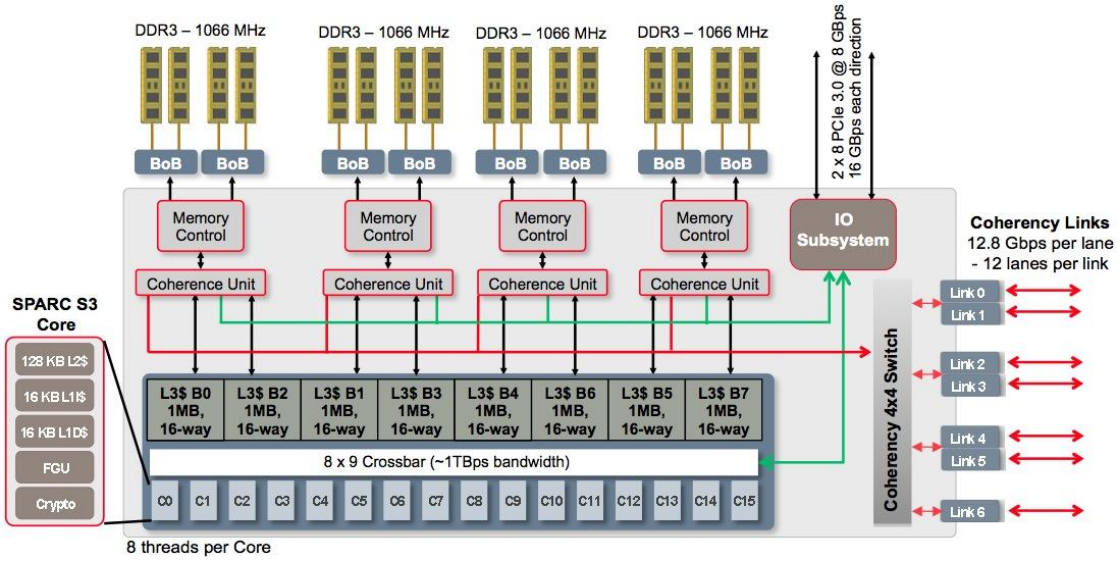

Figure 4. The SPARC T5 processor provides seven coherence links to connect to up to four other processors.

The SPARC T5 processor is a single chip multiprocessor (CMP) and contains 16 physical processor cores. Each physical processor core has full hardware support for eight strands, two integer execution pipelines, one floating-point execution pipeline, and one memory pipeline.

In addition to outstanding multithreaded performance, the SPARC T5 processor offers greatly improved single-thread performance. In particular, the SPARC T5 processor provides a robust out-of-order, dual-issue processor core that is heavily threaded among eight strands. It has a 16-stage integer pipeline to achieve high operating frequencies, advanced branch prediction to mitigate the effect of a deep pipeline, and dynamic allocation of processor resources to threads. This allows the SPARC T5 processor to achieve very high single-thread performance (about 30 percent higher than the SPARC T4 processor), while still scaling to very high levels of throughput.

Each physical core has a 16-KB, 4-way associative instruction cache (32B lines), a 16 KB, 4-way associative data cache (32B lines), a 64-entry fully-associative instruction translation lookaside buffer (TLB), and a 128-entry fully associative data TLB that are shared by the eight strands. It also includes a private, 128-KB, 8-way inclusive write-back L2 cache with 32B lines. Each physical core also includes cryptographic acceleration hardware, accessible via user-level instructions.

The SPARC T5 processor has coherence link interfaces to allow communication between up to eight SPARC T5 chips in a system without requiring any external hub chip. There are seven coherence links, each with 12 lanes in each direction running at 153.6 Gb/sec. The SPARC T5 processor has seven coherence link units (CLUs), four coherence units, and a cross bar (CLX) between coherence units and CLUs.

The SPARC T5 processor interfaces to external DDR3 DIMMs via an external buffer-on-board (BoB) chip using proprietary unidirectional high-speed links. There are four memory links on the SPARC T5 processor. Each memory link is 12 lanes southbound and 12 lanes northbound and operates at 12.8 Gb/sec. Each BoB chip has a DDR3 channel for a total of up to eight DDR3 channels per SPARC T5 processor. Each DDR3 channel has two DIMMs providing up to 16 DDR3 DIMMs per SPARC T5 processor.

The SPARC T5 processor can support one-, two-, four-, and eight-socket implementations. A typical two-socket implementation is shown in Figure 5. Dual-socket, as well quad-socket SPARC T5 implementations interconnect the processors' seven coherence links. No additional circuitry is required.

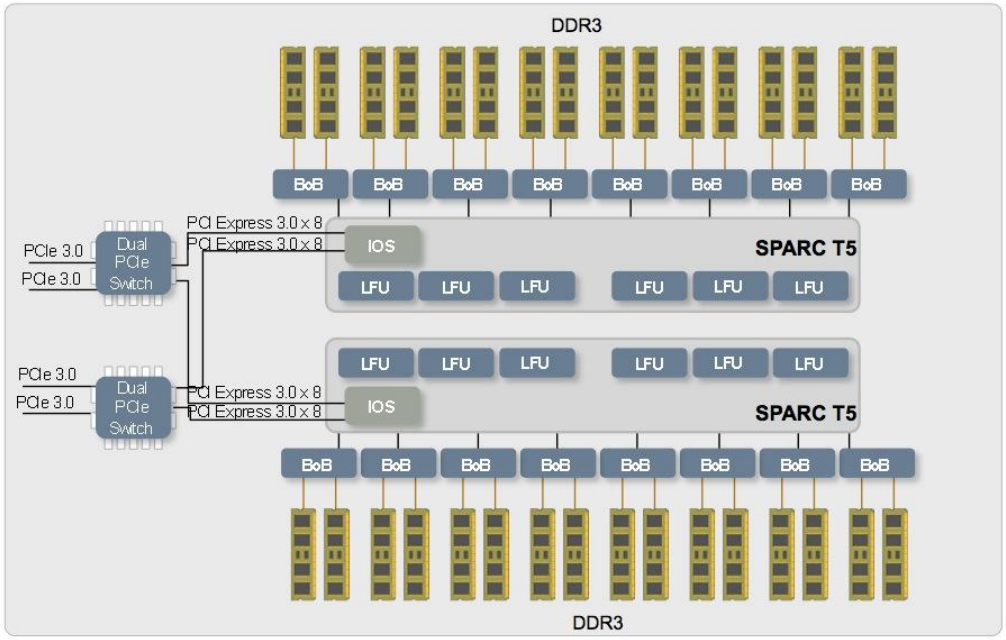

Figure 5. A typical dual-socket SPARC T5 configuration.

SPARC T5 Processor Cache Architecture

The SPARC T5 processor has a three-level cache architecture. Levels 1 and 2 are specific to each core, that is, these two levels of cache are not shared with other cores. Level 3 is shared across all cores of a given processor. Cache sharing does not occur across another processor even though that processor may be in the same physical system. The SPARC T5 processor has Level 1 caches that consist of separate data and instruction caches. Both are 16 KB and are per core. A single Level 2 (L2) cache, again per core, is 128 KB. The Level 3 (L3) cache is *shared* across all sixteen cores of the SPARC T5 processor and is 8 MB, has eight banks, and is 16-way set associative. Figure 6 illustrates the relationship between the L2 and Level 3 caches and shows them connected by a 8x9 crossbar:

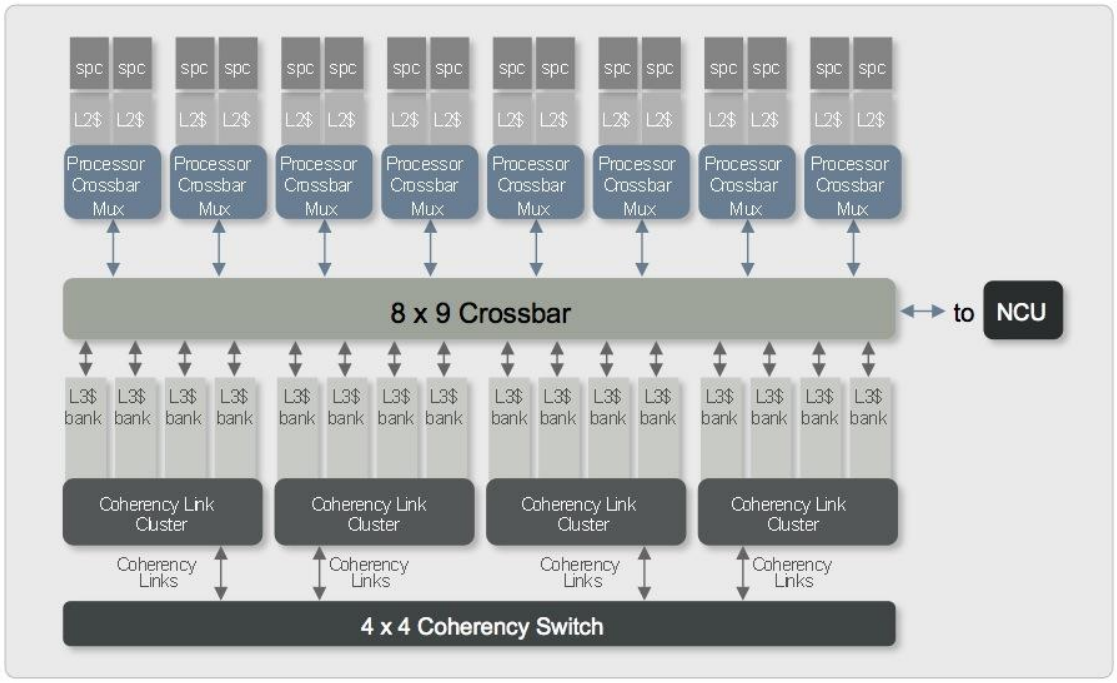

Figure 6. The relationship between the L2 and L3 caches.

SPARC T5 Core Architecture

The SPARC T5 processor represents a fundamental redesign of the core within a SPARC multicore architecture and a continuation of the previous SPARC T4 processor. Now included within the core are the following aspects that are more conventionally associated with superscalar designs:

- Out-of-Order (OoO) instruction execution
- Sophisticated branch prediction
- Prefetching of both instructions and data
- Much deeper pipelines (relative to previous versions of multicore processors from Sun/Oracle)
- Three levels of cache
- Support for a much larger memory management unit (MMU)page size (2 GB)
- Multiple instruction issue

There are many functional units, pipelines, and associated details that are present within the SPARC T5 core but are beyond the scope of this paper. However, due to the significantly new characteristics and features of the SPARC T5 core, this paper does attempt to touch upon the major exposed features or characteristics (that is, those that are visible to either programmers or users of SPARC T5–based systems).

One aspect by which the designers of the SPARC T5 architecture were able to achieve a physical space savings of chip real estate was to reuse many physical pieces of a given core for widely varying functionality. For example, for each of the four major pipelines present within each core, the first 14 stages of each pipeline are actually shared. This represents a major space utilization efficiency by making each of the first 14 stages identical. Thus, they can be used by one of two integer instructions, a floating-point graphics instruction, or a load-store instruction. In Figure 7, the first six blocks represent the 14 identical stages, which are specifically defined in Figure 8.

Dynamic Threading

The SPARC T5 processor is dynamically threaded. While software can activate up to eight strands on each core at a time, hardware dynamically and seamlessly allocates core resources such as instruction, data, and L2 caches and TLBs, as well as out-of-order execution resources such as the 128-entry reorder buffer in the core. These resources are allocated among the active strands. Software activates strands by sending an interrupt to a halted strand. Software deactivates strands by executing a HALT instruction on each strand that is to be deactivated. No strand has special hardware characteristics. All strands have identical hardware capabilities.

Since the core dynamically allocates resources among the active strands, there is no explicit single-thread mode or multithread mode for software to activate or deactivate. If software effectively halts all strands except one on a core via critical thread optimization (described later in this document), the core devotes all its resources to the sole running strand. Thus, that strand will run as quickly as possible. Similarly, if software declares six out of eight strands as noncritical, the two active strands share the core execution resources.

The extent to which strands compete for core resources depends upon their execution characteristics. These characteristics include cache and TLB footprints, inter-instruction dependencies in their execution streams, branch prediction effectiveness, and others. Consider one process that has a small cache footprint and a high correct branch prediction rate such that when running alone on a core, it achieves two instructions per cycle (the SPARC T5 processor's peak rate of instruction execution). This is termed a high IPC process. If another process with similar characteristics is activated on a different strand on the same core, each of the strands will likely operate at approximately one instruction per cycle. In other words, the single-thread performance of each process has been cut in half. As a rule of thumb, activating N high-IPC strands will result in each strand executing at 1/N of its peak rate, assuming each strand is capable of executing close to two instructions per cycle.

Now consider a process that is largely memory-bound. Its native IPC will be small, possibly 0.2. If this process runs on one strand on a core with another clone process running on a different strand, there is a good chance that both strands will suffer no noticeable performance loss, and the core throughput will improve to 0.4 IPC. If a low-IPC process runs on one strand with a high-IPC process running on another strand, it's likely that the IPC of either strand will not be greatly perturbed. The high-IPC strand might suffer a slight performance degradation (as long as the low-IPC strand does not cause a substantial increase in cache or TLB miss rates for the high-IPC strand).

The guidelines above are only general rules-of-thumb. The extent to which one strand affects another strand's performance depends upon many factors. Processes that run fine on their own, but suffer from destructive cache or TLB interference when run with other strands, might suffer unacceptable performance losses. Similarly, it is also possible for strands to cooperatively improve performance when run together. This might occur when the strands running on one core share code or data. In this case, one strand may prefetch instructions or data that other strands will use in the near future.

The same discussion can apply between cores running in the chip. Since the L3 cache and memory controllers are shared between the cores, activity on one core can influence the performance of strands on another core.

Figure 7 is a block-level diagram representing a single SPARC core on the SPARC T5 processor. Sixteen cores are supported per processor.

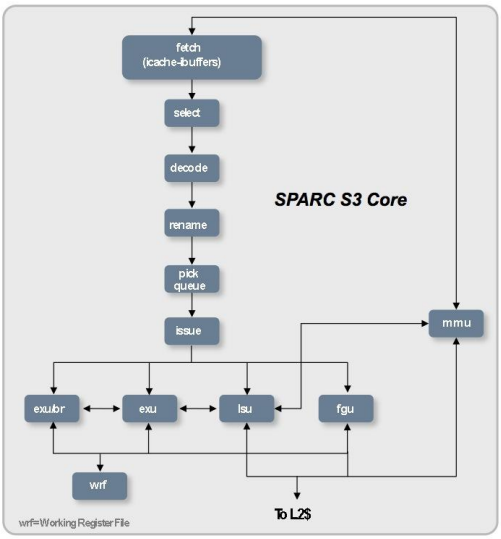

Figure 7. Block-level diagram of a single core of the SPARC T5 processor.

Components implemented in each core include the following:

- **Trap logic unit (not shown).** The trap logic unit (TLU) updates the machine state as well as handling exceptions and interrupts.
- **Instruction fetch unit.** The instruction fetch function is responsible for selecting the thread, fetching instructions from instruction cache (icache) for the selected thread, and providing up to four instructions to the select stage every cycle. On the SPARC T5 processor, it performs the following major functions:
  - Select the thread to be fetched.
  - Fetch instructions from icache for the selected thread, and place them in the instruction buffers for the select unit.

  - Predict direction and target of delayed control transfer instructions (DCTI) on the thread being fetched.
  - On icache miss, fetch data from the L2 cache (L2$), pre-decode it, and store it in icache.
- **Select unit.** The primary responsibility of the select unit is to schedule a thread for execution on the SPARC T5 processor's pipeline for each cycle. For each cycle, up to one thread out of eight threads total can be selected for execution. A thread is in one of two states: Ready or Wait. Threads can be in a Wait state due to postsync conditions, mispredicted branches, lack of valid instructions, or other instruction-related wait conditions. For each cycle, the select unit selects one thread for execution from among the ready threads using a least-recently-used (LRU) algorithm for fairness. For the selected thread, up to two instructions are sent to the decode unit per cycle.
- **Decode unit.** The decode unit on the SPARC T5 processor is responsible for the following:
  - Identifying illegal instructions
  - Decoding integer and FP sources and sinks for up to two instructions per cycle as well as detecting source/sink dependencies
  - Generating flat mapping of integer and FP registers
  - Decoding condition-code sources and destinations
  - Generating micro-ops for complex instructions
  - Generating instruction slot assignments
  - Detecting DCTI (delayed control transfer instruction) couples
  - Creating NOOPs when exceptions or annulling are detected
  - Maintaining speculative copies of window registers and executing certain window register instructions
  - Decoding up to two instructions every cycle
  - Preparing the data and the addressing for the Logical Map Tables (LMTs), which are part of the rename unit (RU)
- **Rename unit.** The rename unit is responsible for renaming the destinations of instructions and resolving destination-source dependencies between instructions within a thread as well as creating age vector dependency based on issue-slot. Renaming takes three cycles: R1, R2, and R3. For each cycle, the rename unit gets up to two instructions from the decode unit at the end of the D2 cycle. Each group of instructions is called a decode group. The rename unit does not break the decode group of instructions received from decode unit.
- **Pick unit**. The pick unit schedules up to three instructions per cycle out of a 40-entry pick queue (PQ). Up to three instructions (two instructions plus one store data acquisition op) are written into the PQ during the second phase of the R3. The PQ is read during the first phase of the pick cycle.

- **Issue unit.** The primary responsibility of the issue unit is to provide instruction sources and data to the execution units. The SPARC T5 processor has six execution units corresponding to the three issue slots, as shown in Figure 8.

Issue Slot	*Unit*
0	Load/Store Unit Integer Execution Unit 0
1	Integer Execution Unit 1 Branch Unit FGU SPU
2	Store data operation

Figure 8. Relationship between issue slots and execution units.

- **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. A fused floating point Mul/Add instruction is implemented. In addition, the integer Fused Mul/Add instruction from the SPARC64 VII instruction set has been added. This also performs part of the cryptographic calculations based upon the algorithm being executed.

  A newly designed core for the SPARC T5 processor implements a 16-stage integer pipeline, a 20-stage load-store pipeline, and a 27-stage floating-point graphics pipeline. All are present in each of the sixteen cores of a SPARC T5 processor (Figure 9).

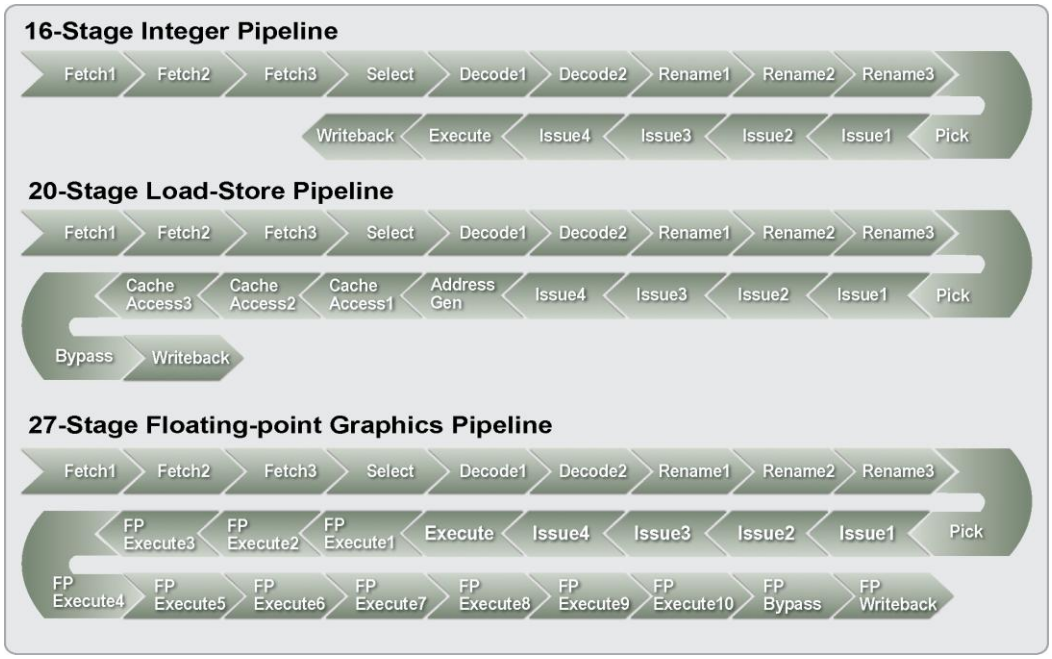

Figure 9. A 16-stage integer pipeline, a 20-stage load-store pipeline, and a 27-stage floating-point graphics pipeline are provided by each SPARC T5 processor core.

- **Stream processing unit.** Each core contains a stream processing unit (SPU) that provides cryptographic processing. This functionality has been implemented within the core pipelines in the SPARC T5 processor and is accessible by 29 new user-level instructions.
- **Load-store unit.** The load-store unit (LSU) is responsible for processing all memory reference instructions and properly ordering all memory references. The LSU will receive load and store instructions out of order as they are picked by the pick unit. Loads might be issued out of order with respect to other loads and stores might be issued out of order with the respect to other loads and stores. However, loads will not be issued ahead of previous stores. In addition to the memory references required by the instruction set, the LSU also contains a hardware prefetcher, which prefetches data into the L1 cache based upon detected access patterns.
- **Memory management unit.** The memory management unit (MMU) provides a hardware table walk (HWTW) and supports 8-KB, 64-KB, 4-MB, 256-MB, and 2-GB pages.
- **Integer execution unit.** The integer execution unit (EXU) is capable of executing up to two instructions per cycle. Single-cycle integer instructions are executed in either the EXU0 (slot0) or EXU1 (slot1) pipeline. Load and store address operations go to EXU0 (slot0). Branch instructions are executed in EXU1 (slot1). Floating point, multicycle integer, and SPU instructions go through the EXU1 (slot1) pipeline. Store data operations go to EXU0 (slot2), but are not considered separate instructions by the EXU since the store address operation must also occur for the same instruction.

  To illustrate how the dual integer pipelines function, Figure 10 depicts the dual EXUs with the working register files (WRFs), floating-point register files (FRFs), and integer register files (IRFs) shown along with the various data paths.

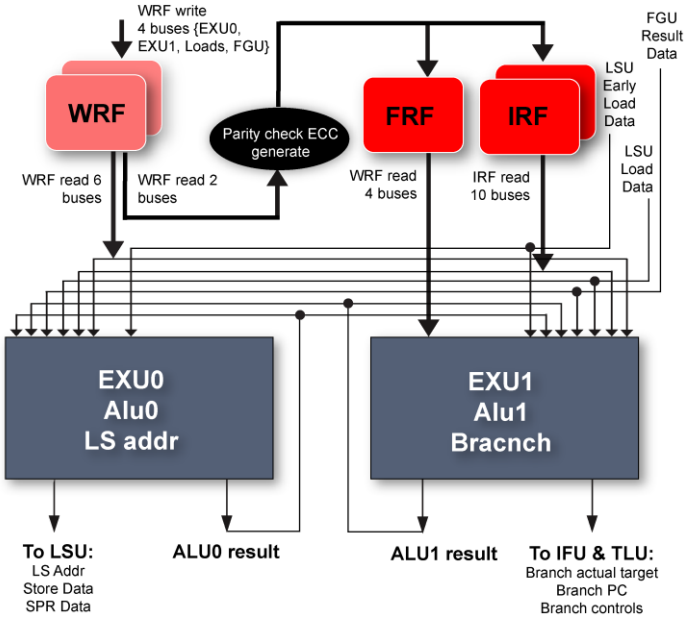

Figure 10. Threads are interleaved between the two integer pipelines and are restricted to EXU0 or EXU1 according to which type of integer operation is to be executed.

Stream Processing Unit

The SPU on each core is implemented within the core as part of the pipelines themselves and operates at the same clock speed as the core. The SPARC T5 processor supports the following cryptographic algorithms:

- DH, DES/3DES
- AES-128/192/256
- Kasumi, Camellia
- CRC32c, MD5
- SHA-1, SHA-224, SHA-256, SHA-384, SHA-512
- RSA via MPMUL/MONTMUL/MONTSQR instructions

A cryptographic algorithm (that is supported in hardware from the groups previously listed) actually uses parts of the FGU and the integer pipelines. Figure 11 illustrates the basic logical pipeline of the SPU.

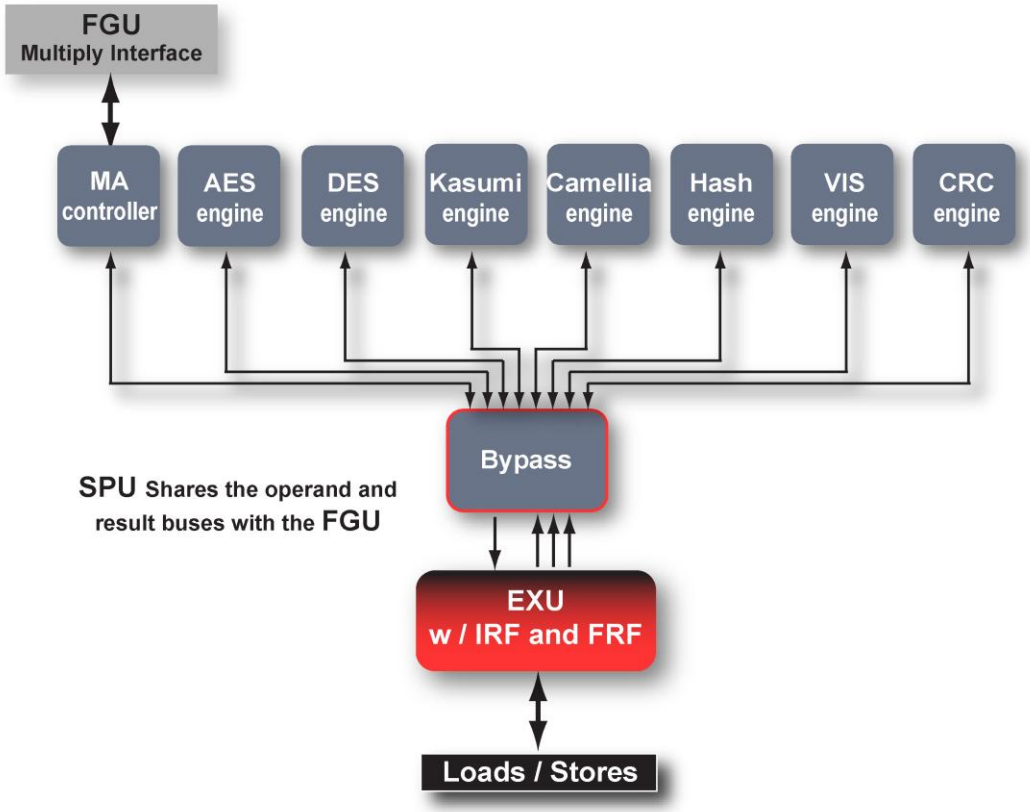

Figure 11. A logical depiction of the SPU pipeline that is in each core.

Integral PCIe Generation 3 Support

SPARC T5 processors provide dual on-chip PCIe Generation 3 (Gen3) interfaces. Each operates at 8 Gb/sec per x1 lane bidirectionally through a point-to-point dual-simplex chip interconnect, meaning that each x1 lane consists of two unidirectional bit-wide connections, one for northbound traffic and the other for southbound traffic. An integral IOMMU supports I/O virtualization and process device isolation by using the PCIe BUS/Device/Function (BDF) number. The total theoretical I/O bandwidth (for an x8 lane) is 16 GB/sec, with a maximum payload size of 256 bytes per PCIe Gen3 interface. The actual realizable bandwidth is more likely to be approximately 14.8 GB/sec. An x8 SerDes interface is provided for integration with off-chip PCIe switches.

Oracle Solaris for Multicore Scalability

Oracle Solaris 10 and 11 are specifically designed to deliver considerable resources for SPARC T5 processor-based systems. In fact, Oracle Solaris provides key functionality for virtualization, optimal use high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. Oracle Solaris runs on a broad range of SPARC- and x86-based systems, and compatibility with existing applications is guaranteed. One of the most attractive features of systems based on SPARC T5 processors is that they appear as a familiar symmetric multiprocessing (SMP) system to Oracle Solaris and the applications it supports. In addition, Oracle Solaris 10 and 11 have incorporated many features to improve application performance on Oracle's multicore/multithreaded architectures.

- **Accelerated cryptography.** Accelerated cryptography is supported through the cryptographic framework in Oracle Solaris (initial release) as well as the SPARC T5 processor. The SPARC T5 processor permits access to cryptographic cypher hardware implementations. For the first time, through user-level instructions, the cyphers are implemented within the appropriate pipeline itself rather than as a coprocessor. This means a more efficient implementation of the hardware-based cyphers as well as no privilege-level changes, resulting in large increase in efficiency in cryptographic algorithm calculations. In addition, database operations can make much more efficient use of the various cryptographic cyphers that are implemented within the instruction pipeline itself. Security is a critical concern at Oracle, and by using the built-in encryption on the SPARC T5 processor across all layers of the Oracle stack, customers gain greater data security with almost no loss in performance.
- **Critical thread optimization.** Oracle Solaris 10 and 11 permit either a user or a programmer to allow the Oracle Solaris Scheduler to recognize a critical thread by means of raising its priority to 60 or above through the CLI or through system calls to a function. If this is done, that thread will run by itself on a single core, garnering all the resources of that core for itself. The one condition that would prevent this single thread from executing on a single core is when there are more runnable threads than available CPUs. This limit was put into place to prevent resource starvation of other threads. Further enhancements to critical thread optimization are planned for Oracle Solaris.

- **Multicore/multithreaded awareness.** Oracle Solaris 10 and 11 are aware of the SPARC T5 processor hierarchy, so the scheduler can effectively balance the load across all available pipelines. Even though it exposes each of these processors as 128 logical processors, Oracle Solaris understands the correlation between cores and the threads they support, and it provides a fast and efficient thread implementation.
- **Fine-granularity manageability.** For the SPARC T5 processor, individual cores and threads (logical processors) can be enabled or disabled by Oracle Solaris. 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, as required or desired.
- **Support for virtualized networking and I/O.** Oracle Solaris contains technology to support and virtualize components and subsystems on the SPARC T5 processor, including support for the on-chip PCIe interface. As part of a high-performance network architecture, Oracle multicore/multithreaded–aware device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices.
- **Non-uniform memory access optimization in Oracle Solaris.** With memory managed by each SPARC T5 processor on the SPARC T5-2, T5-4, and T5-8 servers, these implementations represent a non-uniform memory access (NUMA) architecture. In NUMA architectures, the time needed for a processor to access its own memory is slightly shorter than that required to access memory managed by another processor. Oracle Solaris provides the following technology, which can specifically help to decrease the impact of NUMA on applications and improve performance on NUMA architectures:
  - **Memory placement optimization (MPO).** Oracle Solaris uses MPO to improve the placement of memory across the physical memory of a server, resulting in increased performance. Through MPO, Oracle Solaris helps ensure that memory is as close as possible to the processors that access it, while still maintaining enough balance within the system. As a result, many database applications are able to run considerably faster with MPO.
  - **Hierarchical Lgroup Support (HLS).** HLS improves the MPO feature in Oracle Solaris by optimizing 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 on 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 does not require arbitrary sacrifices. Existing SPARC applications continue to run unchanged on SPARC T5 platforms, protecting software investments. Certified multilevel security protects Oracle Solaris environments from intrusion. The fault management architecture in Oracle Solaris 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 Oracle Solaris DTrace, help organizations tune their applications to get the most out of the system's resources.

Oracle Solaris 11 Operating System

The SPARC T5 servers support both Oracle Solaris 10 and Oracle Solaris 11. Data centers not yet ready for Oracle Solaris 11 can deploy SPARC T5 servers using only Oracle Solaris 10. However, it is highly recommended to deploy using Oracle Solaris 11, because it provides many new features that control administrative costs, increase configuration flexibility, and allow for deployment in to a cloud infrastructure.

Oracle Solaris includes the following features:

- **Advanced reliability**—Uptime is enhanced through comprehensive testing across an integrated solution stack and features such as Predictive Self Healing for hardware and software faults, data integrity with Oracle Solaris ZFS, and live observability with Oracle Solaris DTrace.
- **Superior performance**—Oracle Solaris is optimized for throughput and scalability for the latest SPARC processor technologies and has achieved record-setting benchmarks for the Transaction Processing Performance Council (TPC) TPC-H and TPC-C benchmarks, Oracle's PeopleSoft, Oracle Business Intelligence Enterprise Edition, and many others.
- **Built-in virtualization**—Oracle Solaris Zones and Oracle VM Server for SPARC (formerly called Sun Logical Domains), along with other OS and network virtualization capabilities, enable efficient consolidation for flexibility and performance without significant overhead.
- **Pervasive security infrastructure**—Oracle Solaris provides the compartmentalization and control needed for multitenancy environments and enables governments and financial institutions to meet their strict requirements.
- **Committed support**—Oracle offers sustaining support for Oracle Solaris releases for as long as customers operate their systems, making it possible to keep software infrastructures in place for as long as it makes business sense.

Oracle Solaris Predictive Self Healing, Fault Management Architecture, and Service Management Facility

Oracle Solaris provides an architecture for building and deploying systems and services capable of fault management and predictive self-healing.

- The Predictive Self Healing feature in Oracle Solaris 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.
- The Fault Management Architecture in Oracle Solaris collects data relating to hardware and software errors. This facility automatically and silently detects and diagnoses the underlying problem, with an extensible set of agents that automatically respond by taking the faulty component offline.
- The Service Management 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, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem.

Predictive Self Healing and the Fault Management Architecture can offline processor threads or cores in faults, retire suspect pages of memory, log errors or faults from I/O or any other issue detected by the system.

Oracle Solaris Cryptographic Frameworks

Oracle Solaris cryptographic frameworks at the user and kernel level provide applications with a direct interface to several security ciphers that can be implemented at the software or hardware levels. With Oracle's multicore/multithreaded processors and with these ciphers implemented in the chip, applications running on Oracle Solaris can transparently get to the cryptographic service at the hardware level, which drastically reduces not only the development time of secured applications but also the load on the system when the applications are deployed. This helps developers create secured applications for large-scale deployments.

End-to-End Virtualization Technology

Virtualization technology is increasingly popular as organizations strive to consolidate disparate workloads onto fewer, more-powerful systems, while increasing use. The SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module are designed specifically for virtualization, providing very fine-grained division of multiple resources—from processing to virtualized networking and I/O. Most important, Oracle's virtualization technology is provided as a part of the system, not an expensive add-on.

A Multithreaded Hypervisor

Like the prior generations of UltraSPARC and SPARC processors, the SPARC T5 processor offers a multithreaded hypervisor—a small firmware layer that provides a stable virtual machine architecture that is tightly integrated with the processor. Multithreading is crucial, because the hypervisor interacts directly with the underlying multicore/multithreading processor. This architecture is able to context-switch between multiple threads in a single core, a task that would require additional software and considerable overhead in competing architectures.

Corresponding layers of virtualization technology are built on top of the hypervisor, as shown in Figure 12. The strength of Oracle's approach is that all the layers of the architecture are fully multithreaded, from the processor up through applications that use the fully threaded Java application model. Far from being new technology, Oracle Solaris has provided multithreading support since 1992. This experience has helped to inform technology decisions at other levels, ultimately resulting in a system that parallelizes and virtualizes at every level. In addition to the processor and hypervisor, Oracle provides fully multithreaded networking and the fully multithreaded Oracle Solaris ZFS file system. Oracle VM Server for SPARC, Oracle Solaris Zones, and multithreaded applications are able to receive exactly the resources they need.

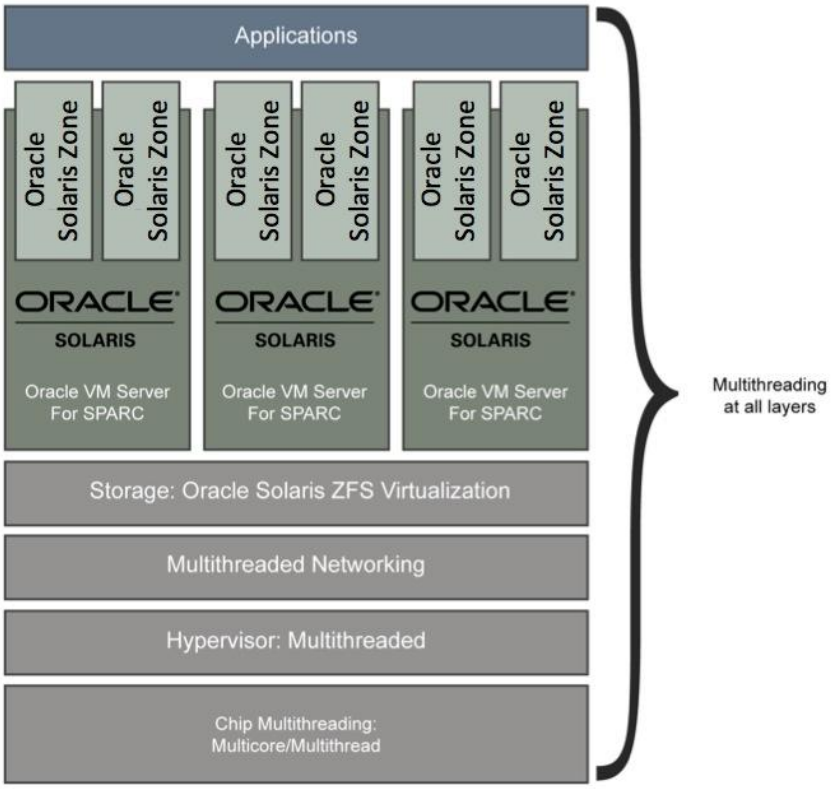

Figure 12. Oracle provides parallelization and virtualization at every level of the technology stack.

Oracle VM Server for SPARC

Oracle VM Server for SPARC is supported in all servers from Oracle that use Oracle's multicore/multithreaded technology, and it provides full virtual machines that run an independent operating system instance. Each operating system instance contains virtualized CPU, memory, 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 domain. Each domain is completely isolated, and the maximum number of virtual machines created on a single platform relies upon the capabilities of the hypervisor, rather than on the number of physical hardware devices installed in the system. For example, the SPARC T5-1B server module with a single SPARC T5 processor supports up to 128 domains, and each individual domain can run a unique OS instance.

Oracle VM Server for SPARC 3.0 has the ability to perform a live migration from one domain to another. As the term *live migration* implies, the source domain and application no longer need to be halted or stopped. Migration of a running application from one domain to another is now possible with Oracle VM Server for SPARC 3.0. This allows a logical domain on the server to be live-migrated to a PDom on a SPARC M5-32 server from Oracle or to another SPARC T3– or SPARC T5–based server.

By taking advantage of domains, organizations gain the flexibility 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. Domains can also host Oracle Solaris Zones to capture the isolation, flexibility, and manageability features of both technologies. By deeply integrating Oracle VM Server for SPARC with the SPARC T5 processor, Oracle Solaris increases flexibility, isolates workload processing, and improves the potential for maximum server utilization.

Oracle Solaris Zones

Oracle Solaris 11 provides a unique partitioning technology called Oracle Solaris Zones, which was called Oracle Solaris Containers in Oracle Solaris 10. This technology can be used to create an isolated and secure environment for running applications. An Oracle Solaris Zone is a virtualized operating system environment created within a single instance of Oracle Solaris. Oracle Solaris 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 Oracle Solaris Zone are prevented from interfering with processes running in another Oracle Solaris Zone.

CPUs in a multiprocessor system (or threads in the SPARC T5 processor) can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to an Oracle Solaris Zone. Resource pools provide the capability to separate workloads so that the consumption of CPU resources does not 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.

Oracle Solaris Zones technology is an excellent tool for consolidating older environments on to newer platforms. This allows applications to benefit from the increased performance of the latest CPU, memory, and I/O technology, as well as enabling applications to be deployed on systems with higher levels of reliability, availability, and serviceability (RAS). Figure 13 shows a typical consolidation scenario.

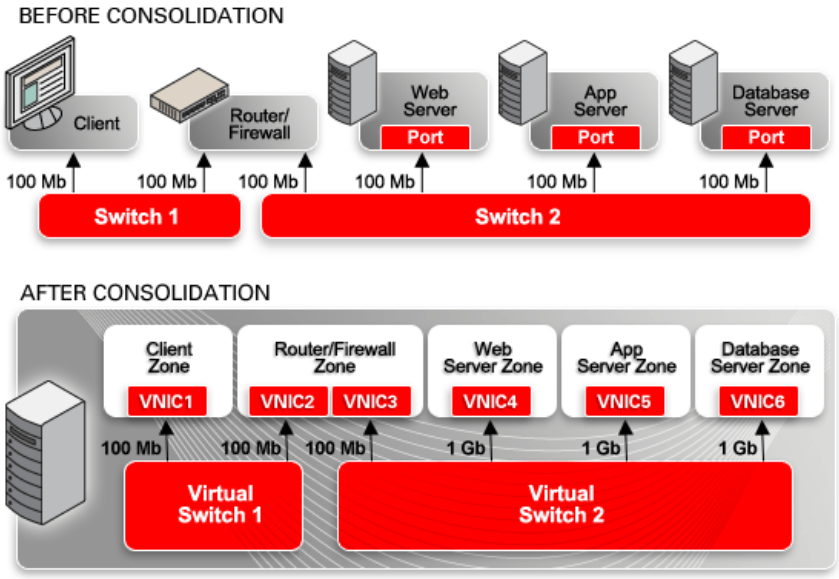

Figure 13. Advantages of server consolation with Oracle Solaris Zones.

Reduced TCO and business goals are achieved when multiple virtualization technologies are combined. For example, multiple Oracle Solaris Zones can run inside each logical domain. In Figure 14, each logical domain separates different Oracle Solaris releases. This is usually done when there are different costs for running the latest OS. The next step is to isolate applications inside their own Oracle Solaris Zone. This allows for fine-grained resource allocation and process isolation. The use of Oracle Solaris Legacy Containers is for applications certified only for Oracle Solaris 8 or 9 and allows the administrator to take advantage of the new SPARC T5 performance and features, while not being forced to upgrade all the software until the business dictates the change.

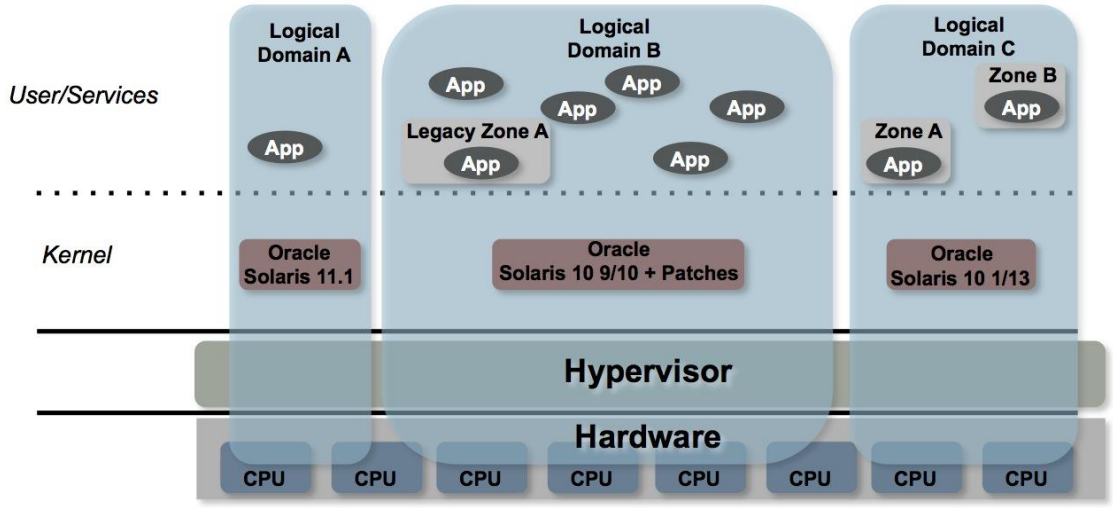

Figure 14. Running Oracle Solaris Zones inside each logical domain.

Enterprise-Class Management

Although 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 SPARC T5 processor technology, the SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module provide full binary compatibility with earlier SPARC systems and are delivered ready to run with preloaded tools and the solid foundation of Oracle Solaris. 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 the SPARC T5 processor.

System Management Technology

As the number of systems grows in any organization, managing an increasingly complex infrastructure throughout its lifecycle becomes 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.

Oracle Integrated Lights Out Manager

The Oracle Integrated Lights Out Manager (Oracle ILOM) service processor is provided across all of Oracle's servers and acts as a system controller, facilitating remote management and administration of the SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module. The service processor is full-featured and is similar in implementation to that used in Oracle's other modular and rackmount x86 servers. As a result, these servers integrate easily with existing management infrastructure. Critical to effective system management, Oracle ILOM

- Implements an IPMI 2.0–compliant service 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 Oracle ILOM 3.2 Ethernet management interface. The service processor also provides visibility to the environmental sensors 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, and SNMP access to this data.
- Supplies remote textual console interfaces.
- Provides a means to download upgrades for all system firmware.

The Oracle ILOM 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 ILOM can send e-mail alerts to each server about hardware failures, warnings, and other events related. Its circuitry runs independently from the server, using the server's standby power. As a result, Oracle ILOM 3.2 firmware and software continue to function when the server operating system goes offline or when the server is powered off. Oracle ILOM monitors the following server conditions:

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

Power Management

The power and cooling costs for servers are becoming significant, and lowering these costs is a top challenge in the corporate data center. Limitations in the availability of power and space to expand data centers force customers to look closely at the power efficiency of servers. Contracts with power providers, which specify penalties for exceeding stated power consumption, require servers to be able to cap their power consumption under customer control. Power efficiency and carbon footprint have become factors when customers evaluate servers.

Beyond the inherent efficiencies of Oracle's multicore/multithreaded design, the SPARC T5 processor incorporates 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, which allows extra stall cycles to be injected into the decode stage

In a virtualized environment using Oracle VM Server for SPARC, the power management manager performs the following tasks when managing logical domain guests:

- Determining which power savings features to enable based on the power management policy
- Handling hypervisor event notifications for its resources
- Calling the Power Management engine to initiate power state changes on its resources to achieve a power adjustment or utilization level (for resources not owned by an Oracle Solaris guest) or telling the hypervisor to enable or disable hypervisor/hardware–managed power states
- Coordinating power state changes with its power management manager peer in each Oracle Solaris guest, including
  - Contacting one or more Oracle Solaris guests to request a power increase or reduction (for example, for power capping)
  - Contacting one or more Oracle Solaris guests to power on or power off a specific resource (for example, for power capping)
  - Handling Oracle Solaris notifications of an offlined resource or requests to bring a resource online (for example, to assist in the offline/online action and notify the affinity engine)
  - Contacting one or more Oracle Solaris guests to provide hints about a power savings policy for specific resources (for example, to manage a resource-sharing relationship)

On systems that support only one physical domain, a power policy is set for the entire system via the existing interface: SP/powermgmt policy. On systems with multiple physical domains, the power policy can be set for each physical domain.

An important element of controlling power utilization is the ability to cap power consumption. Both "hard caps" (power limit with grace period of 0) and "soft caps" (power limit with grace period > 0) are supported. A power cap enabled at the physical domain is measured against the power consumption of boards fully owned by the physical domain. In the case of SPARC T5 systems, the physical domain is the entire server. The logical domain Power Manager adjusts the power states of the physical domain resources to converge to the cap.

Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center delivers a converged hardware management solution for SPARC T5 servers that integrates management across the infrastructure stack. With advanced virtualization management and reporting capabilities, application-to-disk management, intelligent configuration management and more, Oracle Enterprise Manager Ops Center helps IT managers reduce complexity and streamline and simplify infrastructure management. Other solutions on competitive platforms require sometimes three, four, or even five different management applications, which places a heavy burden on administrators to reconcile messages and alerts.

Oracle Enterprise Manager Ops Center enables data center administrators to monitor and manage the storage, network, servers, Oracle Solaris, and virtualized environments from a single interface. This improves operational efficiency and lowers operational costs.

Figure 15. Management of the Oracle stack.

Oracle Enterprise Manager Ops Center is the most comprehensive management solution for Oracle servers and engineered systems infrastructure. Offering a single console for managing multiple server architectures and myriad operating systems, Oracle Enterprise Manager Ops Center can manage the components in SPARC T5 servers using asset discovery, provisioning of firmware and operating systems, automated patch management, patch and configuration management, virtualization management, and comprehensive compliance reporting.

Oracle Enterprise Manager Ops Center automates workflow and enforces compliance via policy-based management—all through a single, intuitive interface. With Oracle Enterprise Manager Ops Center, IT staff can implement and enforce data center standardization and best practices, regulatory compliance, and security policies while efficiently deploying infrastructure to meet business requirements. Figure 16 shows the intuitive browser-based user interface for Oracle Enterprise Manager Ops Center.

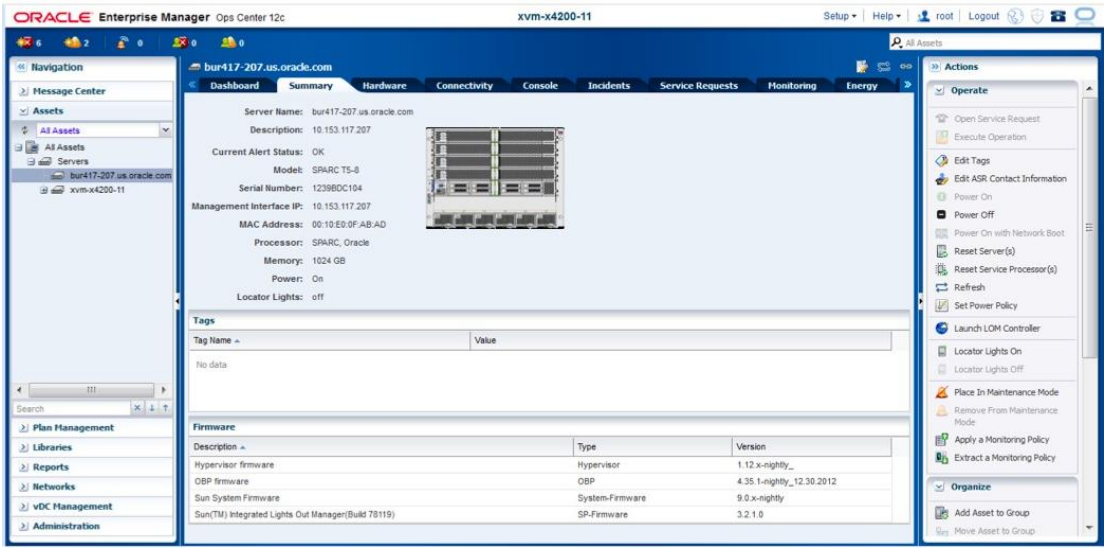

Figure 16. The Oracle Enterprise Manager Ops Center interface.

For further information and a detailed description of Oracle Enterprise Manager Ops Center, please refer to http://www.oracle.com/technetwork/oem/ops-center.

Conclusion

Delivering the demands of IT services, applications, and virtualized eco-efficient data centers requires a comprehensive approach that includes innovative processors, system platforms, and operating systems, along with leading application, middleware, and management technologies. 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 being 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 upon the successful base of previous-generation SPARC T-Series processors, the SPARC T5 processor serves as the industry's next-generation massively threaded RISC processor. It delivers high application throughput and efficiency while offering single-threaded performance gains. With 128 threads per processor, on-chip memory management, two PCIe Generation 3 root complexes, and on-chip cryptographic acceleration, the SPARC T5 processor fundamentally redefines the capabilities of a modern processor. By incorporating cache coherency for multiprocessor support, SPARC T5 processors allow these capabilities to be multiplied incrementally. The SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module leverage these strengths to provide powerful and highly scalable server platforms while delivering even higher levels of performance in a compact rackmount chassis. The result is a data center infrastructure with a very small footprint that can truly scale to meet new challenges.

The SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module provide the computational, networking, and I/O resources needed by the most demanding databases, IT services, enterprise applications, and Web services while 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. With innovations such as Oracle VM Server for SPARC, Oracle Solaris Zones, and Java technology, organizations can adopt these radical new systems for their most important projects—acting responsibly toward the environment and the bottom line.

For More Information

For more information on Oracle's SPARC T5 server line and related software and services from Oracle, please see the references listed in Table 3.

TABLE 3. REFERENCES

SPARC systems	http://www.oracle.com/us/products/servers-storage/servers/sparc-enterprise
Oracle Solaris	http://www.oracle.com/us/products/servers-storage/solaris
Oracle Solaris Cluster	http://www.oracle.com/us/products/servers-storage/solaris/cluster/overview/index.html
Oracle Enterprise Manager Ops Center software	http://www.oracle.com/technetwork/oem/ops-center
Oracle Support	http://www.oracle.com/us/support/index.html

Appendix A: Server Architectures

The SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module have been designed to module provide breakthrough performance while maximizing reliability and minimizing power consumption and complexity. This section details the physical and architectural aspects of these systems.

Motherboard and Memory Subsystem

A separate motherboard design is used in each of the SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module. Although each system deploys the same technologies (SPARC T5 processor, DDR3 memory, PCIe Gen3 switches, SAS-2), each motherboard has its own unique design and form factor. Common features of the SPARC T5 motherboard (Processor Module for the SPARC T5-4 server) are

- A minimum of one socket for a SPARC T5 processor
- Memory slots to supply memory for the SPARC T5 server
- A remote-to-local memory latency ratio of 1.47 (in the case of a two- or four-processor system running the `lmbench` read access test)

In the SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module, the SPARC T5 processor provides on-chip memory controllers that communicate indirectly to DDR3 DIMMs via newly designed buffer-on-board (BoB) memory interfaces through four high-speed serial links. Four dual-channel memory controller units (MCUs) are provided on the SPARC T5 processor. Each MCU can transfer data at an aggregate rate of 12.8 Gb/sec. There are 16 motherboard memory socket locations for each SPARC T5 processor. The SPARC T5-2 server's motherboard has 32 DIMM slots broken up over four memory risers, each of which plug into the motherboard. There are up to 64 motherboard memory socket locations on the SPARC T5-4 server's motherboard and 16 motherboard memory socket locations on the SPARC T5-1B server module's motherboard, providing sufficient board space for two 1066-MHz DDR3 DIMMs per channel.

I/O Subsystem

Each SPARC T5 processor interfaces through two (x8) PCIe Gen3 ports capable of operating at 14 GB/sec bidirectionally. In each server, these ports natively interface through the I/O devices through PCIe Gen3 switch chips, connecting either to PCIe card slots or to bridge devices that interface with PCIe, such as the following:

- **Disk controller.** Disk control is managed by an LSI Logic SAS2008 SAS/SATA controller chip. RAID levels 0, 1, and 10 are supported. The LSI controller chip also drives the DVD (optical drive) in the SPARC T5-2 server. (DVDs are not supported in the SPARC T5-4 or T5-8 servers or in the SPARC T5-1B server module.)

- **Modular disk backplanes.** Depending on the system, a six- or eight-disk backplane is attached to the LSI disk controller by one or more x4 SAS-2 links. The SPARC T5-2 server supports a SAS-2-capable six-disk backplane, and the SPARC T5-4 and T5-8 servers support a SAS-2-capable eight-disk backplane.
- **10 GbE.** The SPARC T5-2, T5-4, and T5-8 servers provide four 10 Gb/sec Ethernet interfaces on the rear of each chassis. Two 10/100/1000 Mb/sec Ethernet interfaces are provided on the SPARC T5-1B server module.
- **USB.** On all servers, a single-lane PCIe port connects to a PCI bridge device. A second bridge chip converts the 33-bit 66-MHz PCI bus into multiple USB 2.0 ports.

Enhanced System and Component Serviceability

Finding and identifying servers and components in a modern data center can be challenging. The SPARC T5-2, T5-4, and T5-8 servers and the SPARC T5-1B server module are optimized for lights-out data center configurations so they are easy to identify. Color-coded operator panels provide easy-to-understand diagnostics and the systems are designed for deployment in hot-aisle/cold-aisle 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 systems straightforward. All hot-pluggable components are tool-less and easily available for serviceability. For instance, easy access to fan modules enables fans to be serviced without exposing sensitive components or causing unnecessary downtime.

Robust Chassis, Component, and Subassembly Design

Several of Oracle's 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 maximum electronic attenuation. Next-generation hard disk drive carriers leverage the hexagonal ventilation of the chassis and provide a small front plate for greater storage density while increasing airflow to the system.

In spite of their computational, I/O, and storage density, Oracle's servers are able to maintain adequate cooling using conventional technologies. Minimized DC-DC power conversions also contribute to overall system efficiency. 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 are employed, appropriate to each chassis. These infrastructure boards serve various functions in the SPARC T5-2, T5-4, and T5-8 servers.

- Power distribution boards distribute system power from multiple power supplies to the major components of the system.

- Fan boards provide connections for power and control for both the primary and secondary fans in the front or rear of the chassis. No cables are required because every fan module plugs directly into one of these boards.
- The disk backplane mounts to the disk cages in the chassis, delivering disk data through one or two four-channel, discrete mini-SAS cables from the motherboard.
- The SPARC T5-2, T5-4, and T5-8 servers all support USB 2.0 and 3.0 interfaces: USB 2.0 on the front and USB 3.0 on the back of the chassis. The SPARC T5-2, T5-4, T5-8 servers and the SPARC T5-1B server module also have an internal USB port. The SPARC T5-1B server module supports two external USB 2.0 interfaces on the front via a dongle cable.

Overview of Oracle's SPARC T5-2 Server

The expandable SPARC T5-2 server is optimized to deliver transaction and Web services, including Java 2 Platform and Enterprise Edition (J2EE platform) technology application services, enterprise application services such as enterprise resource planning (ERP), customer relationship management (CRM), supply chain management (SCM), and distributed databases. With considerable expansion capabilities and integrated virtualization technologies, the SPARC T5-2 server is also an ideal platform for consolidated Tier 1 and Tier 2 workloads. The SPARC T5-2 comes in two configurations: 1) with a single SPARC T5 processor installed, or 2) with two SPARC T5 processors installed.

The SPARC T5-2 server block diagrams are shown in Figure 17.

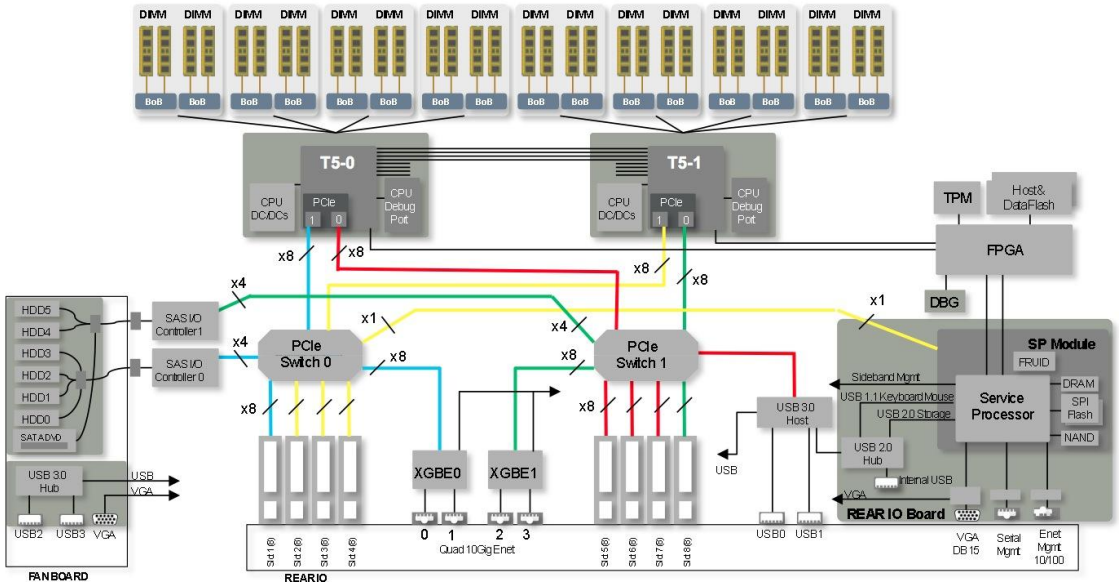

Figure 17a. The SPARC T5-2 server block diagram (two processors).

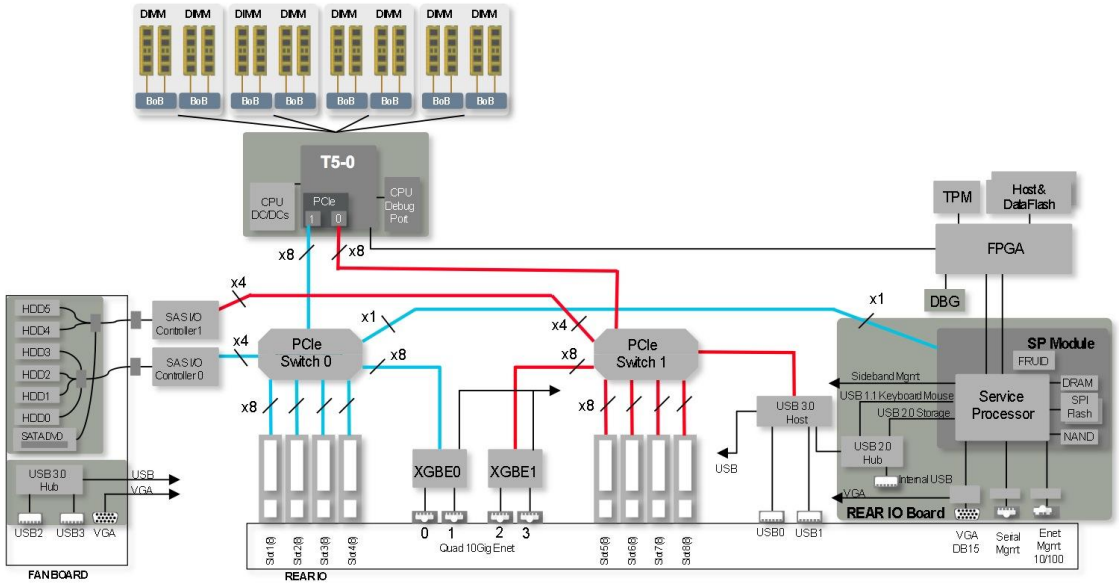

Figure 17b. The SPARC T5-2 server block diagram (one processor).

Enclosure

The SPARC T5-2 server features a compact, expandable 3RU rackmount chassis, giving companies the flexibility to scale processing and I/O requirements without wasting precious space (Table 4).

TABLE 4. DIMENSIONS AND WEIGHT OF THE SPARC T5-2 SERVER

SERVER/DIMENSION	US	INTERNATIONAL
Height	5.11 inches (3RU)	129.85 millimeters
Width	17.18 inches	436.5 millimeters
Depth	28.81 inches	732 millimeters
Weight (without PCIe cards or rackmounts)	80 pounds	36.28 kilograms

The SPARC T5-2 server includes the following major components:

- Single or dual SPARC T5 processors with sixteen cores per processor operating at 3.6 GHz
- Up to 512GB of memory in 16 DDR3 DIMM slots (single processor), or up to 1TB of memory in 32 DDR3 DIMM slots (dual processors). 8 GB, 16 GB, and 32 GB DDR3 DIMMs are supported.
- Four onboard 10 GbE ports

- Eight dedicated low-profile PCIe x8 slots
- Three USB 2.0 ports (two forward, one internal that is restricted to the thumb drive)
- Two USB 3.0 ports (two rear-facing)
- Six available disk drives slots that support SAS-2 commodity disk drives
- Oracle ILOM 3.2 system controller
- Two (N+1) hot-pluggable/hot-swappable high-efficiency 2060 watt AC power supplies
- Six fan assemblies under environmental monitoring and control, N+1 redundancy

Front and Rear Perspectives

Figure 18 illustrates the front and back panels of the SPARC T5-2 server.

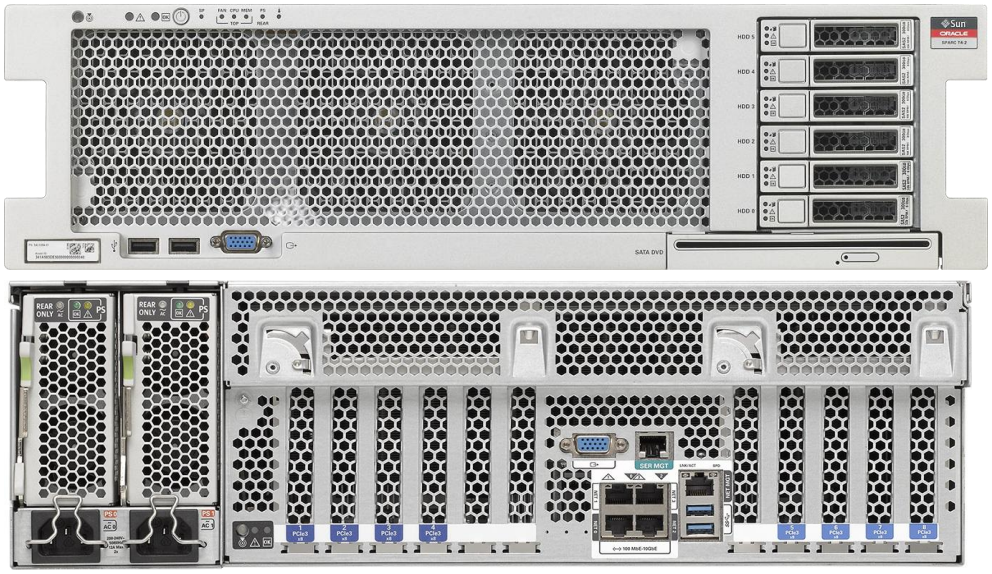

Figure 18. The front and back panels of the SPARC T5-2 server.

External features of the SPARC T5-2 server include the following:

- Front and rear system and component status indicator lights that provide locator (white), service required (amber), and activity status (green) information for the system
- Two management ports for use with the Oracle ILOM 3.0 system controller
  - An RJ-45 serial management port provides a default connection to the Oracle ILOM 3.0 controller.
  - A network management port supports an optional RJ-45 10/100Base-T connection to the Oracle ILOM 3.2 system controller.

Overview of Oracle's SPARC T5-4 Server

With support for four SPARC T5 processors and 512 threads, the compact SPARC T5-4 server provides breakthrough computational power in a space-efficient, 5RU rackmount package. With breakthrough levels of price/performance, this server is ideally suited to the delivery of horizontally scaled transaction and Web services as well as medium-to-large database applications. It presents many opportunities as a consolidation and virtualization server due to its large capacity. The server is designed to address the challenges of modern data centers that require a compact footprint combined with a greatly increased performance capability. Uniquely setting this processor apart from past multicore processors is its ability to deliver a 5x performance increase on single-threaded workloads.

The SPARC T5-4 server has a unique motherboard design (see Figure 19).

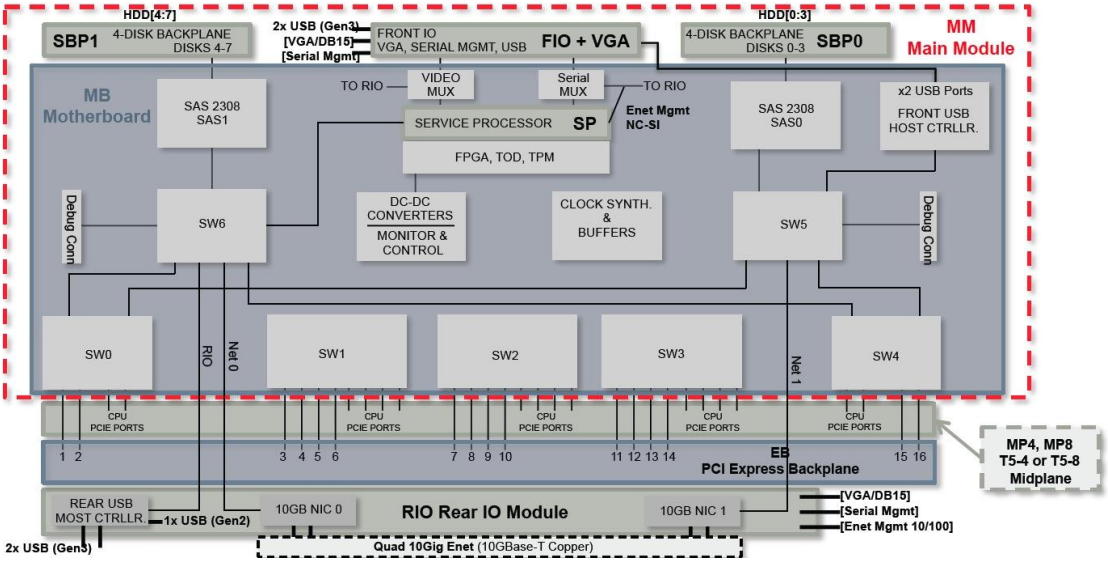

Figure 19. The SPARC T5-4 server's motherboard design.

In Figure 20, the I/O paths are shown for a fully configured SPARC T5-4 server.

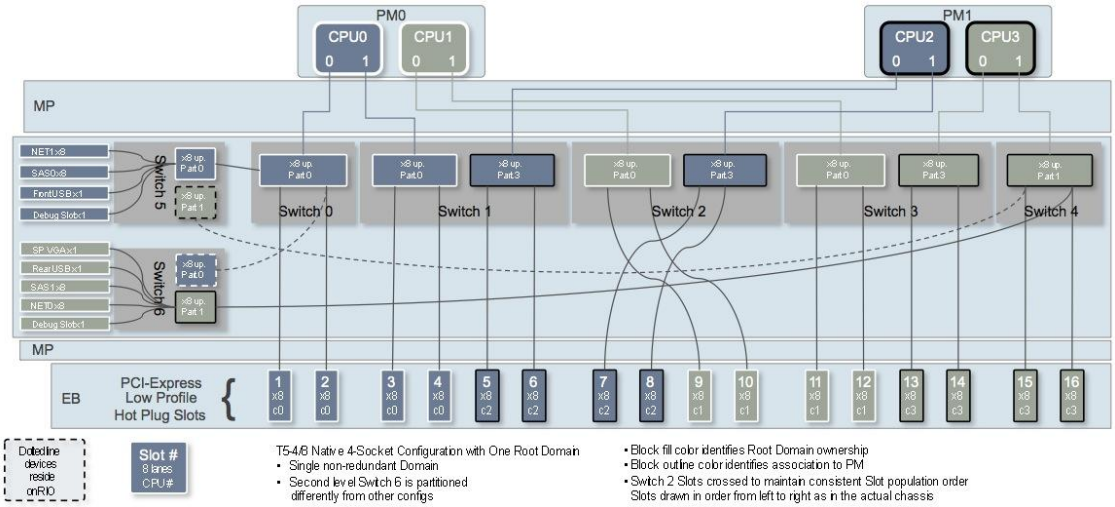

Figure 20. The SPARC T5-4 server's I/O layout.

Enclosure

The 5RU SPARC T5-4 server enclosure is designed for use in a standard 19-inch rack (Table 5).

TABLE 5. DIMENSIONS AND WEIGHT OF THE SPARC T5-4 SERVER

DIMENSION	US	INTERNATIONAL
Height	8.62 inches (5RU)	219 millimeters
Width	17.5 inches	445 millimeters
Depth	27.6 inches	700 millimeters
Weight (without PCIe cards or rackmounts)	175 pounds	79 kilograms

The SPARC T5-4 server includes the following major components:

- Four SPARC T5 processors, sixteen cores per processor operating at clock speed of 3.6 GHz
- Up to 2 TB of memory in 64 DDR3 DIMM slots (16-GB and 32-GB DDR3 DIMMs are supported)
- Four onboard 10 GbE ports
- Sixteen x8 PCIe Gen3 low-profile slots, with hot-pluggable carriers
- Four USB ports (two in front are USB 2.0, two in rear are USB 3.0)
- Eight available disk drives slots that support SAS-2 commodity disk drives
- Oracle ILOM 3.2 system controller
- Two (1+1) hot-swappable, high-efficiency 3000 watt AC power supplies
- Five fan assemblies, under environmental monitoring and control, 2 + 2 redundancy

Front and Rear Perspectives

Figure 21 illustrates the front and rear panels of the SPARC T5-4 server.

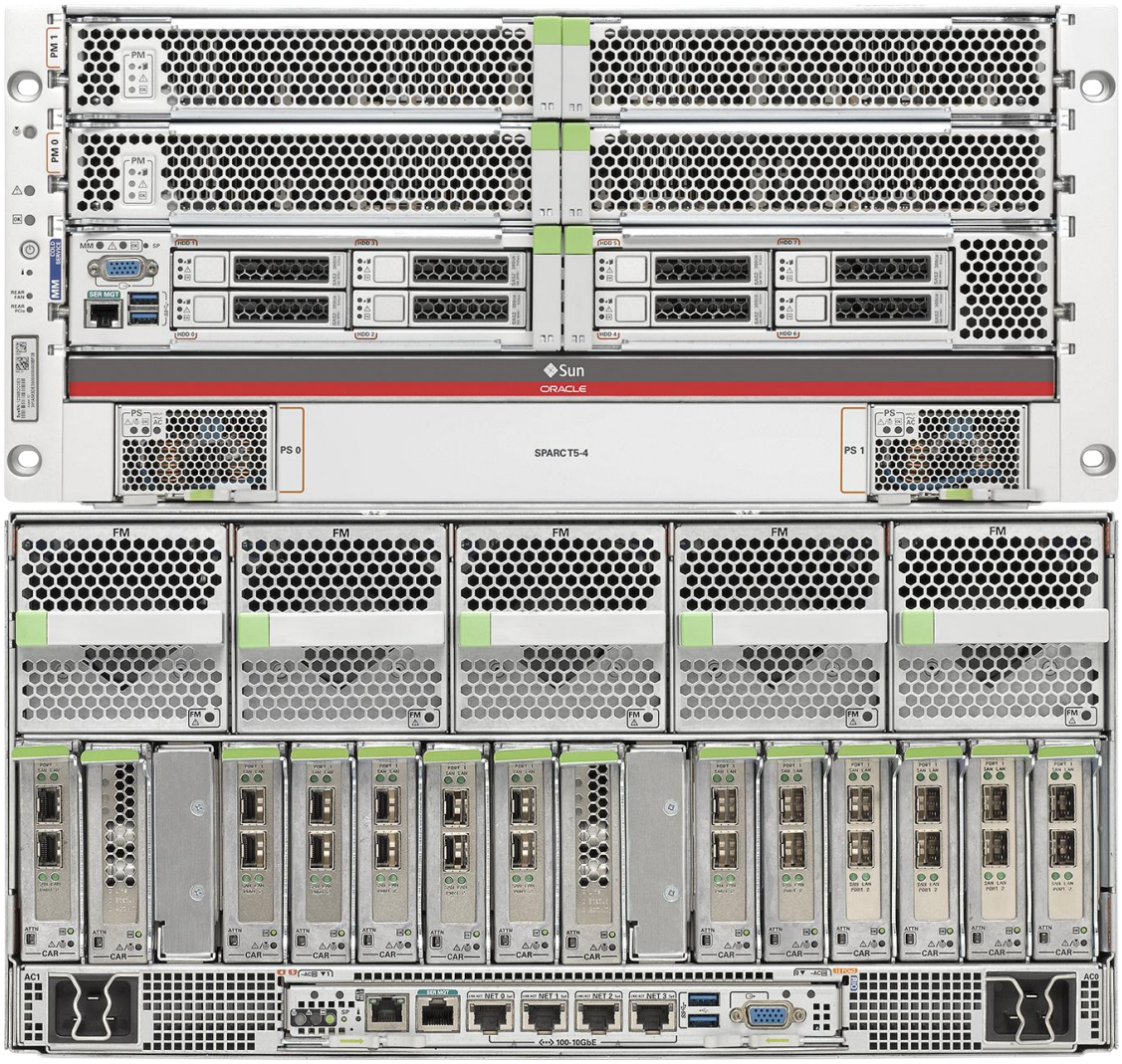

Figure 21. The front and rear panels of the SPARC T5-4 server.

External features on the front of the SPARC T5-4 server include the following:

- Serviced front/rear only (no sliding rack rails)
- Two Processor Modules, each with 2x SPARC T5 CPUs, with 32 DDR3 DIMMs (16-GB or 32-GB DDR3 DIMMs are supported)
- Eight HDDs
- Two power supplies (1+1)
- An HD-15 VGA video port
- Two USB ports
- Console serial port

External features on the rear of the SPARC T5-4 server include the following:

- Serviced front/rear only (no sliding rack rails)
- Sixteen hot-pluggable PCIe carriers
- Rear I/O module with 4x 10Gb networks
- Five fans (N+1)
- Two (1+1) AC cords (200–240V)
- Two USB ports
- Console serial port (duplicate of front)
- Console 10/100 network port
- LEDs and indicators

Overview of Oracle's SPARC T5-8 Server

With support for eight SPARC T5 processors and 1024 threads, the compact SPARC T5-8 server provides breakthrough computational power in a space-efficient, 8RU rackmount package. With breakthrough levels of price/performance, this server is ideally suited to the delivery of horizontally scaled transaction and Web services as well as medium-to-large database applications. It presents many opportunities as a consolidation and virtualization server due to its large capacity. The server is designed to address the challenges of modern data centers that require a compact footprint combined with a greatly increased performance capability compared to the previous-generation SPARC T4-4 system from Oracle.

The SPARC T5-8 server has a unique motherboard design (see Figure 22).

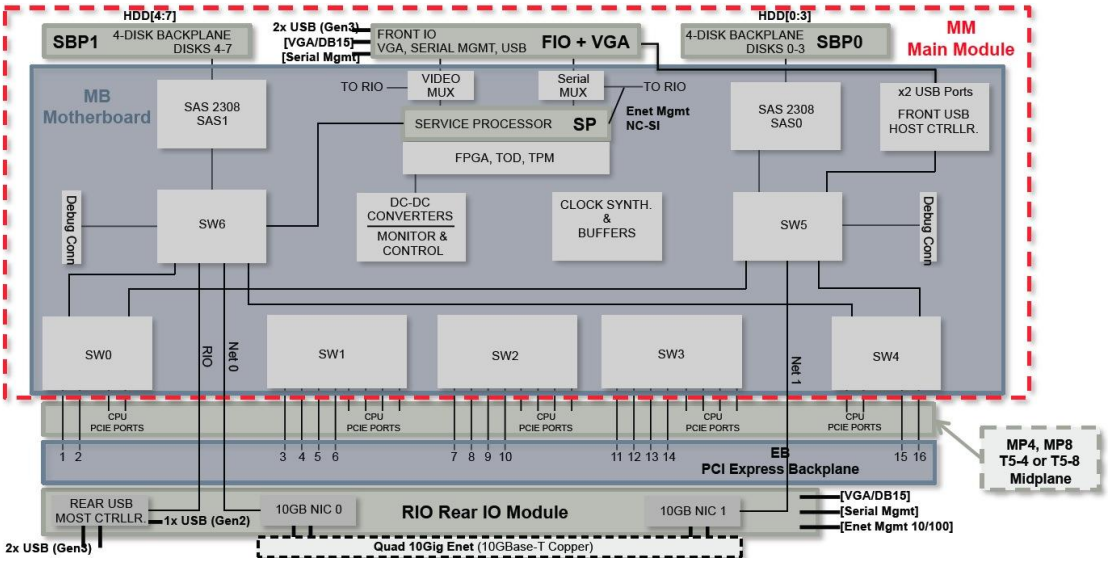

Figure 22. The SPARC T5-8 server motherboard design.

In Figure 23, the I/O paths are shown for a fully configured SPARC T5-8 server.

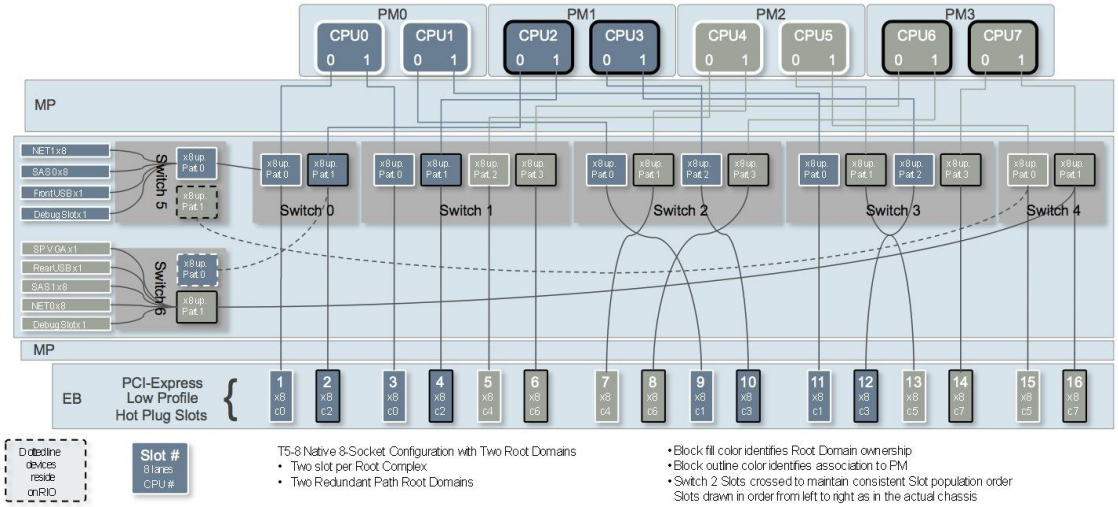

Figure 23. The SPARC T5-8 server's I/O layout.

Enclosure

The 8RU SPARC T5-8 server enclosure is designed for use in a standard 19-inch rack (Table 6).

TABLE 6. DIMENSIONS AND WEIGHT OF THE SPARC T5-8 SERVER

DIMENSION	US	INTERNATIONAL
Height	13.8 inches (8RU)	350 millimeters
Width	17.5 inches	445 millimeters
Depth	31.5 inches	800 millimeters
Weight (without PCIe cards or rackmounts)	261.5 pounds	118.6 kilograms

The SPARC T5-8 server includes the following major components:

- Eight SPARC T5 processors, sixteen cores per processor operating at clock speed of 3.6 GHz
- Up to 4 TB of memory in 128 DDR3 DIMM slots (16-GB and 32-GB DDR3 DIMMs are supported)
- Four onboard 10 GbE ports
- Sixteen x8 PCIe Gen3 slots with hot-pluggable carriers for low-profile cards
- Four USB ports (two in front are USB 2.0, two in rear are USB 3.0)
- Eight available disk drives slots that support SAS-2 commodity disk drives
- Oracle ILOM 3.2 system controller

- Four (N+N) hot-swappable, high-efficiency 3000 watt AC power supplies
- Five fan assemblies, under environmental monitoring and control, 2 + 2 redundancy

Front and Rear Perspectives

Figure 24 illustrates the front and rear panels of the SPARC T5-8 server.

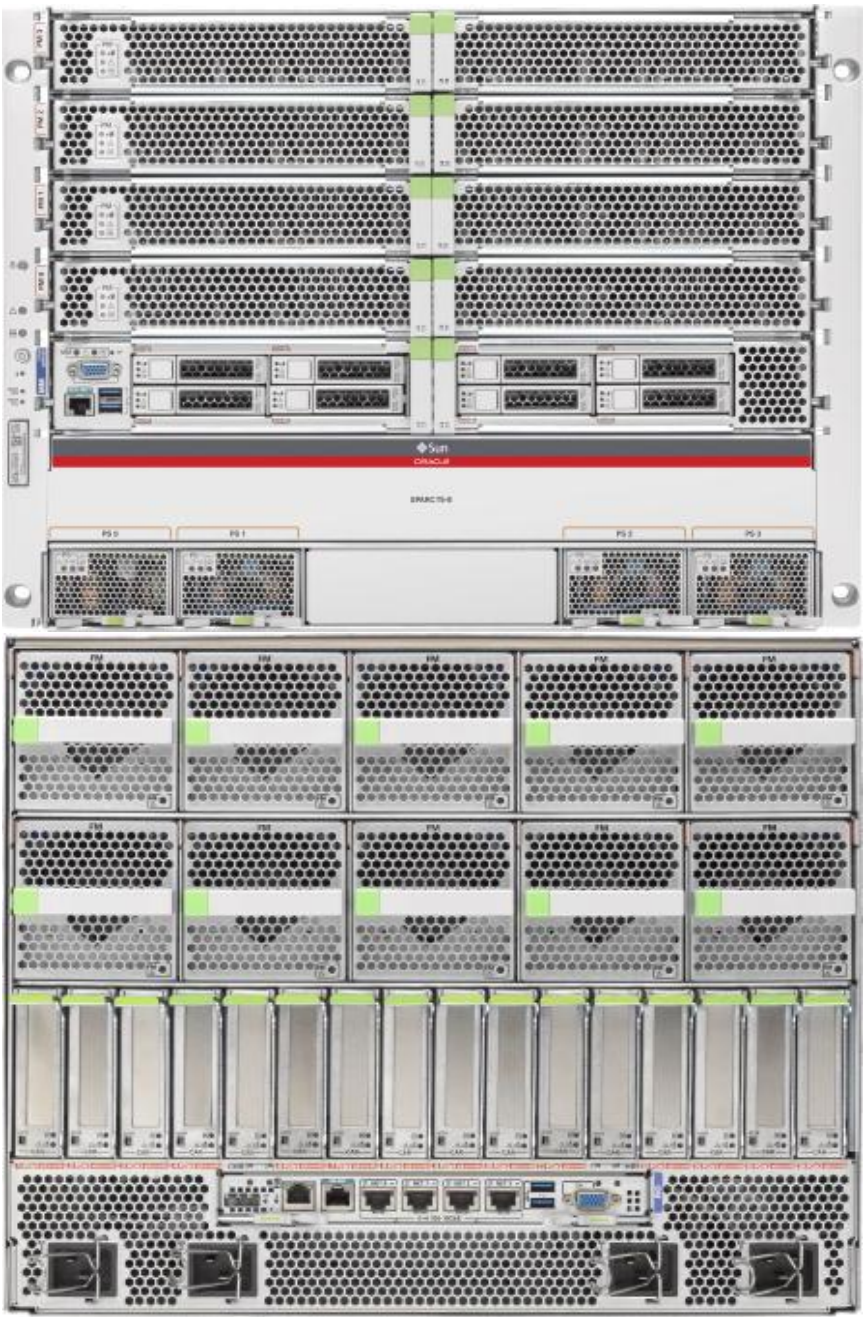

Figure 24. The front and rear panels of the SPARC T5-8 server.

External features on the front of the SPARC T5-8 server include the following:

- Serviced front/rear only (no sliding rack rails)
- Four Processor Modules, each with 2x SPARC T5 CPUs, with 32 DDR3 DIMMs (16-GB or 32-GB DDR3 DIMMs are supported)
- Eight HDDs
- Four power supplies (2+2)
- An HD-15 VGA video port
- Two USB 2.0 ports
- Console serial port

External features on the rear of the SPARC T5-8 server include the following:

- Serviced front/rear only (no sliding rack rails)
- Sixteen hot-swappable low-profile PCIe 3.0 carriers
- Rear I/O module with 4x 10 Gb networks
- Five fans (N+1)
- Four (2+2) AC cords (200–240V)
- Two USB 3.0 ports
- Console serial port (duplicate of front)
- Console 10/100 network port
- LEDs and indicators

Overview of Oracle's SPARC T5-1B Server Module

The SPARC T5-1B server module is optimized to deliver strong performance for streaming media, virtualization and consolidation, and other back-office applications such as Java application servers, OLTP databases, ERP, CRM, SOA, and business integration. Providing support for a SPARC T5 processor in a blade server form factor, it is ideal for expansion capabilities and integrated virtualization technologies. The SPARC T5-1B server module is also an ideal platform for consolidating Tier 1 and Tier 2 workloads.

The SPARC T5-1B server module has a unique motherboard design (Figure 25).

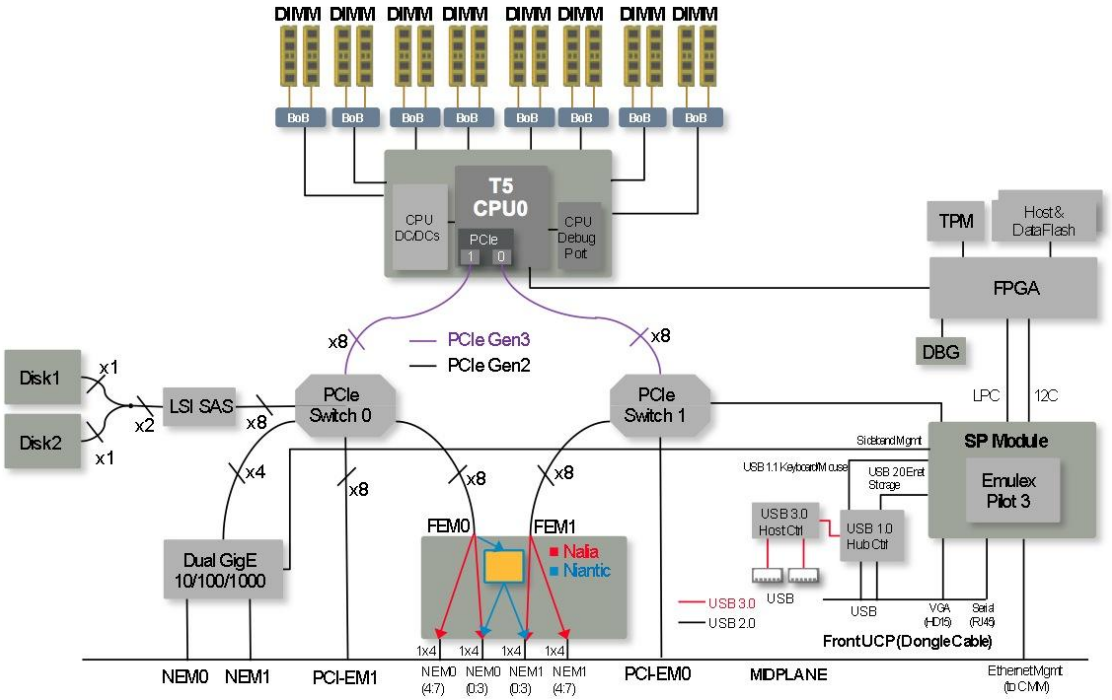

Figure 25. The SPARC T5-1B server module's motherboard design.

Enclosure

The SPARC T5-1B server module features a compact blade server form factor, giving organizations the flexibility to scale their processing and I/O by simply adding the SPARC T5-1B server module to an existing Sun Blade 6000 chassis (Table 7).

TABLE 7. DIMENSIONS AND WEIGHT OF THE SPARC T5-1B SERVER MODULE

SERVER/DIMENSION	US	INTERNATIONAL
Height	1.75 inches (blade)	44.45 millimeters
Width	12.88 inches	327.15 millimeters
Depth	19.56 inches	496.82 millimeters
Weight (with 2 disks, full memory)	20 pounds	9.1 kilograms

The SPARC T5-1B server module includes the following major components:

- One SPARC T5 processor with sixteen cores operating at 3.6 GHz
- Up to 512 GB of memory in 16 DDR3 DIMM slots (8 GB, 16 GB, and 32 GB DDR3 DIMMs supported)
- Two onboard 10/100/1000 Mb/sec Ethernet ports

- Two dedicated x8 PCIe Express Module slots
- Two x8 PCIe slots for use by optional Fabric Expansion Modules (use with appropriate Network Expansion Modules)
- Three USB 2.0 ports (two external via dongle, one internal that is restricted to the thumb drive)
- Up to two available disk drive slots supporting commodity SAS-2 disk drives
- Oracle ILOM 3.2 system controller

Front and Rear Perspectives

Figure 26 illustrates the front and back panels of the SPARC T5-1B server module.

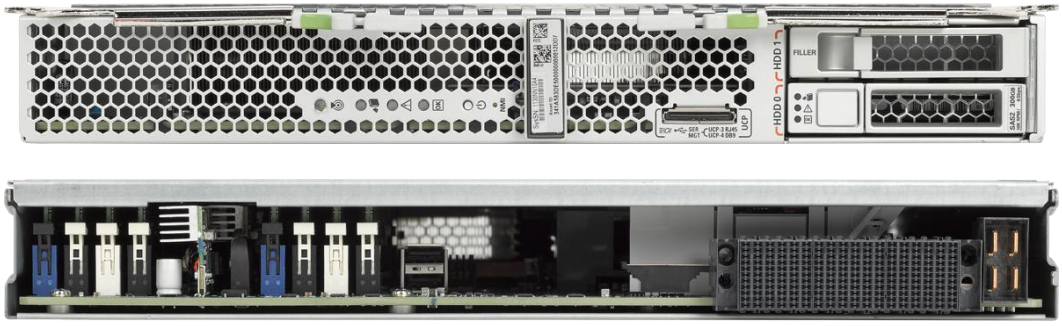

Figure 26. The front and back panels of the SPARC T5-1B server module.

External features of the SPARC T5-1B server module include the following.

- Front component status indicator lights that provide locator (white), service required (amber), and activity status (green) information for the system
- Two hot-pluggable SAS-2 disk drives accessible through the front panel of the system
- Two USB 2.0 ports accessible via dongle
- An HD-15 VGA video port
- Two management ports for use with the Oracle ILOM 3.0 system controller
  - An RJ-45 serial management port (attached via dongle) provides default connection to the Oracle ILOM 3.2 controller.
  - A network management port is connected to the Sun Blade 6000 Chassis Management Module (CMM) switch and accessed via the network management port on the CMM.

Oracle's SPARC T5-2, SPARC T5-4, SPARC T5-8, and SPARC T5-1B Server Architecture
February 2014, Version 2.0
Author: Gary Combs

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

Hardware and Software, Engineered to Work Together