ORACLE®

SUN NETRA

An Oracle White Paper
January 2012, Revised January 2013

# Oracle's Netra SPARC T4-1 and T4-2 Server Architectures

ORACLE®

# Introduction

Many telecommunications providers are moving toward a consolidated framework to converge telecom, Internet, and multimedia services using a single standards-based architecture. However, expanding voice and data services and escalating transaction volumes are putting a strain on existing IT infrastructures.

Even as the IT demand increases, economic realities are reshaping the requirements for the IT infrastructure. Real-estate constraints—along with rapidly rising energy costs—are significant factors that discourage the past practice of adding racks upon racks of servers. The cost and complexity of managing large numbers of systems creates formidable challenges, including low utilization levels that reduce return on investment. Exacting service-level agreements (SLAs) are also increasingly common because service availability is often closely tied to economic success.

In managing the IT infrastructure, organizations are looking to control capital and operational costs more precisely and are making deployment decisions that take into account overall system lifecycle and total cost of ownership (TCO). Oracle's Netra SPARC T4-1 and T4-2 servers can help telecommunications companies address these challenges by providing extreme scalability, energy efficiency, and carrier-grade reliability in a rack-mountable server with a compact form factor. (The Netra SPARC T4-1 server is 2RU and the Netra SPARC T4-2 server is 4RU.)

## Overview of Oracle's Netra SPARC T4-1 and T4-2 Servers

The Netra SPARC T4-1 and T4-2 servers expand on the capabilities of the previous generation Netra SPARC T3-1 server by offering new levels of performance and scalability across a variety of workloads. The servers are based on the new SPARC T4 processor, which provides approximately five times the single-threaded throughput of the previous generation SPARC T3 processor while preserving similar levels of throughput for concurrent, multithreaded applications. The accelerated performance for single-threaded workloads enables telecommunications companies to deliver new kinds of applications or consolidate existing workloads onto fewer servers.

The SPARC T4 processor takes the industry's first massively threaded system on a chip (SoC) to the next level by designing an entirely new core from the ground up. Fifth-generation multicore, multithreading technology supports up to 64 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 a wide range of IT services and application types. Uniformity of management interfaces and adoption of standards also help reduce administrative costs.

In addition to breakthrough performance and scalability, the Netra SPARC T4-1 and T4-2 servers are certified to meet Network Equipment Building Specification (NEBS) Level 3 requirements for environments that demand continuous availability and simplified management. NEBS Level 3 certification (not merely compliance) indicates that the system has undergone a series of rigorous, independently conducted tests to prove that it can continue to operate under severe environmental conditions.

Numerous architectural features—including key redundant and hot-swappable components—enable the Netra SPARC T4-1 and T4-2 servers deliver outstanding levels of availability and reliability that contribute to continuous system operation. Redundant hot-swappable DC or AC power supplies and hot-pluggable disk drives, for example, help to enhance system uptime. A built-in service processor with Integrated Lights Out Manager (ILOM) capabilities enables remote monitoring and system management, providing the kind of reliability and serviceability needed by telecommunications companies or military organizations that operate in severe environments.

Table 1 provides a high-level comparison of the features of the Netra SPARC T4-1 and T4-2 systems.

**TABLE 1. NETRA SPARC T4-1 AND T4-2 SERVER FEATURES**

| FEATURE | NETRA SPARC T4-1 SERVER | NETRA SPARC T4-2 SERVER |
|---|---|---|
| CPUs | • One 8-core 2.85 GHz SPARC T4 processor<br>or<br>• One 4-core 2.85 GHz SPARC T4 processor | • Two 8-core 2.85 GHz SPARC T4 processors |
| Threads | • Up to 64 | • Up to 128 |
| Memory Capacity | • Up to 256 GB (16 GB DDR3 DIMMs) | • Up to 512 GB (16 GB DDR3 DIMMs) |
| Maximum Internal Disk Drives | • Up to 4 HDD (2.5-inch SAS-2 300/600 GB disk drives)<br>• RAID 0/1, (5/6 + BBWC) | • Up to 8 HDD (2.5-inch SAS-2 300/600 GB disk drives)<br>• RAID 0/1 |
| Video | • One HD-15 VGA port | • One HD-15 VGA port |
| Removable, Pluggable I/O | • Slimline DVD+R/-W<br>• Four USB 2.0 ports | • Slimline DVD+R/-W<br>• Four USB 2.0 ports |
| PCI | • Five x8 PCI Express (PCIe) Gen2 slots | • Eight x8 PCIe Gen2 slots<br>• Two x4 PCIe Gen2 slots |
| Ethernet | • Four onboard Gigabit Ethernet (GbE) ports (10/100/1000)<br>• Two 10 GbE ports via XAUI combo slots (shared with PCIe) | • Four onboard GbE ports (10/100/1000)<br>• Four 10 GbE ports with XAUI card installed |
| Power Supplies | • Two hot-swappable AC or DC 1200 W power supplies<br>• N+1 redundancy | • Four hot-swappable AC or DC 1200 W power supplies<br>• N+1 redundancy |
| Fans | • Five hot-swappable fan modules, with counterrotating fans per module<br>• N+1 redundancy | • Four hot-swappable fan trays, with counterrotating fans per module<br>• N+2 redundancy |

# SPARC T4 Processor

The SPARC T4 processor eliminates the need for expensive custom hardware and software development by integrating computing, security, and I/O onto a single chip. It achieves binary compatibility with earlier SPARC processors, and no other processor delivers so much performance in so little space and with such small power requirements. It enables telecommunications providers to rapidly scale the delivery of new network services with maximum efficiency and predictability.

This section provides an overview of the SPARC T4 processor architecture. Additional details can be found on the Web at:

- http://www.oracle.com/us/products/servers-storage/servers/sparc-enterprise/t-series/resources/

- http://www.oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/o11-090-sparc-t4-arch-496245.pdf

- http://www.oracle.com/us/corporate/innovation/deep-dive-hetherington-489126.html

## SPARC T4 Processor Architecture

Each SPARC T4 processor provides eight cores (four cores for the Netra SPARC T4-1 server) with each core able to switch between up to eight threads (64 threads per processor) using a modified LRU (Least Recently Used) algorithm for thread choice. In addition, each core provides two integer execution pipelines, so that a single SPARC core is capable of executing two threads at a time. Unlike the SPARC T3 processor, the SPARC T4 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 2-instruction "decode groups" and proceed through Decode, Rename, and Pick stages before proceeding to Issue, after which they are sent to one of four subsequent execution pipelines, depending upon the type of instruction to be performed.

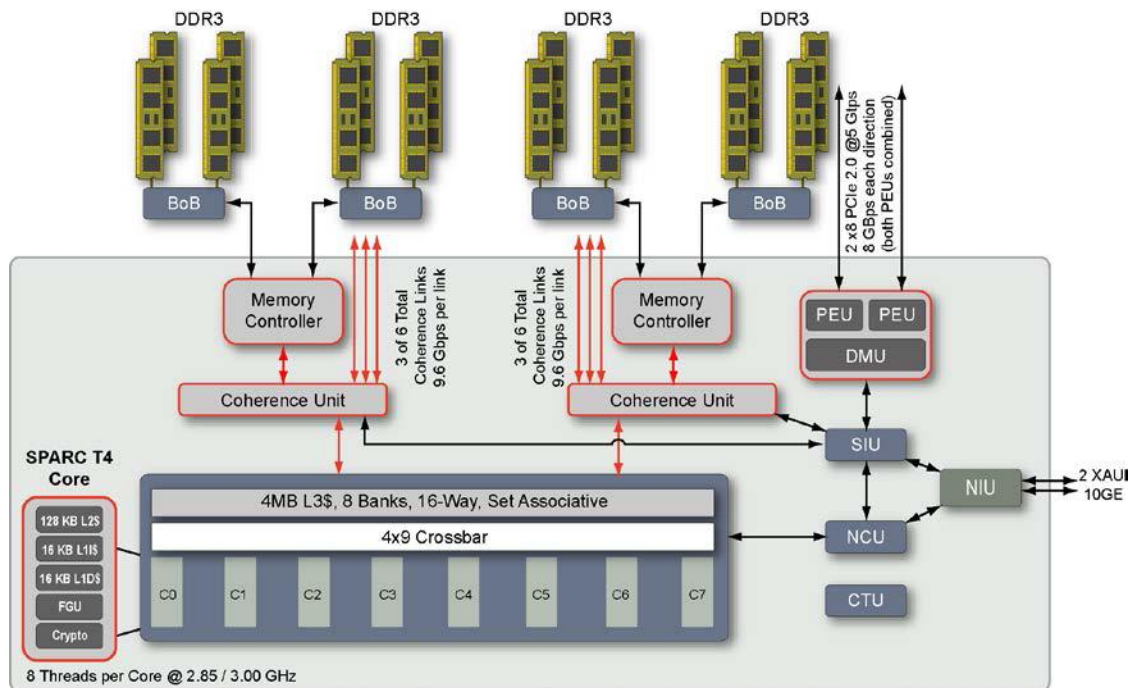Figure 1 provides a block-level diagram of the SPARC T4 processor.

Figure 1. The SPARC T4 processor provides six coherence links to connect to up to four other processors.

The SPARC T4 processor has coherence link interfaces to allow communication between up to four SPARC T4 processors in a system without requiring any external hub chip. There are six coherence links, each with 14 bits in each direction running at 9.6 Gb/sec. Each frame has 168 bits, so the maximum frame rate is 800 M frames per second. The SPARC T4 has two coherence link controllers. Each includes two Coherence and Ordering Units (COU), three Link Framing Units (LFU), and a cross bar (CLX) between COUs and LFUs. Each COU interfaces to two L2 bank pairs. The coherence links run a cache coherence (snoopy) protocol over an FB-DIMM-like physical interface. The memory link speed of the SPARC T4 is maintained at 6.4 Gb/sec, identical to that of the SPARC T3.

The SPARC T4 processor can support one-, two- and four-socket implementations. The newly designed core for the SPARC T4 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 eight cores of a SPARC T4 processor (Figure 2).
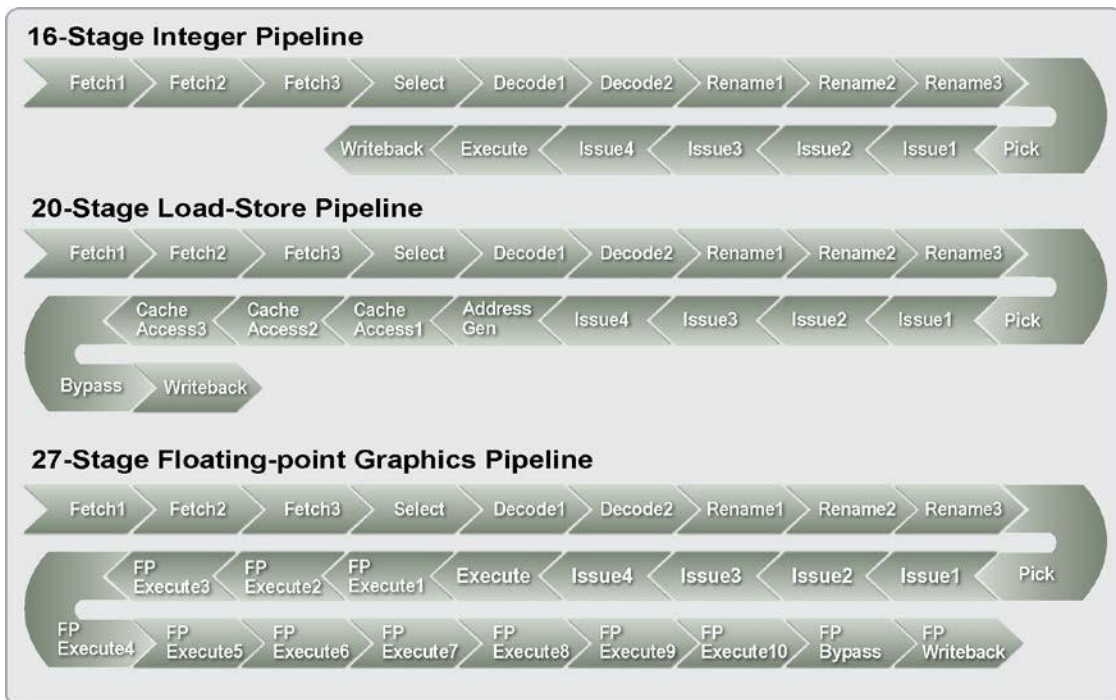
Figure 2. A 16-stage integer pipeline, a 20-stage load-store pipeline, and a 27-stage floating-point graphics pipeline are provided by each SPARC T4 processor core.

**SPARC T4 Processor Cache Architecture**

The SPARC T4 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 might be in the same physical system. SPARC T4 has Level 1 caches that consist of separate data and instruction caches. Both are 16 KB and are per core. A single Level 2 cache, again per core, is 128 KB. The Level 3 cache is *shared* across all eight cores of the SPARC T4 processor and is 4 MB, has eight banks, and is 16-way set associative. Figure 3 illustrates the relationship between the L2 and L3 caches and shows them connected by a 4x9 crossbar.
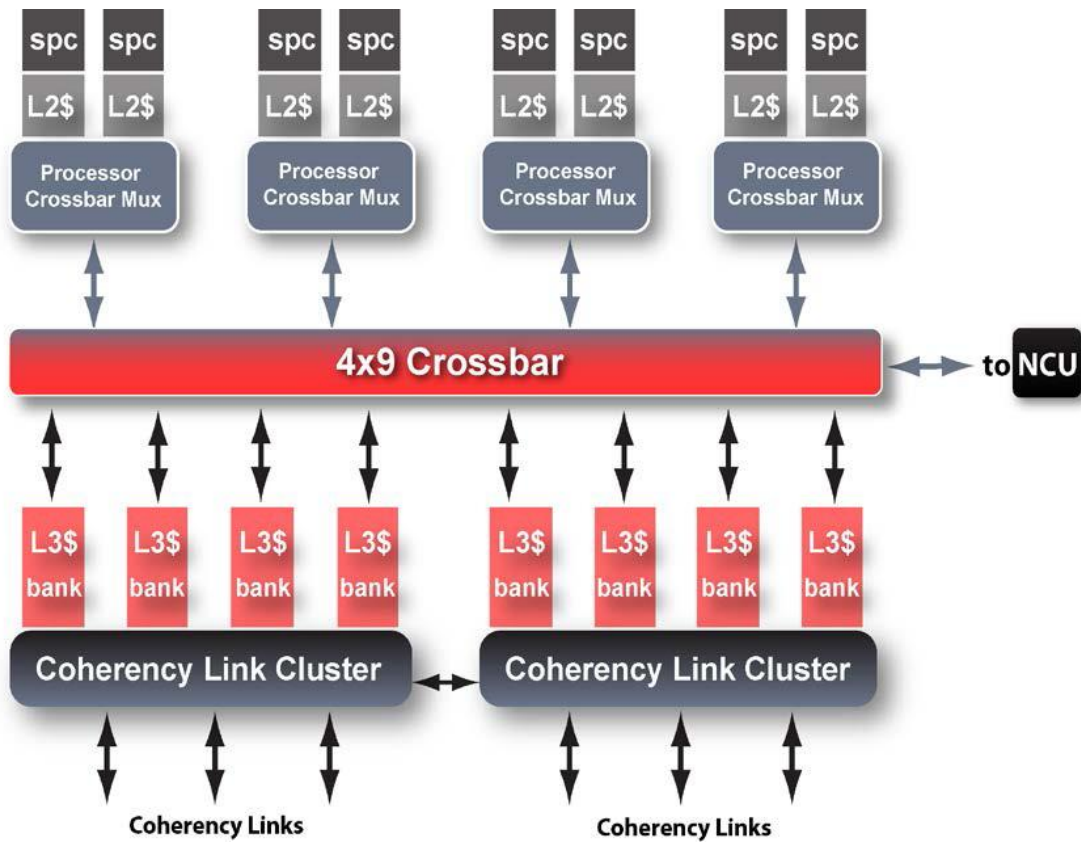
Figure 3. The relationship between Level 2 and Level 3 caches.

**SPARC T4 Core Architecture**

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

- OOO (Out-of-Order) instruction execution

- Sophisticated branch prediction

- Prefetching of both instructions and data

- Much deeper pipelines (relative to previous versions of multicore processors from Sun or Oracle)

- Three levels of cache

- Support for a much larger Memory Management Unit page size (2 GB)

- Multiple instruction issue

All these characteristics in the SPARC T4 have yielded improvements in single-thread performance by five times while retaining networking and throughput performance equal to that of previous multicore processors from Sun or Oracle.

There are many functional units, pipelines, and associated details that are present within the SPARC T4 core but beyond the scope of this paper. However, this paper (due to the significantly new characteristics and features of the SPARC T4 core) does attempt to touch upon the major exposed (that is, having visibility to either programmers or users of SPARC T4 processor-based systems) features or characteristics.

One aspect by which the designers of the SPARC T4 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.

### Integrated Networking

By providing integrated on-chip networking, both the SPARC T3 and SPARC T4 processors provide integrated on-chip networking, enabling better networking performance than earlier processors, such as the UltraSPARC T2+. All network data is supplied directly from and to main memory. Placing networking so close to memory reduces latency, provides higher memory bandwidth, and eliminates inherent inefficiencies of I/O protocol translation. The SPARC T4 processor provides two 10 GbE ports with integrated serializer/deserializer (SerDes), offering line-rate packet classification at up to 30 million packets/second (based on layer 14 of the protocol stack). Multiple DMA engines (16 transmit and 16 receive DMA channels) match DMAs to individual threads, providing binding flexibility between ports and threads. Virtualization support includes provisions for eight partitions, and interrupts can be bound to different hardware threads.

### Stream-Processing Unit (SPU)

Each SPARC T4 processor core contains a stream processing unit (SPU) that provides cryptographic processing. This functionality has been implemented within the core pipelines themselves and is accessible by 29 new user-level instructions.

The SPU operates at the same clock speed as the core and supports the following cryptographic algorithms:

- DES/3DES
- AES-128/192/256
- Kasumi, Camellia
- MD5
- SHA-1
- SHA-256

- SHA-512

- RSA via MPMUL/MONTMUL/MONTSQR instructions

The SPU is designed to achieve wire-speed encryption and decryption on the processor's 10 GbE ports. A cryptographic algorithm (that is supported in hardware from the group previously listed) actually uses parts of the FGU and the integer pipelines.

**Integral PCI Express Generation 2 Support**

SPARC T4 processors provide dual on-chip PCIe Generation 2 interfaces. Each operates at 5 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 4 GB/sec, with a maximum payload size of 256 bytes per PCIe Gen2 interface. The actual realizable bandwidth is more likely to be approximately 2.8 GB/sec. An x8 SerDes interface is provided for integration with off-chip PCIe switches.

**Power Management**

Beyond the inherent efficiencies of Oracle's multicore/multithreaded design, the SPARC T4 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 ability to turn off clocks in both cores and memory to reduce power consumption.

Substantial innovation is present in the areas of:

- Limiting speculation, such as conditional branches not taken

- Extensive clock gating in the data path, control blocks, and arrays

- Power throttling that allows extra stall cycles to be injected into the decode stage

## System-Level Architecture for Netra SPARC T4-1 and T4-2 Servers

The system-on-a-chip (SoC) design of the SPARC T4 processor means that sophisticated system-level functionality can be accomplished with a minimum of high-quality components.

Although each system deploys the same technologies (SPARC T4 processor, DDR3 Memory, PCIe Generation 2 switches, SAS-2), a separate motherboard design is used in Oracle's Netra SPARC T4-1 and T4-2 servers. Common features of the Netra SPARC T4 motherboard are:

- A minimum of one socket for a SPARC T4 processor

- Memory slots to supply memory for the Netra SPARC T4-1 or T4-2 server

- A remote-to-local memory latency ratio of 1.47 (in the case of a processor system running the lmbench read access test)

The sections that follow describe the architecture of components that are common to both the Netra SPARC T4-1 and Netra SPARC T4-2 servers.

## Memory Subsystem

In Oracle's Netra SPARC T4-1 and T4-2 servers, the SPARC T4 processor provides on-chip memory controllers that communicate indirectly to DDR3 DIMMs via newly designed Buffers-on-Board (BoB) memory interfaces through four high-speed serial links. Two dual-channel memory controller units (MCUs) are provided on the SPARC T4 processor. Each MCU can transfer data at an aggregate rate of 6.4 Gb/sec. There are 16 motherboard memory socket locations on the Netra SPARC T4-1 server motherboard. The Netra SPARC T4-2 motherboard has 32 DIMM slots broken up over four memory risers, each of which plug into the motherboard.

## I/O Subsystem

Oracle's SPARC T4 processor provides dual 10 Gigabit Ethernet Attachment Unit Interfaces (XAUI) ports as well as PCIe Gen2 ports that enable connection to PCIe card slots, Express Module slots, or bridge devices that interface with PCIe. The PCIe Gen2 ports are capable of operating at 5 GB/sec bidirectionally and natively interface to I/O devices through PCIe Gen2 switch chips. Bridge devices that connect to these PCIe slots include:

- **Disk controller.** Disk control is managed by an LSI Logic SAS2008 SAS/SATA controller chip. RAID levels 0 and 1 are supported. The LSI controller chip also drives the DVD (optical drive) in Oracle's Netra SPARC T4-1 and T4-2 servers.

- **GbE.** Oracle's Netra SPARC T4-1 and T4-2 servers provide four 10/100/1000 Mb/sec Ethernet interfaces on the rear of each chassis.

- **USB.** A single-lane PCIe port connects to a PCI bridge device on the Netra SPARC T4-1 and T4-2 servers. A second bridge chip converts the 33-bit 66 MHz PCI bus into multiple USB 2.0 ports.

## System and Component Serviceability

Finding and identifying servers and components in a modern telecommunications provider data center can be challenging. Oracle's Netra SPARC T4-1 and T4-2 servers are optimized for lights-out data center configurations with easy-to-identify servers and modules. Color-coded operator panels provide easy-to-understand diagnostics and systems are designed for deployment in hot-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-plug 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

Oracle's Netra SPARC T4-1 and T4-2 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 Netra SPARC T4-1 and T4-2 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 Oracle's Netra SPARC T4-1 and T4-2 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 PCBs.

- The Oracle Netra SPARC T4-1 and T4-2 servers both support USB 2.0 interfaces, two each on the front and back of the chassis. The servers also have an internal USB port.

### Cooling

The power and cooling efficiency of Netra SPARC T4-1 and T4-2 servers exceeds that of many competitive carrier-grade systems configured with similar processing, memory, and storage capacities. The effective front-to-back airflow design helps to lower component temperatures, reducing the number of fans needed to cool the system. The chassis design features hex-shaped, honeycombed air inlet holes that enhance airflow and provide EMI shielding.

The chassis is divided into two distinct airflow chambers. One chamber draws airflow for the motherboard (CPUs, RAM) and PCI trays. The other chamber provides airflow for the hard drives, DVD-RW device (if present), and power supplies. The power supply/storage chamber is cooled by individual fans on the back of each power supply. The motherboard/PCI chamber is cooled by a row of fans mounted in front of the server behind the bezel.

Variable-speed fans run under the control of the onboard service processor, which monitors processor temperatures and system ambient air temperature. Based on these readings, the fans operate at the lowest speeds possible to provide sufficient cooling—conserving power usage, prolonging fan life, and reducing acoustical noise.

Fan failures are indicated by an amber "Service Required" LED on the front panel. Fan modules are designed for redundancy—a backup fan enables system continuity in the event of a fan failure. The fans are also hot-swappable so that a module with a failed fan can be removed and a new fan module can be inserted without shutting down the system.

## Power and Thermal

Engineered for high availability as well as low energy consumption, the Netra SPARC T4-1 and T4-2 servers are equipped with highly efficient, redundant, hot-swappable AC or DC power supplies with separate power cords. The power supplies are rated at 1200 watts and can auto-detect between 120/240V and 50/60Hz.

One power supply is sufficient to run a fully populated Netra SPARC T4-1 server and two power supplies are sufficient to run a fully populated Netra SPARC T4-2 server. Since the Netra SPARC T4-1 server supports two power supplies and the Netra SPARC T4-2 server supports four power supplies, the have N+1 and N+2 redundancy, respectively. For maximum protection against power supply failures, dual power supplies should be installed at all times in the Netra SPARC T4-1 server and all four power supplies should be installed in the Netra SPARC T4-2 server at all times. In normal operation, the power supplies share the power demands of the system equally.

To further reduce power requirements and to meet NEBS certification, the Netra SPARC T4-1 and T4-2 servers can be run on DC power. Using DC power reduces overall operating costs by lowering energy use, reducing heat, and increasing reliability.

Power consumption metrics for Netra SPARC T4-1 and T4-2 servers are defined in Table 2.

TABLE 2. POWER SPECIFICATIONS FOR NETRA SPARC T4-1 AND T4-2 SERVERS

| FEATURE | NETRA SPARC T4-1 SERVER | NETRA SPARC T4-2 SERVER |
|---|---|---|
| Power Supplies | <ul><li>Two hot-swappable AC or DC 1200 W power supplies</li><li>N+1 redundancy</li></ul> | <ul><li>Four hot-swappable AC or DC 1200 W power supplies</li><li>N+2 redundancy</li></ul> |
| Threads | <ul><li>Up to 64</li></ul> | <ul><li>Up to 128</li></ul> |
| AC Power-Maximum operating input current | <ul><li>8.57A @ 100 VAC</li><li>4.2A @ 200 VAC</li></ul> | <ul><li>7.6A @ 200 VAC</li></ul> |
| AC Power-Maximum operating input power at 200 VAC | <ul><li>920 W</li></ul> | <ul><li>1451.7 W</li></ul> |
| DC Power-Power input | <ul><li>48 VDC or 68 VDC (nominal)</li><li>40 VDC to 75 VDC (range)</li></ul> | <ul><li>48 VDC or 68 VDC (nominal)</li><li>40 VDC to 75 VDC (range)</li></ul> |

For additional information about power requirements, readers can utilize Oracle's power calculators, which are available at http://www.oracle.com/us/products/servers-storage/sun-power-calculators/index.html.

**Rackmounting**

The Netra SPARC T4-1 and T4-2 servers ship with a hard rackmount 19-inch 4-post kit. Optional kits are available for 19-inch two-post, 23-inch two-post, and 600 x 600 millimeter hard rackmount, as well as 19-inch 4-post slide rackmount.

# Netra SPARC T4-1 Server Architecture

The compact Netra SPARC T4-1 server provides significant computational power in a space-efficient, low-power 2RU rackmount package. With excellent price-to-performance ratios, low acquisition cost, and tightly integrated high-performance 10 GbE, this server is ideally suited to the delivery of horizontally scaled transaction and Web services that require extreme network performance. The server is designed to address the challenges of modern telecommunications data centers with greatly reduced power consumption and a small physical footprint.

## Front and Rear Perspectives

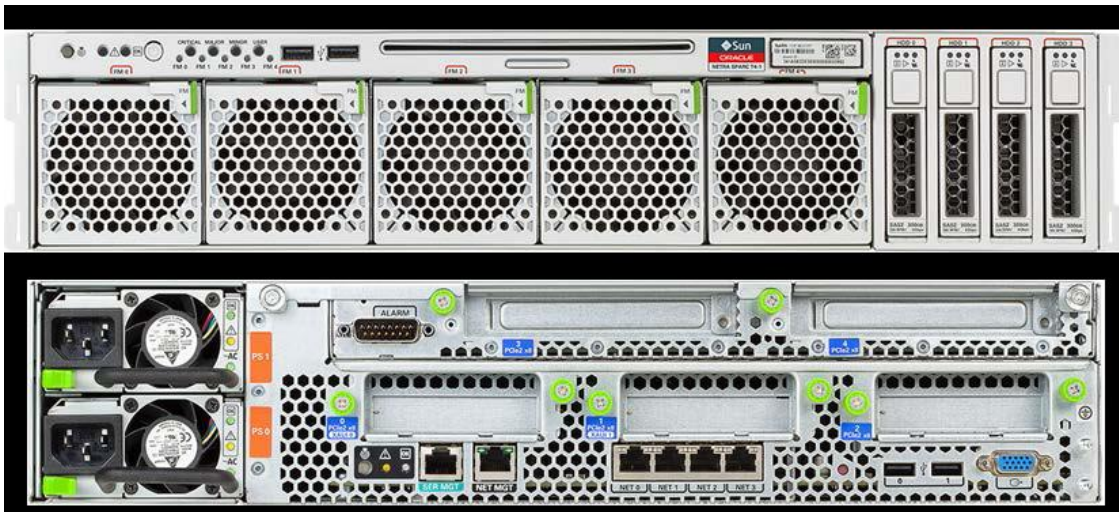Figure 4 illustrates the front and rear panels of the Netra SPARC T4-1 server, showing a 4-disk backplane.



Figure 4. The front and rear panels of the Netra SPARC T4-1 server.

## Enclosure

The 2RU Netra SPARC T4-1 server enclosure is described in Table 3.

**TABLE 3. DIMENSIONS AND WEIGHT OF THE NETRA SPARC T4-1 SERVER**

| DIMENSION | U.S. | INTERNATIONAL |
|---|---|---|
| Height | 3.43 inches (2RU) | 88.65 millimeters |
| Width | 17.52 inches | 445 millimeters |
| Depth: Max to PSU handles | 20.71 inches | 526 millimeters |
| Depth: Max to rear I/O | 19.72 inches | 501 millimeters |
| Weight (approximate, fully configured without PCIe cards or rackmounts) | 41 pounds | 18.6 kilograms |

The Netra SPARC T4-1 server includes the following major components and external features:

- One SPARC T4 processor with eight cores operating at 2.85 GHz or one SPARC T4 processor with four cores operating at 2.85 GHz

- Up to 256 GB of memory in 16 DDR3 DIMM slots (8GB, and 16GB DDR3 DIMMs supported)

- Up to four hot-pluggable SAS-2 disk drives that are insertable through the front panel

- Four onboard 10/100/1000 Mb/sec Ethernet ports

- Supported low-profile and full-height, half-length PCIe. Two slots (x8) with two combination XAUI or low-profile PCIe x8 slots

- Four USB 2.0 ports (two on the front panel and two on the rear)

- Two (N+1) hot-swappable, high-efficiency 1200 watt AC or DC power supplies with integral plugs and fans (insertable from the rear)

- Five fan assemblies (each with two fans) under environmental monitoring and control; N+1 redundancy. Fans are accessed through a dedicated top panel door.

- One Slimline, slot-accessible DVD+R/-W accessible through the front panel

- One HD-15 VGA video port

- Front and rear system and component status indicator lights that provide locator (white), service required (amber), and activity status (green) for the system

- ILOM 3.0 system controller

- Two management ports for use with the ILOM 3.0 system controller. An RJ-45 serial management port provides default connection to the ILOM 3.0 controller and a network management port supports an optional RJ-45 10/100Base-T connection to the ILOM 3.0 system controller.

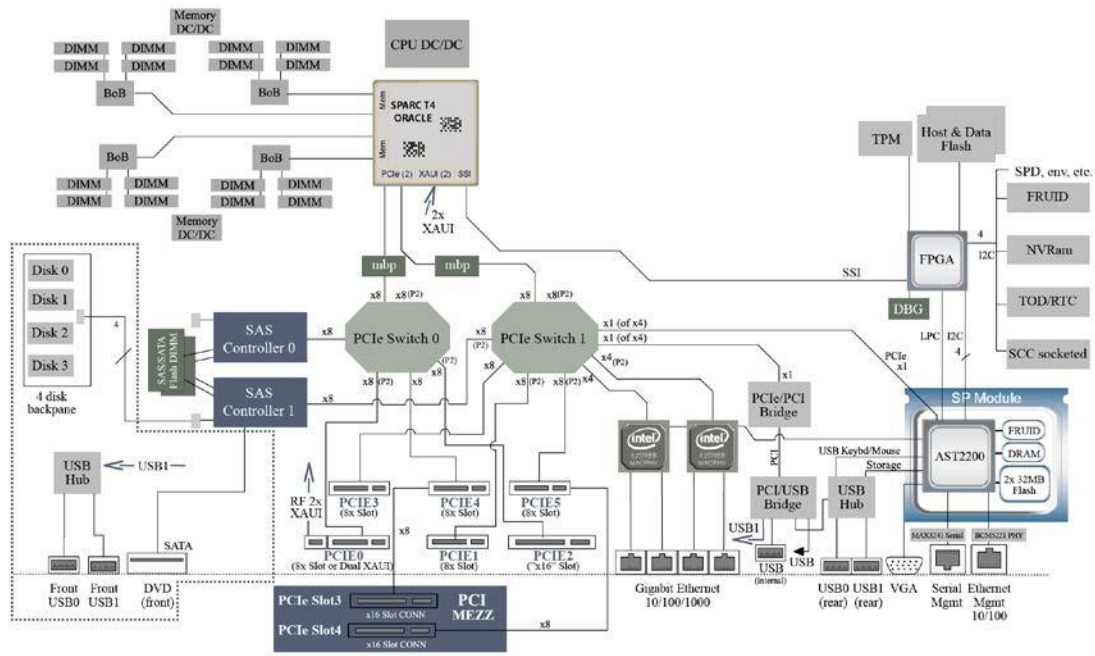The Netra SPARC T4-1 server motherboard design is shown in Figure 5.



Figure 5. Netra SPARC T4-1 server motherboard design.

## Netra SPARC T4-2 Server Architecture

The expandable Netra SPARC T4-2 server is optimized to deliver transaction and telecommunications services, including OSS/BSS technology application services as well as enterprise application services. With considerable expansion capabilities and integrated virtualization technologies, the Netra SPARC T4-2 server is also an ideal platform for consolidated tier 1 and tier 2 workloads.

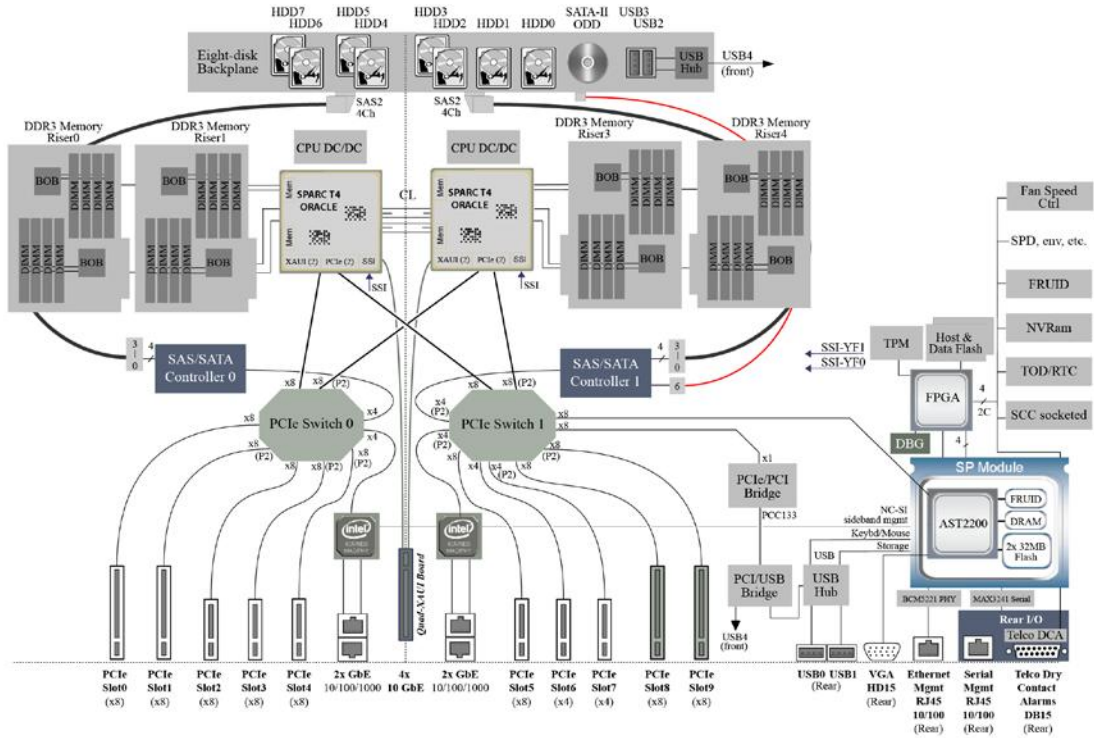The Netra SPARC T4-2 server motherboard design is shown in Figure 6.

Figure 6. Netra SPARC T4-2 server motherboard design.

## Enclosure

Netra SPARC T4-2 servers feature a compact, expandable 4RU rackmount chassis, giving telecommunications companies the flexibility to scale processing and I/O requirements without wasting precious space (Table 4).

**TABLE 4. DIMENSIONS AND WEIGHT OF THE NETRA SPARC T4-2 SERVER**

| DIMENSION | U.S. | INTERNATIONAL |
|---|---|---|
| Height | 6.86 inches (4 RU) | 174.5 millimeters |
| Width | 17.18 inches | 445 millimeters |
| Depth: Max to PSU handles | 20.87 inches | 530 millimeters |
| Depth: Max to rear I/O | 19.88 inches | 505 millimeters |
| Weight (approximate, fully configured without PCIe cards or rackmounts) | 72.6 pounds | 33 kilograms |

The Netra SPARC T4-2 server includes the following major components and external features:

- Dual SPARC T4 processors with eight cores per processor operating at 2.85 GHz

- Up to 512 GB of memory in 32 DDR3 DIMM slots ( 8 GB, and 16 GB DDR3 DIMMs)

- Up to eight hot-pluggable SAS-2 disk drives insertable through the front panel of the system

- Four onboard 10/100/1000 Mb/sec Ethernet ports

- Eight dedicated low-profile PCIe slots and two full-height half-length PCIe slots

- One dedicated XAUI slot for 4x 10 GbE connections (XAUI port cannot be shared with any PCIe card)

- Four USB 2.0 ports (two on the front panel and two on the rear)

- Four (N+2) hot-pluggable/hot-swappable high-efficiency 1200 watt AC or DC power supplies with integral plugs and fans (insertable from the rear)

- Four fan assemblies under environmental monitoring and control; N+2 redundancy

- One Slimline DVD +R/W drive accessible through the front panel

- One HD-15 VGA video port

- Front and rear system and component status indicator lights that provide locator (white), service required (amber), and activity status (green) for the system

- ILOM 3.0 system controller

- Two management ports for use with the ILOM 3.0 system controller. An RJ-45 serial management port provides default connection to the ILOM 3.0 controller and a network management port supports an optional RJ-45 10/100Base-T connection to the ILOM 3.0 system controller.

### Front and Rear Perspectives

Figure 7 illustrates the front and back panels of the Netra SPARC T4-2 server.
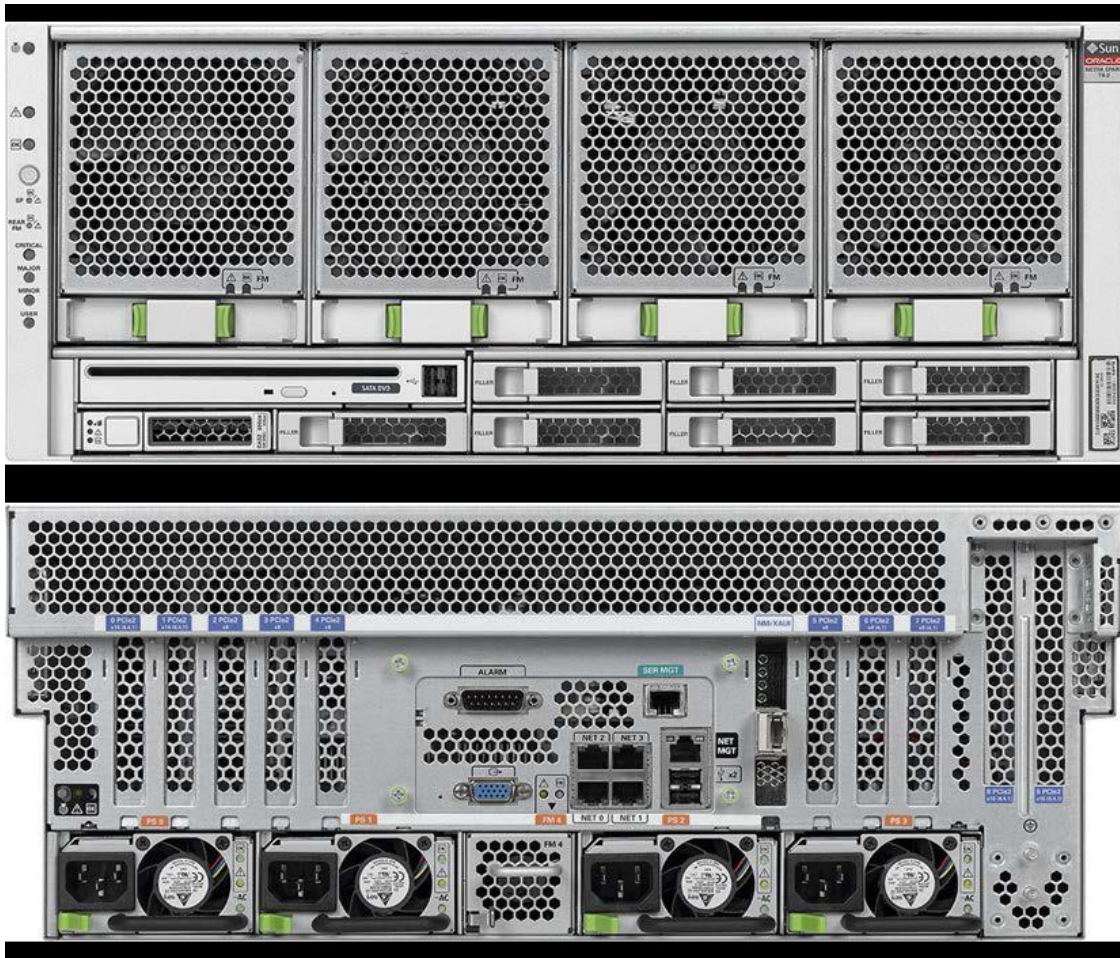


Figure 7. The front and back panels of the Netra SPARC T4-2 server.

## Common RAS Features

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

The Netra SPARC T4-1 and T4-2 servers are engineered for hardware failure prevention, near-continuous operation, fast recovery, and easy serviceability. RAS features for these systems include the following:

- **Reduced component count.** Integration of key functionality into the SPARC T4 processor means fewer components and reduced incidence of component failures to increase overall availability.

- **Hot-swappable redundant components.** Mirrored disks, redundant fan modules, and redundant power supply units can be quickly and easily changed out, increasing system uptime.

- **Accessible components for improved serviceability.** Front-accessible, hot-swappable disk drives can be replaced quickly. The optional DVD-RW drive can also be removed without opening the top cover of the chassis. Fan modules and power supply units can be replaced without completely removing a system from the rack.

- **Indicator LEDs on the back of the chassis.** Easily visible LEDs allow problems to be identified and isolated easily. Diagnostic LEDs are also included on the motherboard.

- **Integrated out-of-band management.** Standard on the Netra SPARC T4-1 and T4-2 servers at no additional charge, ILOM provides powerful tools for remote system management, simplifying administrative tasks, reducing the number of on-site personnel that are needed, and lowering the cost of operations.

- **NEBS Level 3 certification.** Telecordia NEBS Level 3 certification demonstrates that the servers meet stringent requirements for extremely reliable operation, even in harsh environmental conditions.

## Enterprise Scalability and Availability with Oracle Solaris

The SPARC T4 processor technology is mature and well tested, and Netra SPARC T4-1 and T4-2 servers share binary compatibility with earlier SPARC systems. The systems run with the solid and secure foundation of Oracle Solaris 11 or Oracle Solaris 10 8/11 or higher. Moreover, Oracle Solaris includes a variety of sophisticated tools that let organizations easily consolidate and manage workloads while taking advantage of processor technology innovation and availability features that are part of Oracle Solaris.

The subsections that follow provide additional information about some of the most relevant characteristics of Oracle Solaris for telecommunications providers. Additional information about Oracle Solaris capabilities for carrier-grade applications can be found in the white paper, "Oracle Solaris: The Carrier-Grade Operating System," which is available at http://www.oracle.com/technetwork/server-storage/solaris11/documentation/solaris-thecarriergradeos-wp-308726.pdf.

## Scalability and Support for Multicore/Multithreaded Technology

Oracle Solaris 10 and Oracle Solaris 11 are specifically designed to deliver the advanced capabilities of SPARC T4 processor–based systems. In fact, Oracle Solaris provides key functionality for virtualization, optimal utilization, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments.

One significant advantage of a system based on the SPARC T4 processor is that it appears as a familiar symmetric multiprocessing (SMP) system to Oracle Solaris and applications. In addition, Oracle Solaris 10 and 11 have incorporated many features that help improve application performance on multicore/multithreaded architectures:

- **Accelerated Cryptography.** Accelerated cryptography is supported through the cryptographic framework in Oracle Solaris as well as the SPARC T4 processor. The SPARC T4 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 both 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.

- **Critical Thread Optimization.** The new SPARC T4 processor also supports a "critical thread API" that is new in Oracle Solaris 11, enabling the operating system to recognize critical threads and assign them, by themselves, to a single processor core. This allows critical threads to run at the very highest performance levels without competing with other less-critical threads, resulting in faster overall performance for threaded applications.

- **Multicore/multithreaded awareness.** Oracle Solaris 10 and 11 are aware of the SPARC T4 processor hierarchy so that the scheduler can effectively balance the load across all the available pipelines. Even though it exposes each of these processors as 64 logical processors, Oracle Solaris understands the correlation between cores and the threads they support and provides a fast and efficient thread implementation.

- **Fine-granularity manageability.** For the SPARC T4 processor, Oracle Solaris 10 and 11 have the ability to enable or disable individual cores and threads (logical processors). In addition, standard Oracle Solaris features, such as processor sets, provide the ability to define a group of logical processors and schedule processes or threads on them.

- **Binding interfaces.** Oracle Solaris allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set 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 T4 processor, including support for the on-chip 10 GbE ports and 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 T4 processor on the Netra SPARC T4-2 server, this implementation represents 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 technology that can specifically help to decrease the impact of NUMA on applications and improve performance on NUMA architectures:

  - *Memory Placement Optimization*. Oracle Solaris 10 and 11 use Memory Placement Optimization (MPO) to improve the placement of memory across the physical memory of a server, resulting in increased performance. Through MPO, Oracle Solaris works to help 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*. Hierarchical Lgroup Support (HLS) improves the MPO feature in Oracle Solaris. HLS helps Oracle Solaris optimize performance for systems with more-complex memory latency hierarchies. HLS lets Oracle Solaris distinguish between the degrees of memory remoteness, allocating resources with the lowest-possible latency for applications. If local resources are not available by default for a given application, HLS helps Oracle Solaris allocate the nearest remote resources.

- **Oracle Solaris ZFS.** Oracle Solaris ZFS offers a dramatic advance in data management, automating and consolidating complicated storage administration concepts and providing unlimited scalability with the world's first 128-bit file system. Oracle Solaris ZFS is based on a transactional object model that removes most of the traditional constraints 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 T4 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 of the system's resources.

## Important New Features in Oracle Solaris 11

Oracle Solaris 11 provides access to advanced Oracle Solaris features that have been in development for more than five years. Two new Oracle Solaris 11 features are noteworthy for telecommunications providers:

- **Network-based package management.** The Oracle Solaris 11 Image Packaging System (IPS) is a new network-based package management system designed to greatly decrease planned system downtime and provide for completely safe system updates and upgrades. IPS offers a framework for complete software lifecycle management such as installation, upgrade, and removal of software packages. Administrators can install software from network-based package repositories with full automatic dependency checking for any additional libraries that might be required during a software package install.

- **Network virtualization**. A new network stack architecture, also known as the Crossbow project, has been introduced in Oracle Solaris 11. The architecture offers improvements in performance, observability, security, and ease of use. Among other features, it enables administrators to create Virtual NICs (VNICs) and virtual switches, resulting in a virtual network infrastructure that is completely independent from the underlying hardware. This gives IT organizations the ability to deploy a "data center in a box" solution that brings increased levels of flexibility and TCO savings, much like a larger data center consolidation project. Network virtualization also gives telecommunications carriers a method to control network bandwidth and network traffic flow for different services deployed on the system. Thus, services that are very sensitive to network latency can get their share of network bandwidth before others, and the physical network can be spared from being flooded with unwanted traffic levels.

## Fault Management and Predictive Self-Healing

Oracle Solaris 10 introduced a new architecture for building and deploying systems and services capable of fault management and predictive self-healing. Oracle Solaris Predictive Self Healing is an innovative capability in Oracle Solaris 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 configuration problems.

Oracle Solaris Predictive Self-Healing and Oracle Solaris fault management provide the following specific capabilities on Netra SPARC T4-1 and T4-2 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.

- 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 FB-DIMM has been replaced, in which case affected pages are cleared from the resource cache.

- I/O Retirement logs errors and faults.

- fmlog logs faults detected by the system.

## High Availability with Oracle Solaris Cluster

Mission-critical network infrastructure applications in telecommunications often require clustered physical servers to avoid single points of failure and the potential for service interruptions or data loss. Oracle Solaris Cluster is designed to provide the highest availability for applications running on Oracle Solaris. It also allows telecommunication providers to combine high availability clustering with virtualization to offer increased service levels at lower cost.

Oracle Solaris Cluster enables:

- Subsecond failure detection and lower service recovery time through Oracle Solaris kernel-level integration

- Optimized resource use along with high availability and fault isolation by supporting both Oracle VM Server for SPARC and Oracle Solaris Zones

- Consolidation of multiple applications within the same cluster of physical servers using virtual nodes or virtual clusters

- Disaster recovery through multisite, multicluster configurations

For more information on Oracle Solaris Cluster refer to http://www.oracle.com/technetwork/server-storage/solaris-cluster/.

# End-to-End Virtualization Technology

Virtualization technology is increasingly popular as organizations strive to consolidate disparate workloads onto fewer but more-powerful systems to improve utilization. Netra SPARC T4-1 and T4-2 servers are specifically designed for virtualization, providing very fine-grained division of multiple resources—from processing to virtualized networking and I/O. Most importantly, Oracle's virtualization technology is provided as a part of the system, not an expensive add-on.

## A Multithreaded Hypervisor

Like the earlier generations of UltraSPARC and SPARC processors, the SPARC T4 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.

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.

**Oracle VM Server for SPARC**

Supported in all Oracle servers using multicore/multithreaded technology, Oracle VM Server for SPARC provides full virtual machines that can run an independent OS 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 and 11 are written to the hypervisor, which provides a stable, idealized, and virtualizable representation of the underlying server hardware to the operating system in each virtual machine.

By taking advantage of virtual machines, organizations gain the flexibility to deploy multiple OS instances simultaneously on a single platform. In addition, administrators can leverage virtual device capabilities to transport an entire software stack hosted on a virtual machine from one physical machine to another. Virtual machines can also host Oracle Solaris Zones to capture the isolation, flexibility, and manageability features of both technologies.

To support virtualized networking, Oracle VM Server for SPARC implements a virtual Layer 2 switch, to which guest virtual machines (guest OSs) can be connected. Each guest virtual machine can be connected to multiple vSwitches, and multiple guest virtual machines can also be connected to the same vSwitch. vSwitches can be associated with a real physical network port, or they can exist without an associated port, in which case the vSwitch provides only communication between virtual machines in the same server, thus saving valuable network resources.

**Oracle Solaris Zones**

Providing virtualization at the OS level, Oracle Solaris Zones (known as Oracle Solaris Containers in Oracle Solaris 10 and earlier versions) 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. Oracle Solaris Zones can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability because processes in one zone are prevented from interfering with processes running in another zone.

Resource management tools provided with Oracle Solaris help allocate resources, such as CPUs, to specific applications. CPUs in a multiprocessor system (or threads in the SPARC T4 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 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.

## System Management Technology

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

Netra SPARC T4-1 and T4-2 servers can be managed through an integrated lights out management service as well as through Oracle Enterprise Manager Ops Center, an enterprise-level management solution.

## Oracle Integrated Lights Out Manager

The Oracle Integrated Lights Out Manager (ILOM) service processor, which is provided across many of Oracle's x64 servers, acts as a system controller, simplifying the remote management and administration of Netra SPARC T4-1 and T4-2 servers. The service processor is similar in implementation to the implementation used in other Oracle modular and rackmount x64 servers. As a result, these servers integrate easily with existing management infrastructures, including on-site enterprise management frameworks or element managers.

The service processor offers the following capabilities:

- Provides IPMI 2.0-compliant management functions to the server module's firmware, OS, and applications and to IPMI-based management tools that access the service processor via the ILOM Ethernet management interface. It provides visibility of environmental sensors, both on the server module and elsewhere in the chassis.

- Manages inventory and environmental controls for the server module, including CPUs, FB-DIMMs, and power supplies and supplies HTTPS, CLI, or SNMP access to this data.

- Supplies remote textual console interfaces.

- Provides a means to download upgrades to all system firmware.

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

Oracle Integrated Lights Out Manager monitors the following Netra SPARC T4-1 and T4-2 server conditions:

- CPU temperature conditions

- Hard drive presence

- Enclosure thermal conditions

- Fan speed and status

- Power supply status

- Voltage conditions

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

## Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center is infrastructure lifecycle management software for deploying, monitoring, patching, and managing large and small installations of Oracle systems. Oracle Enterprise Manager Ops Center 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 Oracle Enterprise Manager Ops Center to discover bare-metal systems based on a given subnet address or IP range.

- **Group.** Given the number of systems to manage and the constant repurposing of systems, it is critical for IT organizations to find ways to group resources together. Oracle Enterprise Manager Ops Center 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.** Oracle Enterprise Manager Ops Center remotely installs OSs (Oracle Solaris, Red Hat Linux, 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 lifecycle continues, Oracle Enterprise Manager Ops Center can update firmware and provision software packages and patches to selected systems.

- **Monitor.** When systems are up and running, administrators can use Oracle Enterprise Manager Ops Center to monitor system health, helping 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. Oracle Enterprise Manager Ops Center also monitors OS attributes such as swap space, CPU, memory, and file systems. Administrators can define specific threshold levels and set preferred notification methods, including e-mail, pager, or Simple Network Management Protocol (SNMP) traps for each monitored component, as business needs demand.

- **Manage.** Telecommunications companies require that infrastructure lifecycle management extend beyond just deploying and monitoring systems. Oracle Enterprise Manager Ops Center includes lights-out management capabilities, such as powering systems on and off, and remote serial console access to help IT organizations manage their IT infrastructures from remote locations. Leveraging Oracle Enterprise Manager Ops Center software's Role-Based Access Control (RBAC) feature, organizations can grant permissions to specific users to perform specific management tasks.

- **Hybrid user interface.** Oracle Enterprise Manager Ops Center offers users a hybrid user interface (UI) accessible from the Web, which 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.

# Multicore/Multithread Tools Enable Fast Performance and Rapid Time to Market

No matter how compelling new hardware or OS platforms might be, telecommunications organizations must be assured that the costs and risks of adoption are in line with the rewards. In particular, telecommunications organizations want to be able to continue to leverage the considerable advantages of popular commercial and open source software. Developers don't want to have to switch compilers and basic development tools. Administrators can scarcely afford a more-complex support matrix or more time spent getting applications to run effectively in a new environment. Previous generations of multicore/multithreaded servers have relied upon individual tools that provide recommendations as to whether an existing application would execute with maximum efficiency on a multicore/multithreaded server. Developers could employ a variety of individual tools to optimize applications for execution within a multicore/multithreaded environment. However, Oracle, in its ongoing efforts to produce a tightly integrated software platform, has integrated much of what were separate tools into Oracle Solaris and Oracle Solaris Studio.

Oracle Solaris Studio delivers a comprehensive development platform for building scalable, secure, and reliable applications for multicore/multithreaded servers. For more information about Oracle Solaris Studio and Oracle developer tools, visit the following links:

http://www.oracle.com/technetwork/server-storage/solarisstudio/overview/

http://www.oracle.com/us/products/tools/

## Accelerating Development and Reducing Costs with Sun Netra Data Plane Suite

Oracle's Sun Netra Data Plane Suite can help telecommunications carriers accelerate development and reduce R&D and sustaining costs by exploiting the multicore/multithreaded architecture of the SPARC T4 processor for line rate packet processing performance. Comprised of an open unified development environment and a light-weight runtime environment, Sun Netra Data Plane Suite enables better economics through rapid development. It can also leverage virtualization capabilities in the hypervisor using Oracle VM Server for SPARC (previously called Sun Logical Domains or LDoms) for radical consolidation of control and data plane functions into a single unified architecture.

# Conclusion

Oracle's Netra SPARC T4-1 and T4-2 servers combine breakthrough performance and energy efficiency with unprecedented processor core, memory, and I/O densities to help telecommunications companies scale their network infrastructures in a cost-effective manner. The SPARC T4 processor used in Netra SPARC T4-1 and T4-2 servers provides the industry's next generation massively threaded system on a chip. It delivers equivalent network throughput and efficiency compared to previous generation Netra SPARC T3-1 servers while also offering five times the single-threaded performance gains.

With up to 64 threads per processor, on-chip memory management, two 10 GbE interfaces, two PCIe Generation 2 root complexes, and on-chip cryptographic acceleration, the SPARC T4 processor fundamentally redefines the capabilities of a modern processor. By incorporating cache coherency for multiprocessor support, SPARC T4 processors allow these capabilities to be multiplied incrementally.

Oracle's Netra SPARC T4-1 and T4-2 servers 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 highly compact IT infrastructure that can truly scale to meet today's telecommunications challenges and budget constraints.