An Oracle White Paper
February 2011

Enabling End-to-End 10 Gigabit Ethernet in Oracle's Sun Netra ATCA Product Family

Executive Overview

This paper provides an overview of Oracle's 10 Gigabit Ethernet-enabled Advanced Telecom Computing Architecture (ATCA) platforms and the spectrum of available Oracle Sun blades. It outlines the full ecosystem of ATCA products, highlighting how 10 GbE Ethernet is integrated into the Oracle ATCA blade family. The paper provides an architectural overview of the SPARC T3, UltraSPARC® T2, Intel® Xeon®, and AMD Opteron™ processor-powered ATCA blades, the 10 GbE Ethernet ATCA switch, the 10 GbE Ethernet System Platform, available Advanced Rear Transition Modules, and carrier-grade software that can be used to enhance and manage ATCA platforms. An additional appendix discusses considerations for power planning.

Introduction

Communication carriers and Network Equipment Providers (NEPs) are adopting IP networking to support virtually all of their operations, from wireline and mobile communication to rich media, content delivery, and the IP Multimedia Subsystem (IMS). The rapid adoption of VoIP, Internet Protocol Television (IPTV), and Video on Demand is causing IP packet processing workloads to escalate, putting new demands on NEP infrastructures. New access technologies for wireless networks, such as Worldwide Interoperability for Microwave Access (WiMAX), High-Speed Downlink Packet Access (HSDPA), and Long Term Evolution (LTE) of Universal Terrestrial Radio Access Network, increase the bandwidth available to mobile devices and intensify the need for fast, high-capacity IP packet processing solutions. Packets and sessions are quickly replacing circuits and calls, rapidly changing the requirements for telecommunication network infrastructures.

Packet processing at the network's edge requires ever-deeper packet analysis of numerous parallel streams of traffic, and at line speeds of 10 GbE. In the past, the traditional approach to processing such a massive stream of data was to build custom hardware and to engineer

custom ASICs — an approach that drove up costs, increased program risks, and made equipment designs immutable.

Today, 10 GbE Ethernet is integrated throughout Oracle's family of Sun ATCA blade products, bringing unprecedented parallelism, bandwidth, choice, and flexibility. Beginning with 10 Gpbs Ethernet on the Sun Netra CT900 ATCA Blade Server, Oracle adds 10 GbE switching to the server's extended fabric. 10 GbE Ethernet is engineered in Sun server blade offerings, including the blade server based on the UltraSPARC T2 and SPARC T3 processor with CoolThreads technology, blade servers based on dual-socketed quad-core Intel Xeon processors, and blade servers based on a multi-core AMD Opteron processor.

The broad spectrum of available blade processor architectures allows customers to achieve an optimal match of application workloads with computing platforms. The ability of the SPARC T3 processor to execute 96 threads in parallel means that NEPs can develop deep packet-processing applications for the Data Plane on a low-cost, general-purpose compute server, processing 10 GbE flows at line speeds and eliminating the need for custom hardware. Sun x86 blade products enable choice between Intel Xeon and AMD Opteron architectures, and support high-performance x86 applications on Oracle Solaris, carrier-grade Linux, or Windows Advanced Server. The newest Sun blade server offering from Oracle — base on the SPARC T3 — is engineered for high scalability while conserving power consumption, enabling a full-service configuration of 96 compute threads and 32GB within a 200-watt power envelope.

Chapter 1

An Ecosystem Delivering Bandwidth and Flexibility

Oracle's Sun Netra CT900 ATCA Server offers an industry-standard platform for hosting today's applications and for supporting tomorrow's next-generation network infrastructure and services. The Sun Netra CT900 Server is a NEBS Level 3-certified, rack-mountable, -48v powered ATCA blade system that supports PICMG 3.1 Options 1 and 9. As an integrated platform, it enables some of the most highly available applications for the telecommunications market. In the larger marketplace, the ATCA PICMG organization has more than 450 supporting companies delivering products and participating in the standard.

With a long history of supporting communication carriers and NEPs that develop applications for them, the Sun ATCA product line is one of the broadest lines of standards-based, carrier-grade, and commercially-available general-purpose computing platforms available anywhere. In the ATCA market, the Sun Netra CT900 Blade Server supports ATCA processor blades based on the SPARC T3, UltraSPARC T2, UltraSPARC T1, Intel Xeon, and AMD Opteron processors, with support for a spectrum of operating environments — Oracle Solaris, carrier-grade Linux, and/or Windows Advanced Server.

Today, 10 GbE Ethernet is integrated into the full Sun ATCA product line to provide communication carriers and NEPs with a complete, end-to-end solution. Carriers can move applications from proprietary servers onto a more standard, compact form factor with or without the use of virtualization. NEPs can apply the SPARC T3 and the UltraSPARC T2 processor's concurrent threads to Data Plane applications and the balanced processing capabilities of multi-core Intel Xeon or AMD Opteron blades for IMS Control Layer and other communication carrier applications. Product benefits include:

- *10 Gigabit Ethernet*

  10 GbE Ethernet is integrated throughout the ATCA product line at the processor, chip, switch, and midplane level. The UltraSPARC T2 and SPARC T3 processor are true system on a chip, with both a PCI-Express (PCIe) and a 10 GbE Ethernet controller integrated into the processor's memory architecture, speeding data flow from main memory to the network and allowing extremely tight coupling between processor threads and multiple flows supported by the 10 GbE interface. The 10 GbE Ethernet implementation is also embodied in a standard Sun ASIC that brings the same sophisticated processing capabilities to certain other Sun Netra x86 blade products. In addition, the Intel® 82599EB 10 Gigabit Ethernet Controller is used in the Sun Netra CP3270 Intel Xeon ATCA Blade Server to implement extended fabric capabilities. Oracle also offers a Sun ATCA hub switch blade that switches across the PICMG 3.1 Option 9 interconnect.

- *Unprecedented Choice*

  Sun ATCA blades support 10 GbE Ethernet processing through a high degree of parallelism via the SPARC T3 and UltraSPARC T2 processor or through extreme levels of single-threaded performance using multi-core Intel Xeon processors and AMD Opteron processors. Communication carriers and NEPs can select processor blades that best match application needs.

- *Virtualization and Consolidation*

  Built-in virtualization and consolidation capabilities in Sun blades allow a rack's worth of applications to run on a single shelf. The UltraSPARC T2 and SPARC T3 processors support Oracle VM for SPARC (formerly Sun Logical Domains), a virtualization technology that allows multiple OS instances and their applications to run safely on a single ATCA processor blade. For Oracle Solaris applications running on SPARC or x86 processor-based blades, Oracle Solaris Containers support multiple virtual operating environments within the same OS instance.

  Virtualization offers one way to consolidate applications onto the same blade, helping to make more effective use of processing power while reducing costs. The processing and I/O capabilities of 10 Gigabit Ethernet-enabled Sun ATCA blades, combined with their reliability and NEBS compliance, makes them ideal platforms on which applications can be moved from rack-mount servers onto a more compact form-factor, with or without virtualization.

10 GbE Ethernet-Enabled ATCA Blades

Oracle's Sun 10 GbE Ethernet-enabled ATCA blade offerings include the Netra SPARC T3-1BA SPARC T3 ATCA blade, Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server, the Sun Netra CP3270 Intel Xeon ATCA Blade Server, the Sun Netra CP3250 Intel Xeon ATCA Blade Server, the Sun Netra CP3220 Opteron ATCA Blade Server, and the Sun Netra CP3240 10 GbE ATCA Switch Blade.

ATCA Processor Blades

Sun Netra ATCA CP32x0 processor blades support the PICMG 3.0 and 3.1 specifications and can be plugged into any of the user node slots of the Sun Netra CT900 ATCA Server. The only two slots that cannot accommodate processor blades are the switch/hub slots. Sun Netra ATCA processor blades can also be plugged into any ATCA/PICMG 3.0 and 3.1-compliant chassis. The processor blades support the basic PICMG 3.0 system management features and are compatible with other ATCA-compliant blades. Common features of the processor blades include:

- PICMG 3.0 and PICMG 3.1 Option 1 and 9 support
- 10/100/1000 Mbps Ethernet out-of-band management ports front and rear
- Craft RJ-45 serial ports with Cisco pinouts
- Dual USB ports
- Large memory capacities to support memory-intensive applications
- Common Advanced Rear Transition Module (ARTM) pinouts
- Redundant Flash support (8-32 GB Compact Flash socket)
- Single ATCA slot
- Oracle Solaris 10 support

Oracle Netra SPARC T3-1BA SPARC T3 ATCA Blade Server

The Oracle Netra SPARC T3-1BA Blade Server is based on the SPARC T3 processor with 12 Cores and 96 Threads, which integrates PCIe gen2 and dual XAUI Ethernet into this system-on-a-chip (see Figure 1). This ATCA blade can be used to support numerous communication carrier applications, including the following:

- General-purpose computing using Oracle Solaris
- Support for multiple applications, each in its own Oracle Solaris Container
- Oracle VM for SPARC, which can support up to 96 isolated hardware partitions, each running its own OS instance

This processor blade is discussed in detail in "Oracle Netra SPARC T3-1BA SPARC T3 ATCA Blade Server Architecture."

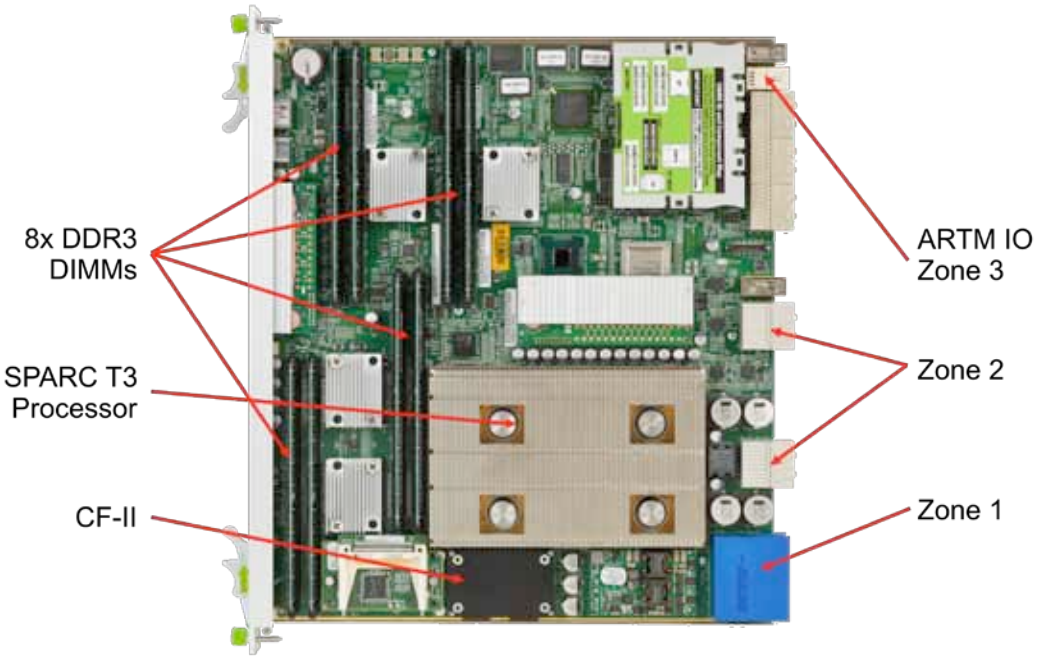

Figure 1. The Oracle Netra SPARC T3 ATCA Blade Server features the SPARC T3 processor with 96 concurrent threads.

Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server

The Sun Netra CP3260 Blade Server is based on the UltraSPARC T2 processor, which integrates PCIe and dual 10 Gigabit Ethernet into this system-on-a-chip (see Figure 2). This ATCA blade can be used to support numerous communication carrier applications, including the following:

- General-purpose computing using Oracle Solaris
- Support for multiple applications, each in its own Oracle Solaris Container

- Oracle VM for SPARC (formerly Sun Logical Domains), which can support up to 64 isolated hardware partitions, each running its own OS instance
- The Netra Data Plane Software Suite, which can be used to dedicate each processor thread to a different phase of packet processing, turning this general-purpose platform into a powerful packet processing engine

This processor blade is discussed in detail in section labeled, "Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server Architecture."

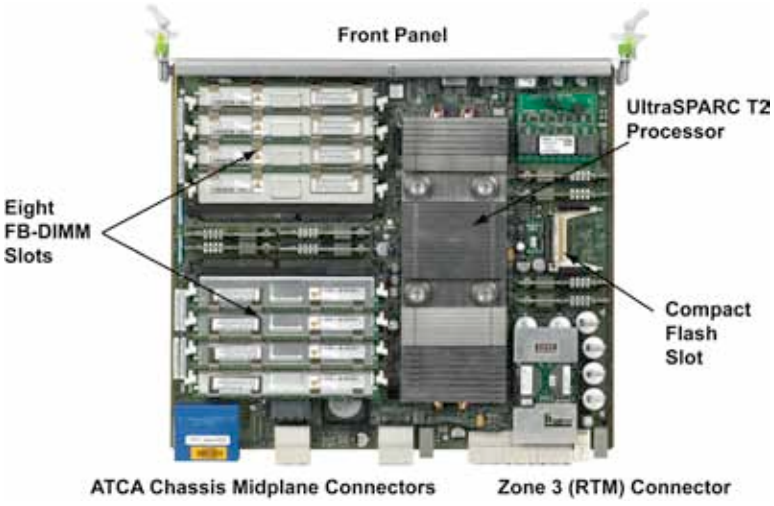

Figure 2. The Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server features the UltraSPARC T2 processor with 64 concurrent threads.

Sun Netra CP3270 Intel Xeon ATCA Blade Server

Engineered with next-generation embedded-class Intel Xeon processor technology, the Sun Netra CP3270 Intel Xeon ATCA Blade Server is designed to minimize power consumption while delivering high performance. This second-generation Intel processor-based blade supports two Intel Xeon LC5518 processors per blade (see Figure 3). These processors follow the next-generation Intel Core Microarchitecture (codenamed "Nehalem"), with enhancements such as Intel QuickPath Interconnect technology and Intel HyperThreading. In addition, this embedded class processor design integrates a DDR3 memory controller and PCIe 2.0 functionality that is normally resident in a Northbridge Memory Controller Hub (MCH) directly into the processor, reducing memory and I/O latency as well as blade power requirements.

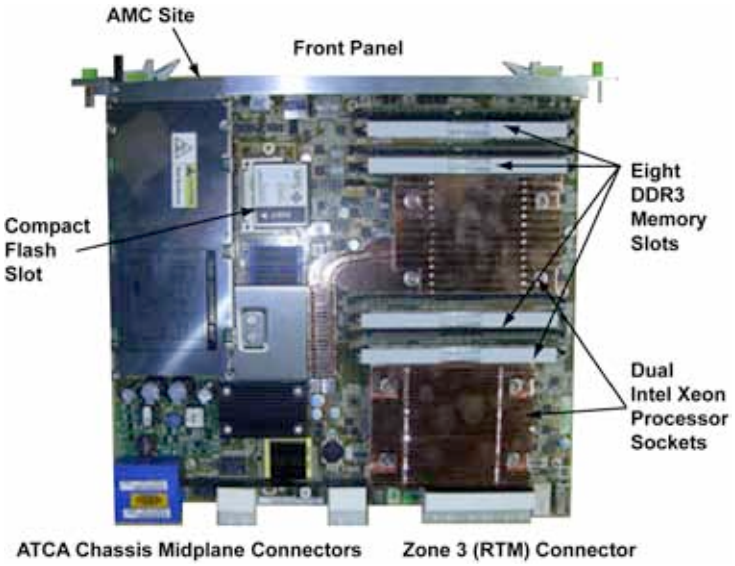

Figure 3. The Sun Netra CP3270 Intel Xeon ATCA Blade Server features dual-socketed Intel Xeon LC5518 processors with 16 threads and up to 32 GB of DDR3 memory.

This blade supports Oracle Solaris, carrier-grade Linux, Windows 2008 Advanced Server, and virtualization through Oracle Solaris Containers. Its design is discussed in more detail in the section labeled, "Sun Netra CP3270 Intel Xeon ATCA Blade Server Architecture."

Sun Netra CP3250 Intel Xeon ATCA Blade Server

The Sun Netra CP3250 Intel Xeon ATCA Blade Server is based on a non-throttling embedded class of quad-core Intel Xeon processors, with two Intel Xeon L5408-LV processor sockets per blade (see Figure 4). In this processor blade, Intel Xeon L5408-LV processors incorporate a low power design that minimizes power use and the need for cooling. This blade supports Oracle Solaris, carrier-grade Linux, Windows 2008 Advanced Server, and virtualization through Oracle Solaris Containers. The blade design is discussed in more detail in "Sun Netra CP3250 Intel Xeon ATCA Blade Server Architecture."

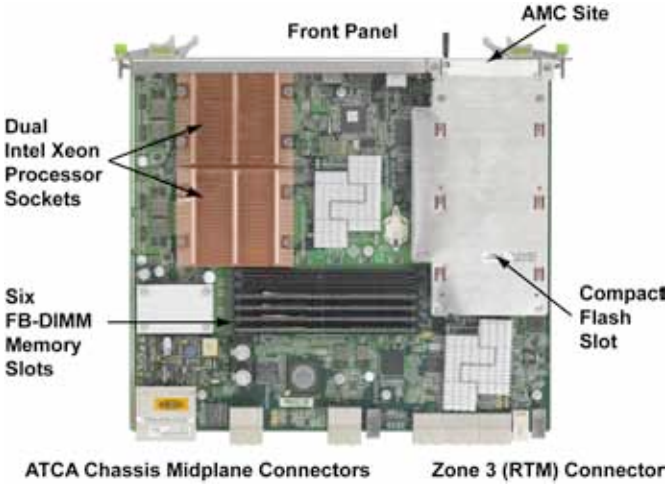

Figure 4. The Sun Netra CP3250 Intel Xeon ATCA Blade Server features dual-socketed quad-core Intel Xeon L5408-LV processors.

Sun Netra CP3220 Opteron Rev. F ATCA Blade Server

The Sun Netra CP3220 Opteron Rev. F ATCA Blade Server is based on AMD Opteron Rev. F processors, with a single socket per blade (see Figure 5), in dual-core or quad-core models. This blade supports Oracle Solaris, carrier-grade Linux, Windows 2003 or 2008 Advanced Server, and virtualization through Oracle Solaris Containers and AMD's hypervisor. This processor blade is discussed in detail in "Sun Netra CP3220 AMD Opteron ATCA Blade Architecture."

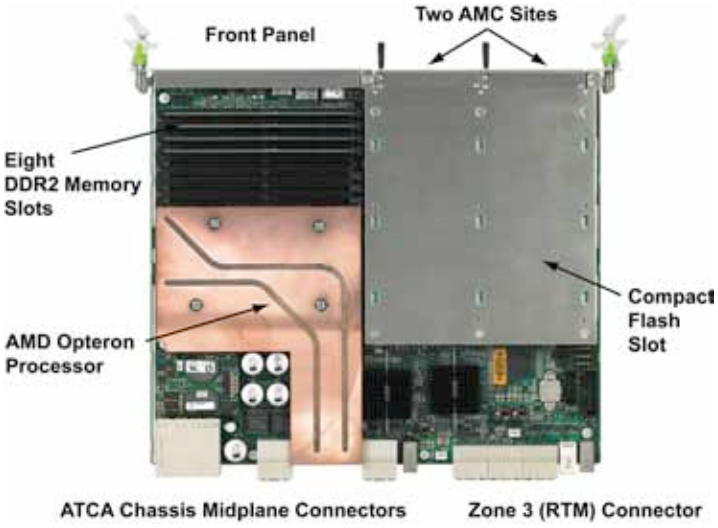

Figure 5. The Sun Netra CP3220 Opteron Rev. F ATCA Blade Server features a multi-core AMD Opteron processor and I/O flexibility for balanced performance.

Sun Netra CP3240 ATCA 10 GbE ATCA Switch Blade

The Sun Netra CP3240 10 GbE ATCA Switch Blade brings 10 Gigabit Ethernet to the Sun Netra CT900 ATCA Server (see Figure 6). The switch blade supports both the Gigabit Ethernet base and the 10 Gigabit Ethernet extended fabric, supporting both existing and 10 GbE Ethernet-enabled processor blades. The switch is capable of line-speed Layer 2 switching and Layer 3 routing with full IPv6 support. This switch blade is discussed in detail in "Sun Netra CP3240 ATCA 10 GbE ATCA Switch Blade."

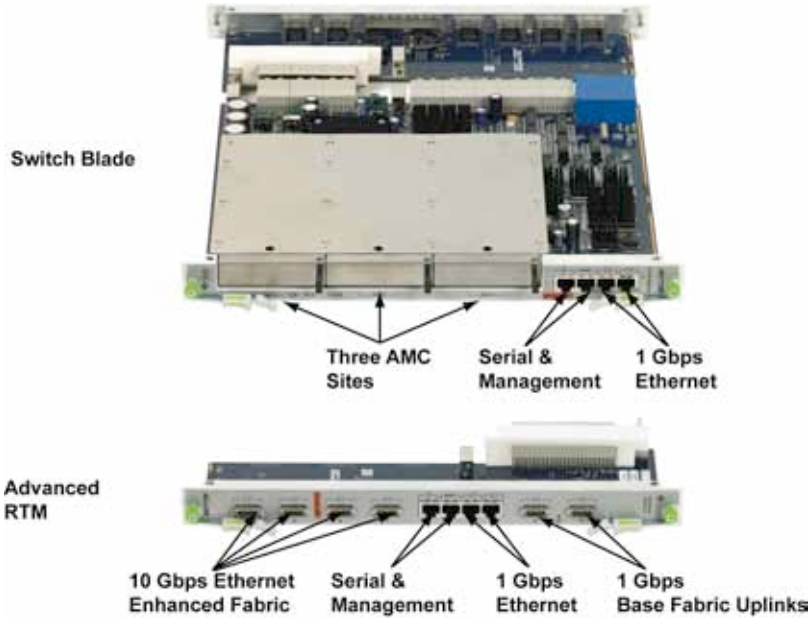

Figure 6. The Sun Netra CP3240 ATCA 10 GbE ATCA Switch Blade brings 10 Gigabit Ethernet to the ATCA chassis' enhanced fabric. A dedicated Advanced Rear Transition Module provides external connectivity.

Advanced Rear Transition Modules

The I/O capabilities of Sun Netra CP32x0 processor blades can be expanded through the use of optional Advanced Rear Transition Modules (RTMs). These modules use the Zone 3 connector to provide expanded I/O capacity through the rear of the Sun Netra CP900 ATCA Server (see Figure 7). This allows carriers and NEPs to tailor I/O capabilities of their Sun Netra processor blades to match requirements closely of supported applications. Carriers and NEPs can configure I/O cabling once and then swap processor blades in and out independently. The Sun Netra CP3200 series Advanced RTMs are designed to work with Sun Netra CP3220, CP3250, CP3270, CP3260 and SPARC T3-1BA Blade Servers and any other ATCA blades that adhere to the Zone 3 specification.

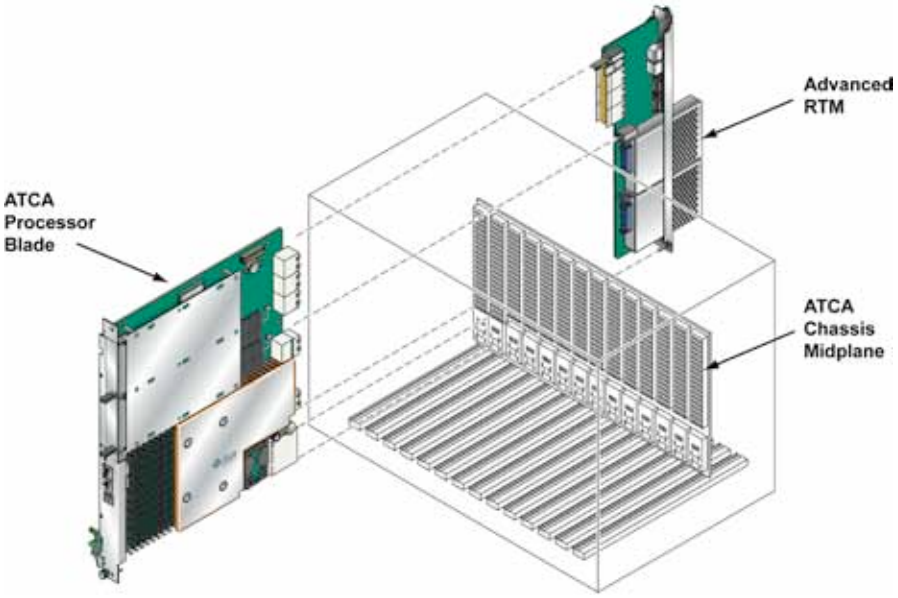

Figure 7. Advanced RTMs connect to Sun Netra CP32x0 Processor Blade Servers through the Zone 3 connector, bringing increased I/O capabilities to these 10 Gigabit Ethernet-enabled processor blades.

Oracle's engineering team has been working with the ATCA community to standardize the Zone 3 connector pinout and to enable other innovative I/O solutions for the ATCA market. Depending on the processor blade, the connector carries some or all of the following interfaces between the processor blade and Advanced RTM:

- 8-lane PCIe
- XAUI (for 10 Gigabit Ethernet)
- SAS (for disk I/O)
- Serial I/O (craft interface)
- Advanced Mezzanine Card (AMC) I/O
- Gigabit Ethernet interfaces
- Internal management buses

Oracle offers Advanced RTMs for Sun ATCA server blades that can enhance I/O capabilities in these areas: multiple Gigabit Ethernet links, multiple 10 Gigabit Ethernet connections, hard disk capacity through Serial-Attached SCSI (SAS) connections, and balanced access to external Fibre Channel disk storage. Table 1 summarizes the available Advanced RTM choices.

TABLE 1. NETRA CP3200 ADVANCED RTMS COMPATIBLE WITH SUN NETRA CP32X0 BLADE SERVERS

FEATURE	NETRA CP3200 ARTM-10G	NETRA CP3200 ARTM-HDD	NETRA CP3200 ARTM-FC
Primary I/O	2x 10 GbE Ethernet	Up to 2x 2.5" SAS disk	2x 4 Gbps Fibre Channel
Secondary I/O	2x 10 GbE Ethernet (With Netra CP3260 UltraSPARC T2 ATCA Blade only)	2x SAS External 2x to AMC slots (With Netra CP3220 Opteron ATCA Blade only)	
Networking	2x Gigabit Ethernet		6x Gigabit Ethernet
Management	1x Gigabit Ethernet	1x Gigabit Ethernet	1x Gigabit Ethernet
Craft	Serial craft interface	Serial craft interface	Serial craft interface
Compatibility	CP32x0 processor blades	CP32x0 processor blades	CP32x0 processor blades

Netra CP3200 ARTM-10G

While the Netra CP3200 ATCA 10 GbE ATCA Switch Blade provides interfaces that bring the extended fabric beyond the system platform, the Netra CP3200 ARTM-10G (see Figure 8) provides a mechanism for bringing multiple 10 GbE Ethernet paths directly from the processor blades. The Netra CP3200 ARTM-10G provides up to four 10 GbE Ethernet interfaces and two 1 Gbps Ethernet interfaces to support intensive network throughput.

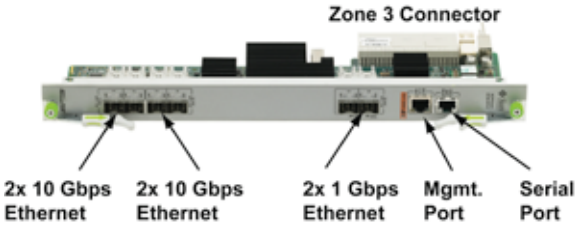

Figure 8. The Netra CP3200 ARTM-10G provides up to four 10 GbE Ethernet interfaces and two 1 Gbps Ethernet interfaces.

The Advanced RTM uses the Zone 3 connector's 8-lane PCIe interface to drive the standard Sun 10 Gigabit Ethernet ASIC. It supports multiple concurrent DMA streams and line-speed packet classification as discussed in "10 Gigabit Ethernet Throughput." The ASIC provides two SFP+ ports supporting 10 GbE Ethernet, and two Small Form Factor Pluggable (SFP) ports supporting 1 Gbps Ethernet.

When used with the Sun Netra CP3260 ATCA UltraSPARC T2 ATCA Blade Server, the Advanced RTM can support up to two additional 10 GbE Ethernet ports through an additional pair of SFP+ ports. This processor blade can connect each of its two on-board 10 GbE Ethernet connections to either the Advanced RTM or to the enhanced fabric through the Zone 3 connector. Each port can be programmed independently, bringing even more flexibility to the platform.

In addition to the interfaces discussed above, the Netra CP3200 ARTM-10G provides a Gigabit Ethernet management interface and a serial craft interface.

Netra CP3200 ARTM-HDD

This Advanced RTM incorporates a six-port PCIe SAS controller that uses the card's x8 PCIe interface (see Figure 9). The Netra CP3200 ARTM-HDD supports two internal, 2.5" SAS hard disk drives with built-in RAID 0 capability. It incorporates two egress SAS ports so that each disk drive can be dual ported and accessed from another Advanced RTM. This provides hardware support through clustering with software such as the Sun Netra High Availability Suite. When used with the Sun Netra CP3220 ATCA Opteron ATCA Blade Server, two SAS ports are translated to the processor blade's Advanced Mezzanine Card (AMC) slots, where two additional AMC-based disk drives can be installed and integrated with the storage system.

Processor blades can be booted from Advanced RTM disk drives. The ARTM-HDD is also compatible with Oracle's NEBS Level 3-certified storage including StorageTek 2530 Arrays.

In addition to providing on-board disk storage and connectivity, the ARTM-HDD provides one Gigabit Ethernet management interface and one serial craft port.

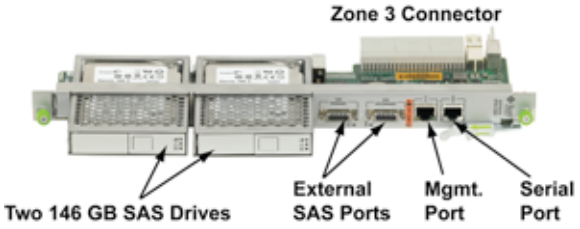

Figure 9. The Netra CP3200 ARTM-HDD makes dual SAS disk drives available to processor blades, as well as external SAS ports and connectivity to internal AMC slots.

Netra CP3200 ARTM-FC

The Netra CP3200 ARTM-FC (see Figure 10) includes an Emulex dual Fibre Channel interface that utilizes the four lanes of the x8 PCIe connection to deliver two 4 Gbps Fibre Channel ports over SFF interfaces. The dual Fibre Channel connections provided by the Advanced RTM can connect to qualified NEBS Level 3-certified storage arrays such as Oracle's StorageTek 2540 and 6140 arrays.

This Advanced RTM also provides six Gigabit Ethernet interfaces to balance network throughput with its Fibre Channel disk interfaces. Three interfaces are passed from the processor blade via one RJ-45 connection and two SFP interfaces. The remaining four lanes of the x8 PCIe interface drive a quad Gigabit Ethernet controller that provides four RJ-45 interfaces. The Advanced RTM includes a PCI x4 expansion port as well as a Gigabit Ethernet management port and a serial craft port.

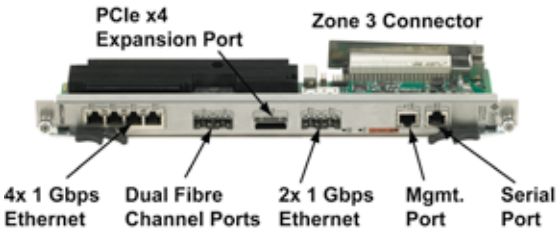

Figure 10. The Netra CP3200 ARTM-FC provides two 4 Gbps Fibre Channel ports to connect to NEBS-certified StorageTek arrays along with six Gigabit Ethernet ports.

Chapter 2

10 Gigabit Ethernet Throughput

10 Gigabit Ethernet networking technology in Oracle's Sun systems is specifically designed to accelerate multi-threaded application performance by optimizing I/O throughput in environments that utilize parallel threads. Providing a shared, multi-homed, 10 Gigabit Ethernet interface, silicon innovations in these platforms represent a revolutionary multi-threaded approach to connecting systems to the network.

10 Gigabit Ethernet is integrated throughout the Sun Netra CT900 ATCA Server via silicon innovations throughout the product line, including in the chassis midplane, dual modular switches, processor blades, and an Advanced RTM (see Table 2). The tight integration between Oracle Solaris, the Netra Data Plane Software Suite, Chip Multi-Threading (CMT) processors, and Oracle's Sun 10 Gigabit Ethernet silicon enables an integrated, optimized network environment with virtually unmatched throughput.

TABLE 2. SUN ATCA PLATFORMS INTEGRATE 10 GIGABIT ETHERNET THROUGHOUT THE PRODUCT LINE

INTEGRATION POINT	HOW SUN INTEGRATES 10 GBE ETHERNET
Sun Netra CT900 Server midplane	Uses PICMG 3.1 Option 9 extended
Extended Fabric	Switched on the midplane with the Sun Netra CP3240 10 GbE ATCA Switch Blade
Sun Netra CP3220 Opteron ATCA Blade Server	Uses standard Sun 10 Gigabit Ethernet ASIC
Sun Netra CP3250 Intel Xeon ATCA Blade Server	Uses standard Sun 10 Gigabit Ethernet ASIC
Sun Netra CP3270 Intel Xeon ATCA Blade Server	Uses Intel 10 Gigabit Ethernet ASIC
Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server	Integrated directly into the UltraSPARC T2 processor
Oracle Netra T3-1BA SPARC T3 ATCA Blade Server	Integrated directly into the SPARC T3 processor
Sun Netra CP3200 ARTM-10GAdvanced RTM	Uses standard Sun 10 Gigabit Ethernet ASIC

10 Gigabit Ethernet Networking Technology

Several Sun systems from Oracle feature 10 Gigabit Ethernet networking technology in the form of an ASIC or in similar logic that is embodied in the UltraSPARC T2 and SPARC T3 processor itself. In addition, an Intel 10 Gigabit Ethernet controller enables 10 Gigabit Ethernet networking in the Sun Netra CP3270 Intel Xeon Server Blade.

10 Gigabit Ethernet Sun ASIC

The first realization of 10 Gigabit Ethernet technology is Oracle's Sun ASIC (see Figure 11) that includes four Ethernet ports configured as two quad speed (10/100/1000/10000 Mbps) and two triple speed (10/100/1000 Mbps) Ethernet ports. Providing a shared, multi-homed, 10 Gigabit Ethernet interface, the ASIC includes multiple ports and advanced processing and virtualization features that help accelerate application performance by optimizing I/O throughput in environments that utilize parallel threads. The ASIC's ability to deliver high performance and enable I/O virtualization is made possible through an architecture that supports line rate packet classification and multiple DMA engines, and provides a wide variety of interrupt handling schemes and several other performance optimizations (see Figure 12).

Figure 11. The 10 GbE Ethernet Sun ASIC is used in several Sun Netra blade servers.

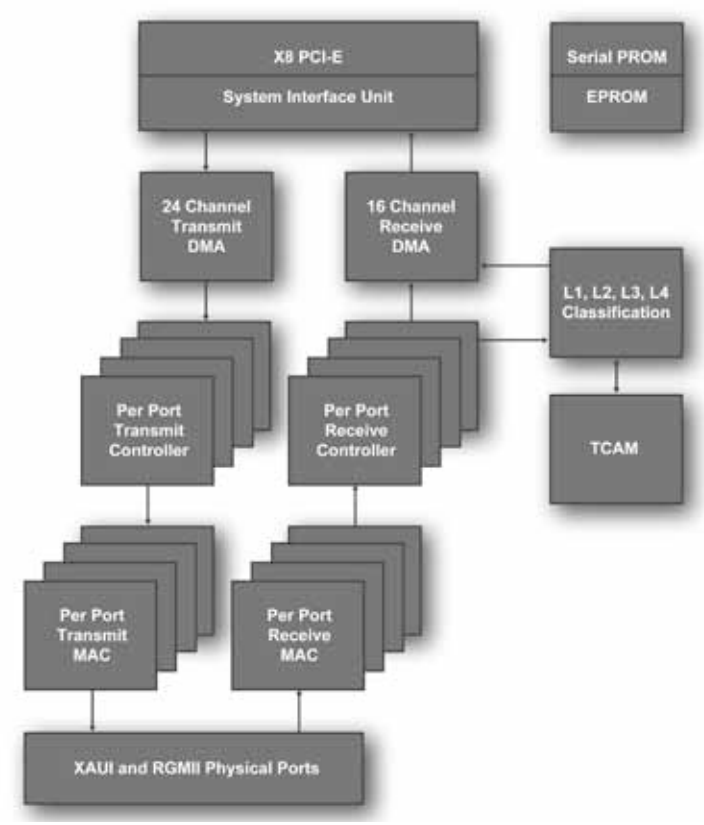

Figure 12. This functional block diagram shows logic used in the Sun 10 Gigabit Ethernet ASIC.

Integrated Into the UltraSPARC T2 and SPARC T3 Processors

The UltraSPARC T2 and SPARC T3 processors provide two 10 GbE Ethernet ports with integrated SERDES, providing line-rate packet classifications at up to 30 million packets/second (based on layers 1-4 of the protocol stack). Multiple direct memory access (DMA) engines match accesses to individual threads, providing binding flexibility between ports and threads. Virtualization support includes provisions for eight partitions, and interrupts may be bound to different hardware threads. These capabilities allow NEPs to create packet-processing applications that dedicate each UltraSPARC T2 processor thread to a specific function.

While the remainder of this chapter discusses the Sun 10 Gigabit Ethernet ASIC, its internal logic is very similar to the functions incorporated into the UltraSPARC T2 processor. The main difference is that the UltraSPARC T2 processor incorporates 16 transmit and receive channels, while the ASIC uses 24 transmit channels and 16 receive channels.

Line Rate Packet Classification

Until recently, the hardware-to-software interface was based on queuing packets into the system, and then distributing and processing those packets. To improve load balancing and performance, the ASIC's 10 Gigabit Ethernet logic transforms the interface by moving to a distribute-then-queue model. Capable of handling over 30 million packets per second, the ASIC statically separates receive traffic into multiple logical queues that can be worked on in parallel. Packets can be classified based on packet classes, ternary matches, or hash functions.

Packet classification extends parallelism to I/O by enabling packets to be separated so that packets that do not depend on one another can be worked on simultaneously. Aggregate network workload is separated into many smaller increments — up to 16-way using Sun technology — enabling multiple TCP streams to be allocated to different processor threads in parallel (see Figure 13). Such a strategy spreads the network load among processing units, and frees CPUs and ports for other tasks.

Figure 13. Packet classification in the ASIC's 10 Gigabit Ethernet Networking Technology enables network packets to be worked on in parallel.

Multiple DMA Channels and CPU and Thread Affinity

Oracle's Sun 10 Gigabit Ethernet networking technology employs multiple DMA channels — 24 transmit and 16 receive channels on the ASIC— to help speed throughput. In addition, the ASIC enables a one-to-one correlation of receive and transmit packets across the same TCP connection, and

binding flexibility between DMA channels and ports. These capabilities keep caches warm, and can help avoid cross-calls and context switching to deliver greater performance while reducing the need for CPU resources to support I/O processing.

Virtualization Support

Support for fine-grained virtualization enables I/O to be virtualized through interface sharing, partitioning, and network virtualization. Resources can be partitioned logically into as many as eight groups. Virtualization can be based on the ingress port, virtual LAN (VLAN), MAC address, or service address. Up to 4,096 VLANs can be attached to a given port, up to 16 MAC addresses can be assigned to a 10 Gigabit Ethernet port, and up to 256 IP addresses or TCP/UDP ports can be associated with a service address. With these capabilities, the Sun ASIC can consolidate network attachments by aggregating attachment points to save switch ports and server I/O slots. In addition, fine-grained virtualization helps limit the compute resources assigned to a given connection. Such a scheme proves invaluable in the event of accidental or malicious high volume traffic, such as a Denial of Service attack, by ensuring the offending actions affect only the resources assigned to the virtual port or address and not the entire system.

Multiple Speed Ethernet Ports

The ASIC includes up to four Ethernet ports configured as two quad speed (10/100/1000/10000 Mbps) and two triple speed (10/100/1000 Mbps) Ethernet ports. With four ports, the ASIC minimizes the number of hops through I/O bridges, thereby reducing the latency between system memory and the network. In addition, the ASIC provides two XAUI 10 GbE ports and four Reduced Gigabit Media Independent Interface (RGMII) 1000 Mbps ports, enabling the mixing and matching of ports and speeds based on system needs.

Interrupts

By handling various interrupts, the technology can accelerate performance for any platform and operating system capable of distributing its network load.

PCIe 1.1 Support

Designed for use in a wide variety of systems, Oracle's Sun 10 Gigabit Ethernet networking technology includes support for the PCIe 1.1 standard and is capable of operating at x1, x4, or x8-lane configurations. In the Sun Netra CP3250 and CP3220 Blade Servers and in the Netra CP3200 ARTM-10G, the PCIe designs use x8 lanes to optimize throughput. Up to 32 outstanding transactions can be handled at any given time. In addition, the ASIC supports relaxed memory access via on-chip transaction re-ordering to hide I/O and memory latencies, and improve throughput and efficiency.

Chapter 3

Oracle Netra SPARC T3-1BA ATCA Blade Server Architecture

The Netra SPARC T3-1BA ATCA Blade is based on the SPARC T3 processor with 12 Cores, 96 Threads running at 1.4GHz. The SPARC T3 is the latest iteration of the Chip Multi-Threading technology, pioneered in UltraSPARC T1 and the server on a chip design introduced in the UltraSPARC T2, with higher scaling than ever before in a single chip design. The SPARC T3 utilizes the Solaris 10 Operating Environment to extend Oracle's commitment to Solaris binary compatibility to the latest processor.

Key Features

The Oracle Netra T3-1BA SPARC T3 ATCA Blade Server is an ATCA node board based on the SPARC T3 processor. It is hot-swappable to an ATCA midplane and supports dual Gigabit Base interfaces and 10 Gigabit XAUI Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. The processor blade's key features include:

- NEBS Level 3-certified
- Compliant to PICMG 3.0 R2 and PICMG 3.1 Option 9
- Single processor socket support with twelve cores, eight threads per core, at 1.4GHz
- Eight DDR3 memory sockets for up to 32 GB of main memory
- Compatible with Sun Advanced RTM options for rear-accessible 10 GbE Ethernet, hard disk drives and interfaces, and Fibre Channel interfaces
- Eight Gigabit Ethernet channels:
- Two channels for Basic fabric (1 Gbps — PICMG 3.0)
- Two channels for Extended Fabric (10 GbE— PICMG 3.1, Option 9)
- Three channels for Zone 3 (1 Gbps— ARTM implementation specific)
- One channel for front panel (1G)
- One RS-32 serial craft port (front panel)
- Two USB ports (front panel)
- 16 or 32 GB Compact Flash socket
- Oracle Solaris 10

Architectural Overview

A more detailed architectural overview can be found in the Oracle Netra SPARC T3-1BA Board User's Guide.

Figure 14 illustrates the core components of the Oracle Netra SPARC T3-1BA SPARC T3 ATCA Blade Server and interfaces to the front panel, the midplane (through Zone 1 and Zone 2 connectors), and to Advanced RTMs (through the Zone 3 connector)

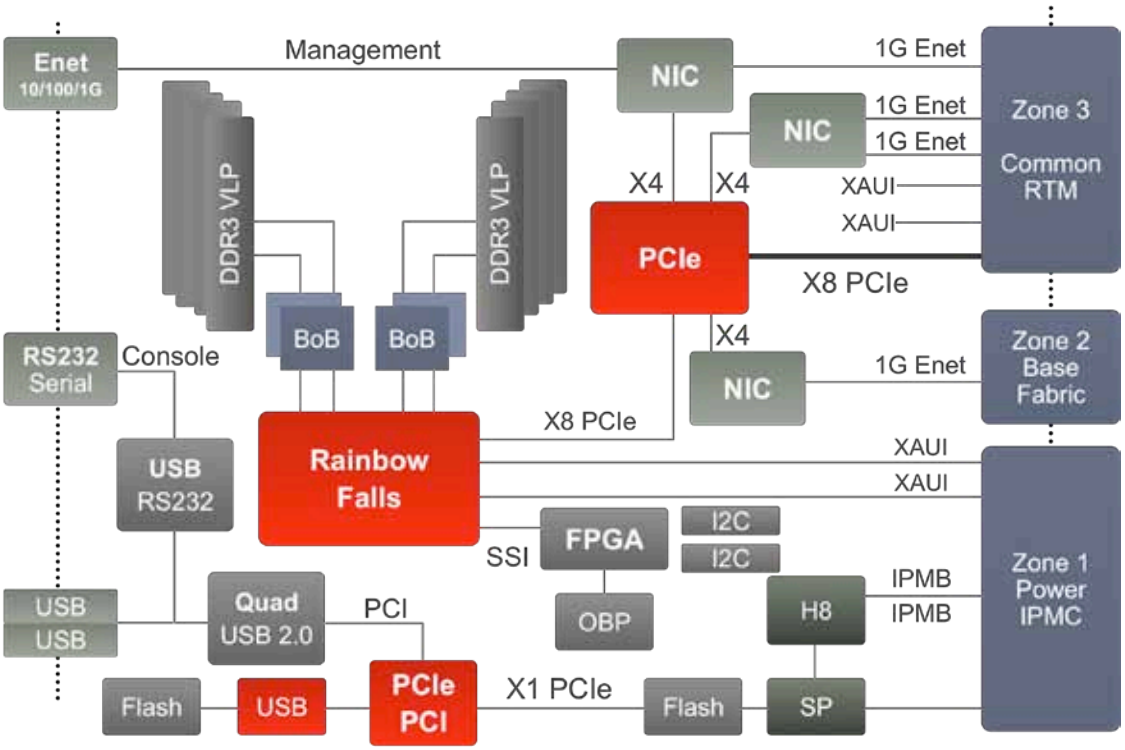

Figure 14. Netra SPARC T3-1BA Blade Server block diagram.

UltraSPARC T3 Processor and Memory

The SPARC T3 processor used in the Oracle Netra SPARC T3-1BA SPARC T3 ATCA Blade Server has twelve physical SPARC V9-compatible processor cores running at a clock rate of 1.4 GHz. Each core has full support for eight threads, and includes:

- Two integer execution pipelines
- One fully pipelined floating-point unit per core
- One stream processing unit for cryptographic functions per core

Each core is connected via crossbar switch to eight banks of 24-way set associative Level 2 cache totaling 6 MB. The eight banks interface to four on-chip memory controllers each of which connect to two DDR3 sockets on the processor blade for a total of eight memory sockets (see Figure 15).

In addition to the memory interface, the SPARC T3 processor incorporates a dual, on-chip 10 GbE Ethernet interface and an integral PCIe 2 interface. These two interfaces form the basis for the processor blade's I/O architecture.

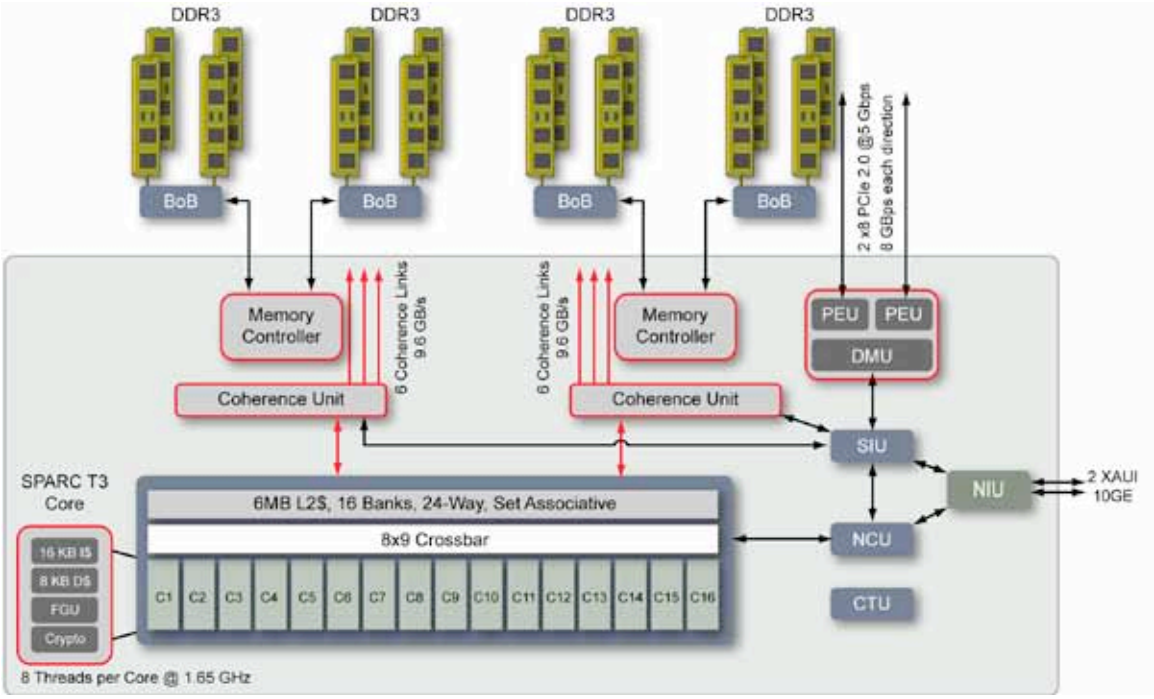

Figure 15. SPARC T3 Processor block diagram

Base and Extended Fabric Interfaces

The processor's x8 PCIe 2 interface is connected to a PCIe switch that, through an x4 connection, powers a dual Gigabit Ethernet controller (Intel 82571) that connects to the Base through a Broadcom 5397 switch. This provides dual Gigabit Ethernet connectivity to the Base interface on the Zone 2 connector.

The Extended Fabric interface is driven by the processor's dual XAUI interface and connects to the midplane to communicate with the Netra CP3240 10G switch.

Additional I/O

The on-board PCIe switch provides connectivity to a number of different I/O devices including the serial craft port, dual USB ports, and a Compact Flash card.

The PCIe to PCI bridge chip provides an on-board local PCI bus to a PCI-to-USB device that provides the USB ports. Two USB ports connect to the front panel and one USB port connects to a USB-to-IDE bridge. The USB-to-IDE bridge supplies the IDE interface for the Compact Flash card.

Advanced RTM Connectivity

Support for Netra CP3200 Advanced RTM modules is provided through Zone 3 connectors. Typical Advanced RTM support includes:

- One serial console port
- One Gigabit Ethernet management port.

To support I/O functionality of Advanced RTMs, the processor blade also provides the following interfaces:

- An x8 PCIe interface from the on-board PCIe switch
- Dual Gigabit SERDES Ethernet channels (from one 82571 dual Ethernet Controller)
- Dual XAUI connections for additional external 10 GbE Ethernet (as described in "Base and Extended Fabric Interfaces.")

Management Infrastructure

The service processor is used for domain configuration and hypervisor interaction, host system reset, and boot support. The service processor runs the BSC firmware, and it uses the XC3S1000 FPGA as a gateway to the SPARC T3 processor via its SSI/JTAG interface. This FPGA also provides support for the IPM controller.

The H8 chip provides the IPM Controller (IPMC) function on the processor blade. The IPMC provides PICMG 3.0 board management functionality, and it interfaces to the host CPU through a serial interface. The IPMC runs on standby power and handles power, reset, and environmental monitoring for the entire processor blade.

Chapter 4

Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server Architecture

Employing the UltraSPARC T2 processor — the industry's first massively threaded system on a chip — the Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server offers breakthrough performance and energy efficiency. This Chip Multi-Threading (CMT) technology processor supports up to 64 threads in a single eight-core processor — providing increased computational density while staying within highly constrained envelopes of power and cooling. Very high levels of integration help reduce latency, lower costs, and improve security and reliability. Balanced system design provides support for a wide range of application types, from communication carrier applications to packet-level processing in Data Plane functions.

Key Features

The Sun Netra CP3260 UltraSPARC T2 ATCA Server is an ATCA node board based on the UltraSPARC T2 processor. It is hot-swappable to an ATCA midplane and supports dual Gigabit Base interfaces and 10 Gigabit XAUI or 1 Gigabit SERDES Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. The processor blade's key features include:

- NEBS Level 3-certified
- Compliant to PICMG 3.0 R2 and PICMG 3.1 Option 9
- Single processor socket support with processor options for six or eight cores, eight threads per core
- Eight FB-DIMM memory sockets for up to 32 GB of main memory
- Compatible with Sun Advanced RTM options for rear-accessible 10 GbE Ethernet, hard disk drives and interfaces, and Fibre Channel interfaces
- Eight Gigabit Ethernet channels:
- Two channels for Basic fabric (1 Gbps — PICMG 3.0)
- Two channels for Extended Fabric (10 GbE— PICMG 3.1, Option 9)
- Three channels for Zone 3 (1 Gbps— ARTM implementation specific)
- One channel for front panel (1G)
- One RS-32 serial craft port (front panel)
- Two USB ports (front panel)
- 8-16 GB Compact Flash socket
- Oracle Solaris 10 and Sun Netra Data Plane Software Suite support

Architectural Overview

A more detailed architectural overview can be found in the Sun Netra CP3260 Board User's Guide.

Figure 16 illustrates the core components of the Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server and interfaces to the front panel, the midplane (through Zone 1 and Zone 2 connectors), and to Advanced RTMs (through the Zone 3 connector).

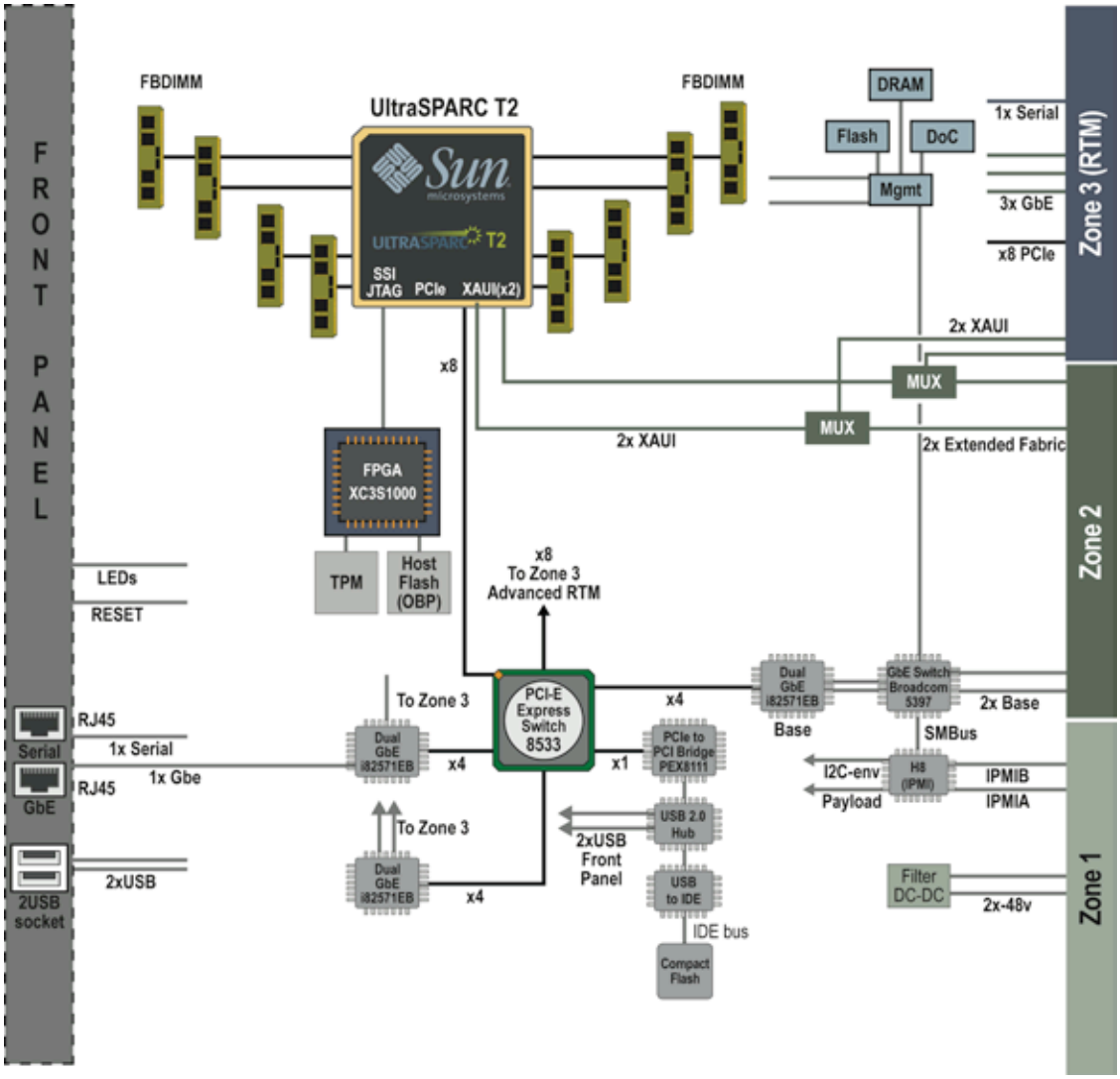

Figure 16. The block diagram for the Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server shows major board components.

UltraSPARC T2 Processor and Memory

The UltraSPARC T2 processor used in the Sun Netra CP3260 UltraSPARC T2 ATCA Blade Server has eight physical SPARC V9-compatible processor cores running at a clock rate of 1.167 GHz. Each core has full support for eight threads, and includes:

- Two integer execution pipelines
- One fully pipelined floating-point unit per core
- One stream processing unit for cryptographic functions per core

Each core is connected via crossbar switch to eight banks of 16-way set associative Level 2 cache totaling 4 MB. The eight banks interface to four on-chip memory controllers each of which connect to two FB-DIMM sockets on the processor blade for a total of eight memory sockets (see Figure 17).

In addition to the memory interface, the UltraSPARC T2 processor incorporates a dual, on-chip 10 GbE Ethernet interface and an integral PCIe interface. These two interfaces form the basis for the processor blade's I/O architecture.

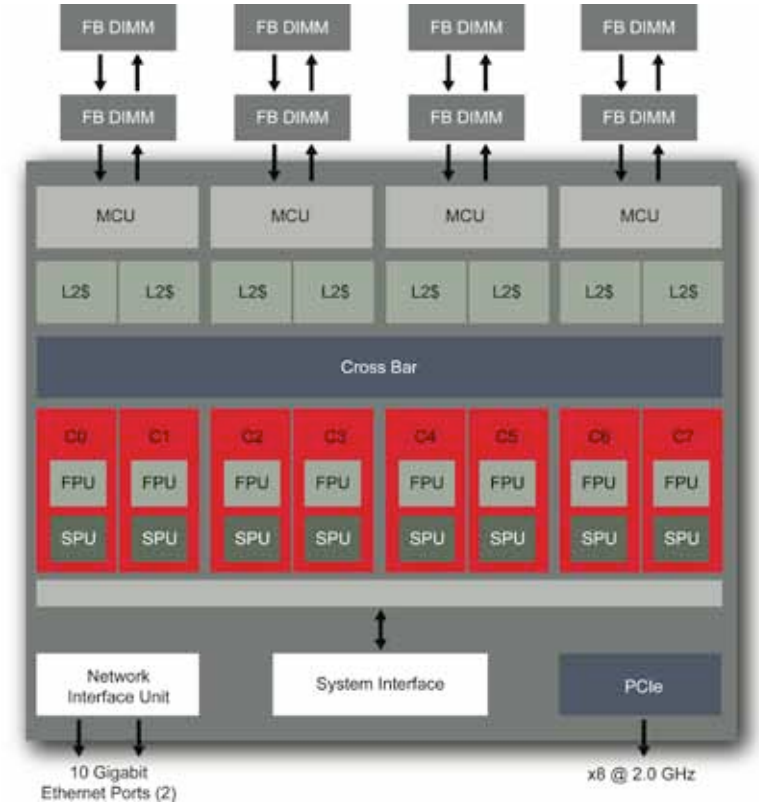

Figure 17. The UltraSPARC T2 processor features 8 cores, each core with two integer pipelines and separate FPU/SPUs.

Base and Extended Fabric Interfaces

The processor's x8 PCIe interface is connected to a PCIe switch that, through an x4 connection, powers a dual Gigabit Ethernet controller (Intel 82571) that connects to the Base through a Broadcom 5397 switch. This provides dual Gigabit Ethernet connectivity to the Base interface on the Zone 2 connector.

The Extended Fabric interface is driven by the processor's dual XAUI interface and connects to the midplane through multiplexors. These multiplexors can be programmed to connect one or both of the

interfaces to the Extended Fabric interfaces, or they can route one or both of the interfaces to the Zone 3 connector.

- When routed to the Zone 3 connector, the Netra CP3200 ARTM-10G Advanced RTM can pass through the XAUI interfaces to a pair of SFP+ connectors, bringing a total of four 10 Gigabit Ethernet interfaces directly to the rear of the system.
- When routed to the Zone 2 connector, the XAUI connections bring dual 10 GbE Ethernet to the midplane, where it can be switched by the Netra CP3240 10 GbE Switch Blade.

Additional I/O

The on-board PCIe switch provides connectivity to a number of different I/O devices including the serial craft port, dual USB ports, and a Compact Flash card.

The PCIe to PCI bridge chip provides an on-board local PCI bus to a PCI-to-USB device that provides the USB ports. Two USB ports connect to the front panel and one USB port connects to a USB-to-IDE bridge. The USB-to-IDE bridge supplies the IDE interface for the Compact Flash card.

Advanced RTM Connectivity

Support for Netra CP3200 Advanced RTM modules is provided through Zone 3 connectors. Typical Advanced RTM support includes:

- One serial console port
- One Gigabit Ethernet management port.

To support I/O functionality of Advanced RTMs, the processor blade also provides the following interfaces:

- An x8 PCIe interface from the on-board PCIe switch
- Dual Gigabit SERDES Ethernet channels (from one 82571 dual Ethernet Controller)
- Dual XAUI connections for additional external 10 GbE Ethernet (as described in "Base and Extended Fabric Interfaces.")

Management Infrastructure

The service processor is used for domain configuration and hypervisor interaction, host system reset, and boot support. The service processor runs the BSC firmware, and it uses the XC3S1000 FPGA as a gateway to the UltraSPARC T2 processor via its SSI/JTAG interface. This FPGA also provides support for the IPM controller.

The H8 chip provides the IPM Controller (IPMC) function on the processor blade. The IPMC provides PICMG 3.0 board management functionality, and it interfaces to the host CPU through a serial interface. The IPMC runs on standby power and handles power, reset, and environmental monitoring for the entire processor blade.

Chapter 5

Sun Netra CP3270 Intel Xeon ATCA Blade Server Architecture

For performance-intensive x86 applications with low power requirements, Network Equipment Providers and Telecom Equipment Manufacturers can deploy Oracle's Sun Netra CP3270 Intel Xeon ATCA Blade Server which doubles the performance while maintaining the same footprint. The Sun Netra CP3270 ATCA Blade Server features dual sockets for embedded-class Intel Xeon LC5518 processors, which follow the next-generation Intel Core Microarchitecture and use a 45nm manufacturing process. In addition, these processors integrate the memory controller and PCIe functionality of a Northbridge Memory Controller Hub (MCH) directly into the processor die, reducing latency and lowering power consumption. As a result, this blade server can support a full-service configuration of 32GB and 16 threads even within a low, 200-Watt power envelope.

Key Features

The Sun Netra CP3270 Server is an ATCA node board with dual processor sockets. Each Intel Xeon LC5518 1.73 GHz processor features four execution cores, two threads per core, and a shared 8MB Level 3 cache to enable fast performance. The board is hot-swappable to an ATCA midplane and supports dual Gigabit Base interfaces and dual 10 Gigabit XAUI Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. Key features include:

- NEBS Level 3-certified
- Compliant to PICMG 3.0 R2 and PICMG 3.1 Option 9
- Dual-socketed, quad-core Intel Xeon LC5518 processors at 48 W, with integrated memory controller support and PCIe 2.0 functionality
- Eight DDR3 memory slots for up to a maximum main memory of 32 GB
- One AMC I/O slot (x8 PCIe) that also supports SAS signaling between the processor blade and a Netra CP3200 ARTM-HDD Advanced RTM
- Compatible with Sun Advanced RTM options for rear-accessible 10 GbE Ethernet, hard disk drives and interfaces, and Fibre Channel interfaces
- Two Gigabit Ethernet channels connected to the Base fabric
- Two 10 Gigabit Ethernet channels for the Extended Fabric
- One front panel-based serial craft port
- One front panel-based Gigabit Ethernet management port
- Two front panel-based USB ports
- One Gigabit Ethernet management port routed to the optional Advanced RTM through the Zone 3 connector

- Redundant 8 MB BIOS
- Compact Flash Type II
- Support for Oracle Solaris, carrier-grade Linux, and Microsoft Windows 2008 Advanced Server

Architectural Overview

A more detailed architectural overview can be found in the Sun Netra CP3270 Board User's Guide.

Figure 18 illustrates the core components of the Sun Netra CP3270 Intel Xeon ATCA Blade Server, including interfaces to the front panel, the midplane (through Zone 1 and Zone 2 connectors), and to Advanced RTMs (through the Zone 3 connector).

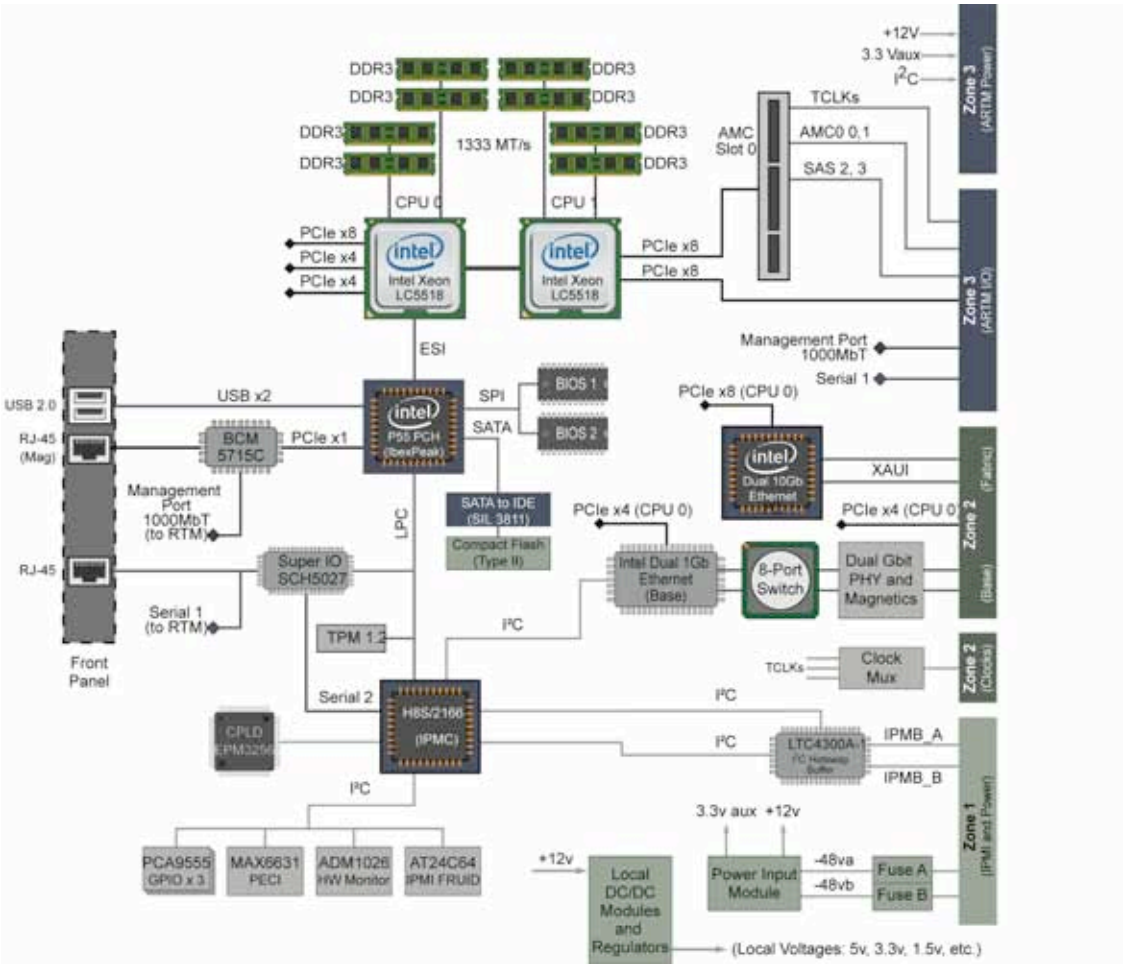

Figure 18. The block diagram for the Sun Netra CP3270 Intel Xeon ATCA Blade Server depicts major components.

Intel Xeon Processors

The Intel Xeon LC5518 processor is an embedded class processor design that uses the next-generation Intel Core Microarchitecture (codenamed "Nehalem"). This microarchitecture features enhancements over previous generation Intel Xeon microarchitectures, including:

- Multiple processor cores. The microarchitecture design scales from 2 to 8 cores per die, with four cores implemented in the Intel Xeon LC5518 processor.
- Intel HyperThreading (HT) technology. This technology provides multiple virtual threads per core, increasing performance for highly threaded applications. On the Intel Xeon LC5518 processor, there are two threads implemented per core, for a total of 16 threads per blade.
- Intel QuickPath technology. This new technology provides a high-speed, point-to-point interconnect between processors, memory, and I/O. The Intel QuickPath interconnect (QPI) links processors in a distributed shared memory design that enables high bandwidth and low latency memory access. Because it is a point-to-point interconnect, processors do not contend for a single bus when accessing memory and I/O, and do not compete for bus bandwidth, which enhances scalability.
- Integrated DDR3 memory controller. Designed as a separate component in earlier architectures, the memory controller is integrated on the processor die. The processor design creates a Non-Uniform Memory Access (NUMA)-style memory architecture since each processor in multi-socketed systems can access local memory connected to the local processor's memory controller as well as remote memory connected to the second processor.
- Advanced cache model. Each core has an associated Level-1 (L1) instruction/data cache (64KB per core) and a large integrated Level-2 (L2) cache (256KB per core). Also, all cores on a die share access to an inclusive Level-3 (L3) cache. On the Intel Xeon LC5518 processor, the four cores on each die share access to an 8MB cache (see Figure 19).
- Extended SSE4 (Streaming SIMD Extensions). These processor extensions improve performance for XML, string, and text processing.
- Virtualization enhancements. Embedded virtualization technologies enable hardware-based assistance for I/O device virtualization, improved virtualization efficiency, and enhanced connectivity within a virtualized server.
- Intel Turbo Boost Technology. For both multi-threaded and single-threaded workloads, this technology increases performance by taking advantage of processor and system power along with thermal headroom. The Turbo Boost feature can increase performance up to two or three speed bins (266 or 400 MHz) above typical performance levels. Turbo Boost and HyperThreading capabilities vary according to specific processor models.
- Intel Intelligent Power Technology. When a processor workload decreases, unneeded components — cores, cache and memory — are put into sleep mode to reduce power consumption.

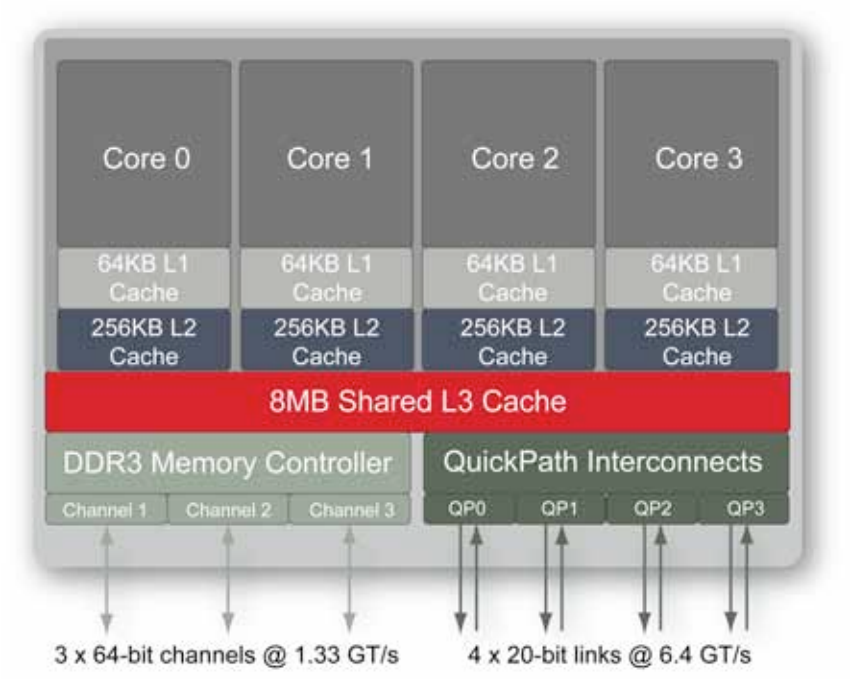

Figure 19. Four cores on the Intel Xeon Processor 5500 Series share an 8MB L3 cache.

Integrated I/O Functionality

In addition to an integrated memory controller in this microarchitecture, the embedded-class Intel Xeon LC5518 processor integrates PCIe 2.0 functionality typically found in a Northbridge chip directly into the processor. The addition of this I/O capability on-die provides significant benefits — first, it reduces I/O latency, which helps to improve performance, and secondly, it eliminates approximately 20 to 25 watts of power that is generally consumed by a separate Northbridge MCH.

Memory

The integrated memory controller supports multiple DDR3 memory channels per processor, providing high memory bandwidth for applications. DDR3 memory components also offer greater density and run at higher speeds (but at significantly lower voltages) than previous generation DDR2 memories.

Each Intel Xeon LC5518 processor supports two memory channels that can be populated with two DDR3 Registered ECC DIMM modules per channel, for a total of eight DIMMs per blade server. Clocked at 1333MHz, 4 GB DDR3 DIMMs are available at initial product release, enabling a maximum memory capacity of 32 GB. Although each Intel Xeon LC5518 processor supports three memory channels natively, only two are used that support two DIMMs per channel (due to space and power constraints) on the Sun Netra CP3270 Intel Xeon ATCA Blade Server.

Base and Extended Fabric Interfaces

As shown in the block diagram in Figure 18, I/O and networking capabilities are managed in conjunction with the Intel BD3420 PCH Platform Controller Hub (PCH, formerly Ibex Peak), which connects to CPU 0 using an ESI link.

The Base fabric is driven by an Intel 82576EB Gigabit Ethernet Controller that integrates with the system through an x4 PCIe interface from CPU 0. This dual controller enables two Gigabit Ethernet connections via the Zone 2 connector.

The Extended Fabric is driven by an Intel 82599EB 10 Gigabit Ethernet Controller that uses an x8 PCIe connection to CPU 0. The ASIC provides two 10 GbE interfaces through XAUI connections to the Zone 2 connector.

Additional I/O

Additional I/O interfaces are provided through integrated I/O functionality in the processors:

- 8-lane PCIe connectivity is provided to the Zone 3 connector from CPU 1
- 8-lane PCIe connectivity is provided to the single AMC slot from CPU 1

Remaining I/O interfaces are supported through the Intel P55 PCH:

- A SATA interface connects to a Compact Flash memory slot, which supports 16 GB or 32 GB flash modules
- Two USB ports are provided on the front panel
- A dual Gigabit Ethernet interface provides management port connections to the front panel and through an optional Advanced RTM to the back panel
- A Low Pin Count (LPC) bus connects to a serial craft interface and to the management subsystem
- An SPI bus connects to the dual redundant 8 MB BIOS chips

Advanced RTM Connectivity

Support for Sun Netra CP3200 Advanced RTM modules is provided through the Zone 3 connectors. Typical Advanced RTM support includes:

- One serial console port
- One Gigabit Ethernet management port

To support the I/O functionality of the Advanced RTMs, the processor blade also provides the following interfaces:

- An x8 PCIe interface provided from CPU 1
- AMC I/O connections

- One pair of SAS connections that allow the Netra CP3200 ARTM-HDD Advanced RTM to control disk drives installed in the AMC slot

Management

An H8 chip provides the IPM Controller (IPMC) function on the processor blade. The IPMC provides PICMG 3.0 board management functionality, and interfaces to most of the monitored components through multiple I^2C buses. The IPMC handles power, reset, and environmental monitoring for the entire blade server.

Chapter 6

Sun Netra CP3250 Intel Xeon ATCA Blade Server Architecture

The Sun Netra CP3250 Intel Xeon ATCA Blade Server brings high performance, availability, and innovation to the central office and telecommunications datacenter. It provides applications with solid performance and incorporates Oracle's Sun ASIC and its advanced 10 Gigabit Ethernet networking technology. This blade uses Intel embedded quad-core processors that employ an advanced 45nm manufacturing process to achieve high transistor density and to operate within a low power envelope.

Key Features

The Sun Netra CP3250 Server is an ATCA node board with dual sockets for quad-core Intel Xeon L5408-LV 2.13 GHz ATCA processors. Each Intel Xeon L5408-LV processor features four execution cores and a 12MB Level 2 cache to enable fast performance. The board is hot swappable to an ATCA midplane and supports dual Gigabit Base interfaces and dual 10 Gigabit XAUI Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. Key features include:

- NEBS Level 3-certified
- Compliant to PICMG 3.0 R2 and PICMG 3.1 Option 9
- Dual-socketed, quad-core Intel Xeon L5408-LV processors at up to 40 W
- Six DDR2 memory slots for up to 24 GB of main memory
- One AMC I/O slot (x8 PCIe) that also support SAS signaling between the processor blade and a Netra CP3200 ARTM-HDD Advanced RTM
- Compatible with Sun Advanced RTM options for rear-accessible 10 GbE Ethernet, hard disk drives and interfaces, and Fibre Channel interfaces
- Two Gigabit Ethernet channels connected to the Base fabric
- Two 10 Gigabit Ethernet channels for the Extended Fabric
- One front-panel-based serial craft port
- One front-panel based Gigabit Ethernet management port
- Two front-panel-based USB ports
- One Gigabit Ethernet management port routed to the optional Advanced RTM through the Zone 3 connector
- Redundant 8 MB BIOS
- Compact Flash Type II
- Support for Oracle Solaris, carrier-grade Linux, and Microsoft Windows 2008 Advanced Server

Architectural Overview

A more detailed architectural overview can be found in the Sun Netra CP3250 Board User's Guide.

Figure 20 illustrates the core components of the Sun Netra CP3250 Intel Xeon ATCA Blade Server, including interfaces to the front panel, the midplane (through Zone 1 and Zone 2 connectors), and to Advanced RTMs (through the Zone 3 connector).

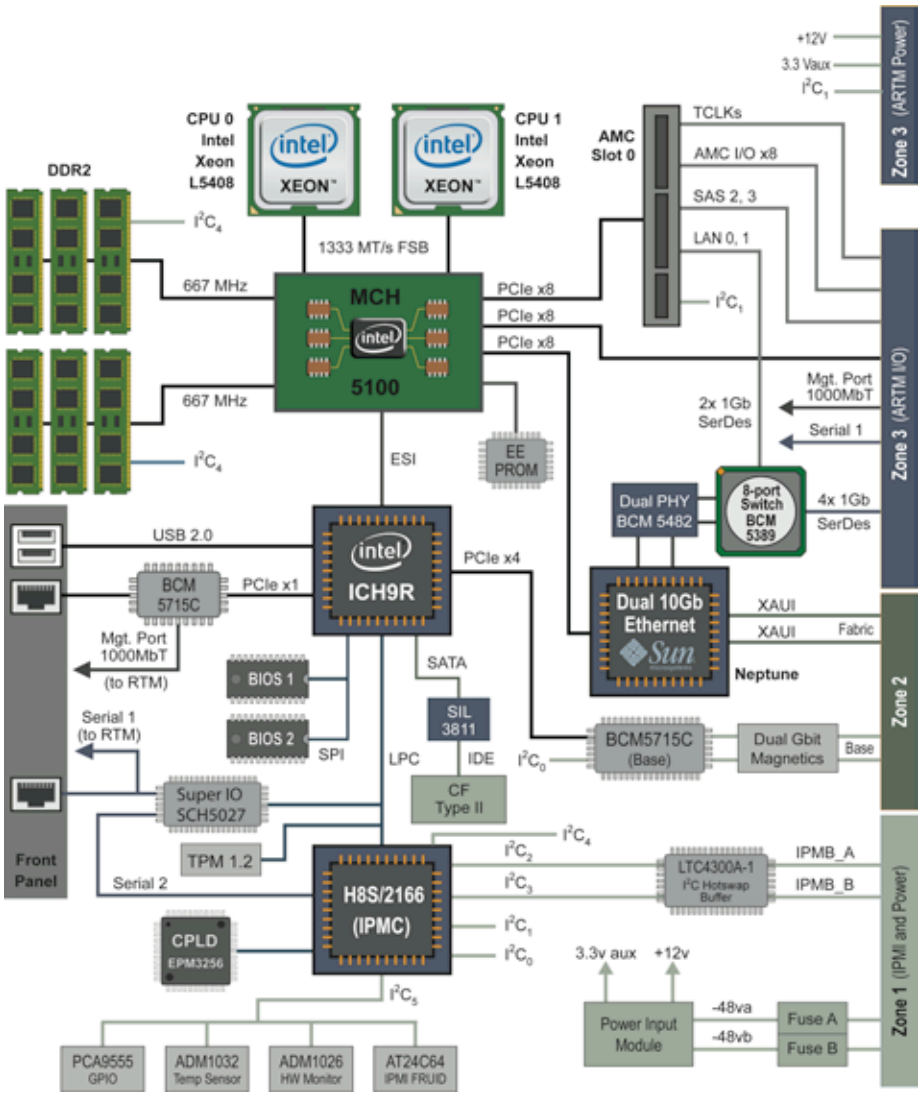

Figure 20. The block diagram for the Sun Netra CP3250 Intel Xeon ATCA Blade Server shows major components.

Intel Xeon Processors and Memory

The Intel Xeon L5408-LV processor is an embedded class processor design that uses a Multi-Chip Package (MCP) to deliver low-power quad-core configurations. The packaging approach increases die yields and lowers manufacturing costs, helping Intel and Sun to deliver high performance at low price

points. Based on the Intel Core Microarchitecture, the Intel Xeon L5408-LV processor features microarchitecture enhancements (such as large Level 2 caches) over previous generation Intel Xeon Processor 5000 Series processors. An advanced 45nm manufacturing process increases transistor density while maintaining a substantially reduced power envelope. These 40-watt, quad-core processors maximize performance-per-watt, providing increased density for central office and telecommunications datacenters.

Figure 21 is a logical diagram for the Intel Xeon L5408 processor that depicts its four execution cores. On each processor, a pair of cores shares a 6MB Level 2 cache (for a total of 12MB per processor). With two sockets in the Sun Netra CP3250 ATCA Blade Server, the design enables eight cores per blade.

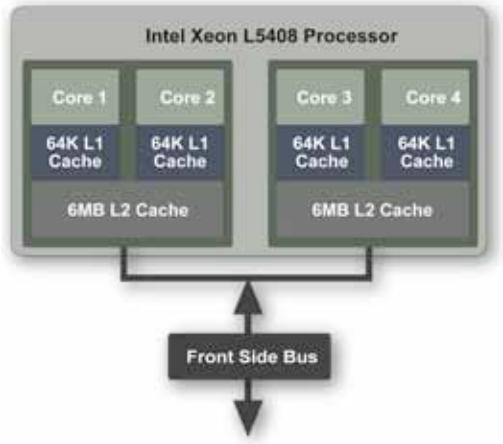

Figure 21. Each Intel Series L5408-LV processor uses two dies and features four cores.

The Intel Core Microarchitecture includes several features designed specifically to enhance performance while maintaining energy efficiency:

- Out-of-order execution conserves processor-to-memory bandwidth, improving memory access and increasing overall processing performance
- Speculative prefetches move data to the L2 cache before an L1 cache request occurs
- Large L2 caches increase system memory efficiency and enhance performance
- Power management capabilities modulate power delivery to execution cores, limiting power to unused components

As shown in Figure 20, dual independent Front Side Buses (FSBs) act as a system interconnect between each processor socket and the Intel 5100 Memory Controller Hub (MCH). For more information on the Intel Xeon Processor 5000 Sequence chipset, see www.intel.com/products/processor/xeon5000/.

Multiple DDR2 memory channels help to support large memory densities and high bandwidth. Each memory channel supports up to three DDR2 DIMMs, enabling up to 6 DDR2 DIMMs per system for a maximum capacity of 24GB using 4GB modules. The memory channels are organized into two

branches. The same 667MHz, PC2-5300 DDR2 DIMM modules — in either 1GB, 2GB, or 4GB capacities — are supported in both server models. Memory modules feature Error Checking and Correcting (ECC) with Chipkill technology for high reliability.

Base and Extended Fabric Interfaces

The blade's I/O and networking infrastructure is built around the Northbridge Intel 5100 Memory Controller Hub (MCH) and the Southbridge I/O Handler (IOH). The IOH (the Intel ICH9R) is interconnected to the Northbridge MCH using an ESI link.

The Base fabric is driven by a Broadcom 5715 dual Gigabit Ethernet controller that integrates with the system through an x4 PCIe interface on the IOH chip. The two Gigabit Ethernet connections are delivered to the Zone 2 connector.

The Extended Fabric is driven by the Sun 10 Gigabit Ethernet ASIC, which uses an x8 PCIe connection to the MCH. The two 10 GbE interfaces are provided through XAUI connections to the Zone 2 connector.

Additional I/O

Remaining I/O interfaces are also supported through the Northbridge Intel 5100 MCH or Southbridge IOH:

- 8-lane PCIe connectivity is provided to the Zone 3 connector
- 8-lane PCIe connectivity is provided to the single AMC slot
- A SATA to IDE interface connects to the Compact Flash memory slot
- Two USB ports are provided on the front panel
- A dual Gigabit Ethernet interface provides management port connections to the front panel and through an optional Advanced RTM to the back panel
- A Low Pin Count (LPC) bus connects to a serial craft interface and to the management subsystem
- An SPI bus connects to the dual redundant 8 MB BIOS chips
- The Sun 10 Gigabit Ethernet ASIC's two 1 Gbps interfaces are switched for LAN connectivity to the AMC slot and to the Zone 3 connector for Advanced RTM access

Advanced RTM Connectivity

Support for Sun Netra CP3200 Advanced RTM modules is provided through the Zone 3 connectors. Typical Advanced RTM support includes:

- One serial console port
- One Gigabit Ethernet management port.

To support the I/O functionality of the Advanced RTMs, the processor blade also provides the following interfaces:

- An x8 PCIe interface provided from the MCH
- Dual Gigabit SERDES Ethernet channels (from the 8-port Ethernet switch)
- AMC I/O connections
- One pair of SAS connections that allow the Netra CP3200 ARTM-HDD Advanced RTM to control disk drives installed in the AMC slot

Management

The H8 chip provides the IPM Controller (IPMC) function on the processor blade. The IPMC provides PICMG 3.0 board management functionality, and it interfaces to most of the monitored components through multiple I²C buses. The IPMC handles power, reset, and environmental monitoring for the entire processor blade.

Chapter 7

Sun Netra CP3220 AMD Opteron ATCA Blade Architecture

The Sun Netra CP3220 AMD Opteron ATCA Blade is built around second-generation, dual and quad-core AMD Opteron processors. This single-socket processor blade architecture is designed for balanced computing, with I/O capacity to match the extreme power and bandwidth of the AMD Opteron processor and its HyperTransport mechanism. The AMD Opteron processor has a very high-speed, low-latency CPU-to-memory interconnect that supports moving and processing data in communication carrier environments (e.g., in call setup, HSS/HLR, and other control layer applications). Support for two AMCs, Advanced RTMs, Base, and Extended Fabric at 10 GbE makes it an excellent match for applications that modulate T1, OC3, OC12, and other carrier protocols to IP.

The processor blade incorporates the Sun 10 Gigabit Ethernet ASIC, which increases the AMD Opteron processor's headroom by offloading functions including traffic classification. This helps increase I/O density per ATCA processor slot, allowing communication carriers and NEPs to get the more value out of each ATCA system.

Key Features

The Sun Netra CP3220 AMD Opteron ATCA Blade is an ATCA node board based on a single AMD Opteron Rev. F processor. Available with a dual- or quad-core AMD Opteron processor, the blade is hot-swappable to an ATCA midplane and supports dual Gigabit Base interfaces and 10 Gigabit XAUI or 1 Gigabit SERDES Ethernet interfaces to the Extended Fabric to support a redundant dual star network topology. The processor blade's key features include:

- NEBS Level 3-certified
- Compliant to PICMG 3.0 R2 and PICMG 3.1 Option 1
- Single socket supports dual and quad-core AMD Opteron HE processors at up to 68 W
- Eight DDR2 memory slots for up to 32 GB of main memory
- Two AMC I/O slots (x4 PCIe) that also support SAS signaling between the processor blade and a Netra CP3200 ARTM-HDD Advanced RTM
- Compatible with Sun Advanced RTM options for rear-accessible 10 GbE Ethernet, hard disk drives and interfaces, and Fibre Channel interfaces.
- Two Gigabit Ethernet channels connected to the Base fabric
- Two 10 Gigabit Ethernet channels for the Extended Fabric
- One front-panel-based serial port
- One front-panel based Gigabit Ethernet management port
- One Gigabit Ethernet management port routed to the optional Advanced RTM through the Zone 3 connector

- Compact Flash II socket for 8-16 GB of Flash memory
- Support for Oracle Solaris and carrier-grade Linux

Architectural Overview

A more detailed architectural overview can be found in the Netra CP3220 Board User's Guide.

Figure 22 illustrates the core components of the Sun Netra CP3220 Opteron ATCA Blade and the interfaces they provide to the front panel, the midplane (through Zone 1 and Zone 2 connectors), and to Advanced RTMs (through the Zone 3 connector).

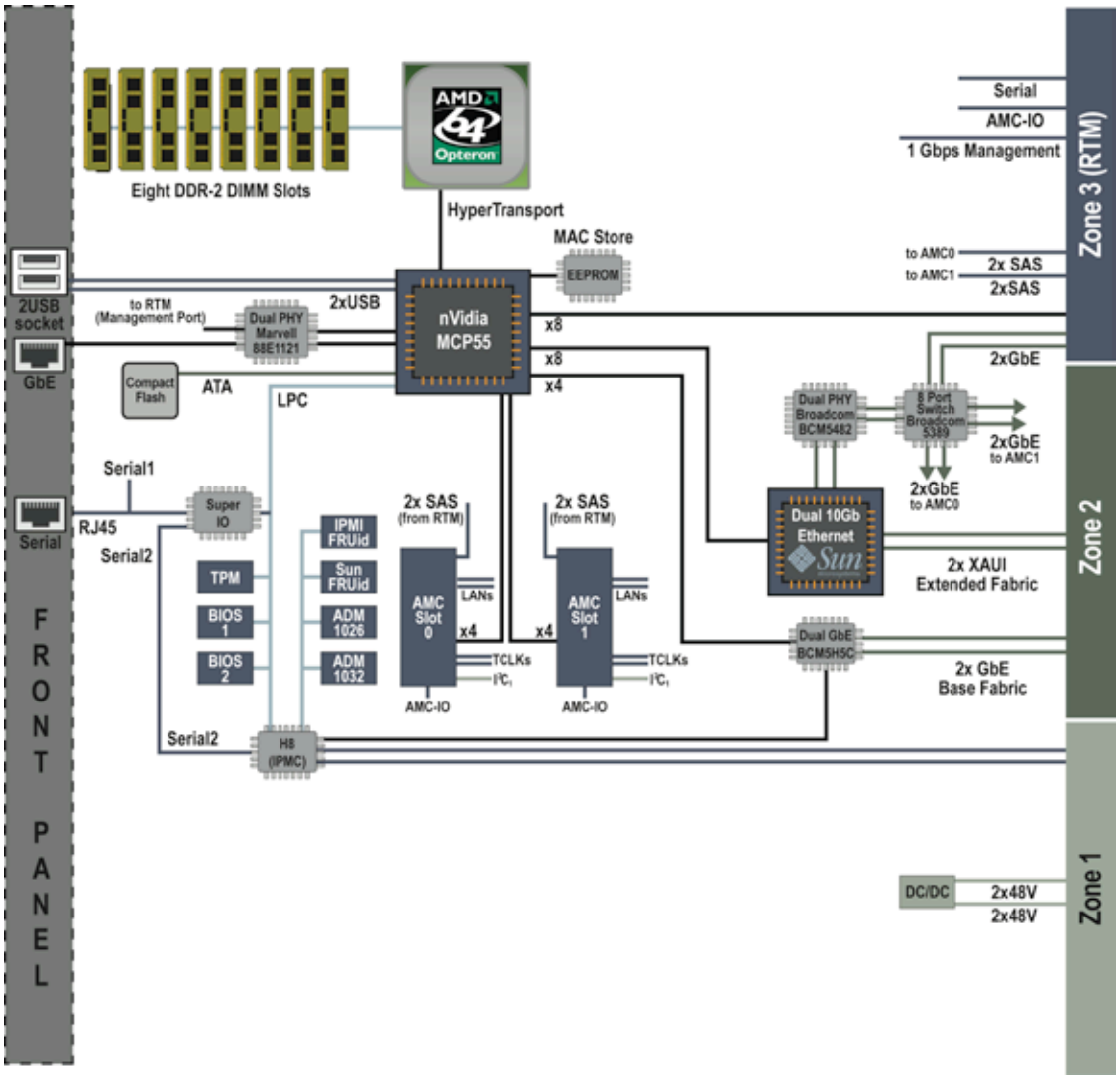

Figure 22. The block diagram for the Sun Netra CP3220 Opteron ATCA Blade shows major components.

The AMD Opteron Processor and Memory

The AMD Opteron Rev F processor with Direct Connect Architecture includes several new features including seamless quad-core upgradeability and DDR2 memory. AMD Opteron processors with DDR2 memory feature a common core architecture that is consistent across dual and quad-core implementations. Both processors utilize the same 1207-pin F socket.

Figure 23 provides a high-level processor block diagram. Features of the second-generation 2.2GHz AMD Opteron processor include:

- x86 Architecture (64-bit extensions) with AMD Direct Connect Architecture using HyperTransport technology supports simultaneous 32-bit and 64-bit computing
- Three HyperTransport links support up to 24 GBps peak bandwidth per processor
- Integrated 128-bit wide DDR memory controller reduces latency and increases performance
- 1 MB L2 cache per core
- Low power usage and cooling requirements — AMD Opteron Rev F processors at 68W offer efficient heat dissipation without reducing performance
- AMD PowerNow! technology can deliver performance on demand while helping to minimize power consumption
- DDR2-based systems can upgrade to quad-core AMD Opteron processors when they are available, within existing power and thermal envelopes
- NEBS-certified

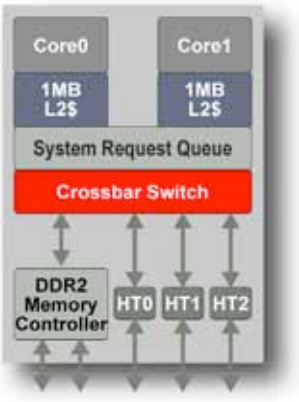

Figure 23. The Rev. F AMD Opteron processor features dual cores and HyperTransport technology.

On the Sun Netra CP3220 Opteron ATCA Blade, the processor's memory controller supports up to 32 GB of main memory using eight 4 GB DDR-2 DIMMS. The controller runs in 144-bit ECC mode at 667 MHz and achieves a memory bandwidth of up to 5.3 GBps.

Base and Extended Fabric Interfaces

The processor blade's I/O and networking infrastructure is built around the nVidia MCP55 media and communication processor that uses a high-performance, scalable architecture with a low-power design that helps to reduce overall power consumption.

The Base fabric is driven by a Broadcom 5715 dual Gigabit Ethernet controller that integrates with the system through an x4 PCIe interface on the MCP55 chip. The two Gigabit Ethernet connections are delivered to the Zone 2 connector.

The Extended Fabric is driven by Oracle's Sun 10 Gigabit Ethernet ASIC, which uses an x8 PCIe connection to the MCP55 chip. The two 10 GbE interfaces are provided through XAUI connections to the Zone 2 connector.

Additional I/O

The remaining I/O interfaces are also supported through the nVidia MCP55 media and communication processor:

- 8-lane PCIe connectivity is provided to the Zone 3 connector
- 4-lane PCIe connectivity is provided to each of the two AMC slots
- An ATA interface connects to the Compact Flash memory slot
- 2 USB ports are provided on the front panel
- A dual Gigabit Ethernet interface provides management port connections to the front panel and through an optional Advanced RTM to the back panel
- A Low Pin Count (LPC) bus connects to a serial craft port interface on the front panel, and to the dual 1 MB BIOS chips and management subsystems
- The Sun 10 Gigabit Ethernet ASIC's two 1 Gbps interfaces are switched for LAN connectivity to each AMC slot and to the Zone 3 connector for Advanced RTM access

Two AMC Slots

The processor blade contains two front-panel accessible AMC slots for I/O expansion. Both slots are single-width, mid-height slots. Each slot has one hot-plug connector for an AMC card that conforms to the PICMG AMC.0 specification, which also dictates support for AMC.1 Type 4E2 cards.

Each AMC slot is provided four lanes of PCIe connectivity, dual Gigabit Ethernet, and dual SAS signaling to an optional Advanced RTM through the Zone 3 connector. This configuration provides a high degree of I/O flexibility:

- Standard AMC cards can be installed with front-panel accessibility.
- When the optional Sun Netra CP3200 ARTM-HDD Advanced RTM is installed, the Advanced RTM's SAS disk controller can drive two hard disk drives on the processor blade using Port 2 or 3

SAS connections. This is in addition to the Advanced RTM's two internal drives and two external SAS interfaces.

- There is no native support for SAS drives on the processor board itself.

Advanced RTM Connectivity

Support for Sun Netra CP3200 Advanced RTM modules is provided through the Zone 3 connectors. Typical Advanced RTM support includes:

- One serial console port
- One Gigabit Ethernet management port.

To support the I/O functionality of the three Advanced RTMs, the processor blade also provides the following interfaces:

- An x8 PCIe interface provided from the MCP55 chip
- Dual Gigabit SERDES Ethernet channels (from the 8-port Ethernet switch)
- AMC I/O connections
- Two pairs of SAS connections that allow the Sun Netra CP3200 ARTM-HDD Advanced RTM to control disk drives installed in the AMC slots

Management

The H8 chip provides the IPM Controller (IPMC) function on the processor blade. The IPMC provides PICMG 3.0 board management functionality, and it interfaces to most of the monitored components through multiple I^2C buses. The IPMC handles power, reset, and environmental monitoring for the entire processor blade.

Chapter 8

Sun Netra CP3240 ATCA 10 GbE ATCA Switch Blade

The Sun Netra CP3240 ATCA 10 GbE Switch Blade (see Figure 24) is a 10 Gigabit Ethernet switch that provides a plug-in upgrade to the Netra CT900 ATCA Blade Server. The switch blade switches the midplane 1Gbps Ethernet Base and 10 GbE Ethernet Extended Fabric. The switch helps to provide sufficient bandwidth for Data Plane applications and for consolidating multiple communication carrier applications onto a single ATCA system platform.

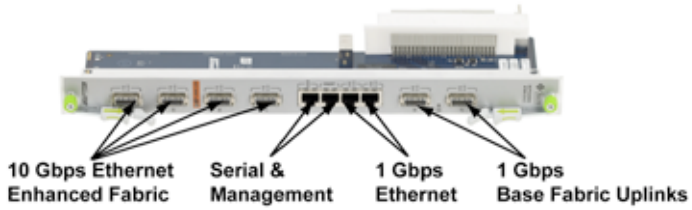

Figure 24. The Sun Netra CP3240 ATCA 10 GbE Switch Blade switches midplane 1Gbps Ethernet base and10 GbE Ethernet extended fabric.

Key Features

The Sun Netra CP3240 ATCA 10 GbE Switch Blade is a PICMG 3.0/3.1 hub board featuring:

- Separate control processors for Base and Extended Fabric
- Fully managed solution with SNMP, serial, Web interface, telnet, and SSH
- Various uplink options via front panel and RTM
- Dual redundant shelf management controller support (cross connect)
- Dual image firmware for disaster recovery
- PICMG 3.0 R2.0 ECN002 compliant
- Base PICMG 3.0 Ethernet interface
- Fabric 3.1 PICMG 3.1 Option 1 and 9 Ethernet interface
- Three front panel-accessible, mid-size AMC slots
- NEBS Level 3-certified

Base Fabric

The switch blade's Base interface supports 1 Gbps Ethernet across the midplane, allowing it to support standard ATCA blades conforming to the PICMG 3.0 Option 1 specification. A pair of switches supports a dual star Base fabric. The switch provides wire-speed performance, Layer 2 switching, Layer 3 routing, full IPv6 support, and Layer 2 and 3 Multicast support. The blade's Base Fabric switching

capability is provided by a Broadcom 24-port Ethernet switch, which provides the following connectivity:

- 12 Gigabit Ethernet connections to the midplane, one connection to each processor blade
- 3 Gigabit Ethernet connections to the three AMC slots
- 2 Gigabit Ethernet connections shared between the front panel and RTM (only one of each can be used at any time)
- 2 Gigabit Ethernet connections cross-connecting to the second switch in a pair
- 1 Gigabit Ethernet connection to the midplane for shelf management
- 1 Gigabit Ethernet connection shared between the front panel and the RTM for management (only one can be used at any time)
- 1 Gigabit Ethernet connection for out-of-band management (internal to the switch)
- 2 10 GbE Ethernet uplinks accessible from the RTM

Extended Fabric

The switch blade's Extended Fabric supports 10 GbE Ethernet switched across the midplane, supporting ATCA blades conforming to the PICMG 3.1 Option 9 specification. A pair of switches supports a dual star Extended Fabric. The switch provides wire-speed performance, Layer 2 switching, Layer 3 routing, full IPv6 support, and Layer 2 and 3 Multicast support. The blade's Extended Fabric switching capability is provided by a Broadcom 20-port 10 GbE Ethernet switch, 18 ports of which provide the following connectivity:

- 12 10 GbE Ethernet connections to the midplane, one connection to each processor blade
- 1 10 GbE Ethernet connection to one of the AMC slots
- 4 10 GbE Ethernet connections (CX4 or SFP+) accessible at the RTM
- 1 10 GbE Ethernet connection cross-connecting to the second switch in a pair

Three AMC Sites

The switch blade provides three mid-height, full-length AMC sites. The sites are enabled for master or line card clocking.

- One site offers 1 Gbps Ethernet connectivity to the Base Fabric (CX4) as well as 10 GbE Ethernet connectivity to the Extended Fabric (XFP).
- Two sites offer 1 Gbps Ethernet connectivity to the Base Fabric (CX4 or SFP)

The AMC sites are AMC.0 R2.0 compliant. The Base Fabric connectivity conforms to AMC.2 Type E1, and the Extended Fabric connectivity conforms to AMC.2 Type 5.

Chapter 9

Sun Netra CT900 ATCA Blade Server

Oracle's Sun Netra CT900 Blade Server is one of the industry's fastest, densest, most reliable ATCA servers. Built on ATCA standards, the Sun Netra CT900 Server allows mixing and matching up to 12 SPARC, Intel Xeon, and AMD Opteron processor-based blades, running Oracle Solaris, Microsoft Windows, or carrier-grade Linux, all in the same enclosure. With up to 30% greater compute density compared to competing systems, and a design that's meant for "six nines" availability, the Sun Netra CT900 Server is an integrated system whose key features make it much more than the sum of its parts. More information about the Sun Netra CT900 ATCA Blade Server is available at http://www.oracle.com/goto/atca.

Key Features

The Sun Netra CT900 ATCA Blade Server offers the following features:

- *Mix and match architectures.* The design allows up to 12 blades to be plugged into a single 12U Sun Netra CT900 ATCA Blade Server. Customers can choose from SPARC, UltraSPARC, Intel Xeon, and AMD Opteron processor blades, and run them side-by-side in the same server. In addition, customers can switch Ethernet on the midplane using the 1 Gbps Ethernet Base Fabric or using the 10 GbE Ethernet Extended Fabric.
- *Operating system support.* Processor blades in the Sun Netra CT900 ATCA Blade Server can support Oracle Solaris, carrier-grade Linux, or Microsoft Windows 2003 or 2008 Advanced Server (depending on the blade). The Sun Netra CT900 Server allows mixing and matching of processors and supported operating systems within the same enclosure, providing flexibility, the opportunity for consolidation, and choice.
- *Redundant components.* The Sun Netra CT900 Blade Server is designed for "six nines" reliability with dual redundant Gigabit Ethernet switches, 10 Gigabit Ethernet switches (optional), shelf managers, power and fan modules. Redundant hardware helps to eliminate single points of failure.
- *Hot-swap components.* Technicians can replace failed components while the server continues to operate. The Sun Netra CT900 Server features hot-swap Gigabit and 10 Gigabit Ethernet switches, shelf managers, and power and fan modules, helping to reduce system downtime to seconds per year.
- *Fault diagnosis software.* The Sun Netra CT900 Blade Server ships with Netra Blade Management Suite (NBMS) software preinstalled. NBMS software enhances system reliability with automatic fault diagnosis, and it interoperates with third-party SAF-compliant products.
- *High density.* The Sun Netra CT900 Server can pack up to up to 144 SPARC T3, 96 UltraSPARC T2, 96 Intel Xeon, or 48 AMD Opteron processor cores into 12 rack units of vertical space, delivering compute density while saving up to 10 inches of space per rack.
- *Shared management environment.* Sun Netra High Availability Suite software supports managing and monitoring blade performance in a mixed OS environment.

- *Common connectors.* Optional Advanced RTMs allow hot swapping of individual ATCA blades without the hassle of recabling.
- *Standards support.* The Sun Netra CT900 server is based on open industry standards including SAF and PICMG, giving customers access to a range of compatible products from independent hardware and software vendors.

Chapter 10

Software and Management

Oracle's Sun Netra CT900 ATCA Blade Server uses standard management interfaces. In addition, Sun software options from Oracle provide the capability to build line-rate packet-processing applications, cluster processor blades in high-availability configurations, virtualize, and consolidate. Of course, all of the Sun ATCA blades are offered with support for the communication carrier operating system of choice — Oracle Solaris.

Shelf Management

The Sun Netra CT900 ATCA Blade Server utilizes two shelf management cards that provide redundant, standard, Intelligent Platform Management Interfaces (IPMI) for controlling platform-level functions and for managing processor blades.

Sun Netra Blade Management Suite software builds on the IPMI infrastructure and automates fault diagnosis and identifies fault causes, helping to improve system reliability and maximize service availability. Sun Netra Blade Management Suite interoperates with third-party SAF-compliant products.

Oracle Solaris

Offering unprecedented flexibility, Oracle Solaris is optimized for UltraSPARC, UltraSPARC T1, UltraSPARC T2, SPARC T3 and AMD Opteron and Intel Xeon processors. Oracle Solaris has its roots in the telecommunications industry, having originated from Bell Labs, and is virtually an industry-standard OS for mission-critical communication carrier applications. Tools such as Oracle Solaris Dynamic Tracing (DTrace) help NEPs to maximize performance for communication carrier applications. Predictive self-healing helps to maximize system and application service availability by automatically diagnosing, isolating, and recovering from system faults as they occur. Oracle Solaris Containers provides an OS-based virtualization platform that provides both security and resource isolation. It allows NEPs to allocate resources to applications as needed to meet peak demand, enabling them to provide predictable service levels even when multiple applications are on the same system. Oracle Solaris security enhancements help to minimize downtime, such as buffer overflow protection at the application and server level, an integrated IP filter, a new cryptographic framework, and process and user rights management. Flash archive capabilities and other new management features enable fast installation of new OS and application software on servers over the network, helping to reduce administrative costs.

Oracle Solaris 10 can be downloaded from http://www.oracle.com/us/products/servers-storage/solaris/index.html. The OpenSolaris open source project software can be downloaded from http://opensolaris.org.

- Oracle VM for SPARC domains — Operating system-independent, virtualized computing environments for Sun UltraSPARC T1, UltraSPARC T2 and SPARC T3 processors.

- MediaLib software — a low-level performance library for multimedia applications, with more than 3,000 functions for image, signal, audio, video processing, and algebra applications. It is accelerated with the VIS and VIS2 features of the SPARC platform, and MMX/SSE/SSE2 on x86 platforms. MediaLib software is bundled with Oracle Solaris 10 or is available via download. Codeclib is optimized with mediaLib software to support a wide range of audio, video, and image CODEC standards. It also includes acoustic echo cancellation, line echo cancellation, Dual-Tone Multi Frequency (DTMF) generator and detector features. A free download is available at: http://www.oracle.com/processors/vis/mlibfiles.html.

In addition, Sun Netra System Management tools provide detailed alarming and fault diagnostic capabilities across multiple blade servers, including Sun blades as well as those from third-party suppliers. The Sun Netra High Availability (HA) Suite supports Service Availability Forum (SAF) compliant carrier-grade high availability.

Furthermore, strict adherence to industry standards and published open interfaces eliminates the vendor lock-in that NEPs often experience with proprietary technologies. The technology within the Sun Unified Network Platform includes OpenSPARC technology, published open interfaces for logical domain virtualization technology, ANSI C as the basis for the common control and Data Plane development environment, and the OpenSolaris operating system, as well as carrier-grade Linux options for the Control Plane environment.

Conclusion

The telecommunications industry is on the threshold of a revolution. As the use of Internet Protocol becomes ever more pervasive in communication carrier networks, central offices, and datacenters, increasing amounts of standard, off-the-shelf technology can be used to provide services that once required costly, purpose-built hardware. The threshold that the industry is about to cross is the ability to use off-the-shelf hardware to build timing-dependent, high-throughput packet-processing applications that typically have required custom hardware.

Oracle is facilitating this transition with a set of products from Sun that provide end-to-end 10 Gigabit Ethernet capabilities within a standard, NEBS Level 3-certified ATCA System Platform. Sun processor blades — based on SPARC T3, UltraSPARC T2, Intel Xeon, and AMD Opteron processors — integrate 10 Gigabit capabilities that use the ATCA midplane to reach the Sun 10 Gigabit switch blade. Sun 10 Gigabit technology is unlike many others, with the ability to support higher line rates by offloading host CPUs with on-chip packet classification. The SPARC T3 processor's capability to support 96 concurrent threads allows developers to program each thread to perform functions that once required custom, proprietary ASICs. For demanding x86 applications, the Sun Netra CP3270 Intel Xeon Server Blade also offers 16 compute threads and 32GB, all within a low 200-watt power envelope, on a single blade.

In addition, three new Advanced RTMs give an unprecedented level of expansion and I/O capabilities to Sun Netra ATCA blades. With Advanced RTM models designed with multiple 10 Gigabit Ethernet links, internal hard disk controllers and drives, and Fibre Channel I/O capabilities, communication carriers and NEPs alike can configure underlying hardware capabilities for virtually any application, including Control Plane, security, media processing, and server computing.

References

WEB SITES	
Oracle's Sun Netra CP32x0 Board User's Guides	www.oracle.com/us/products/servers-storage/servers/netra-carrier-grade/
Sun Netra Carrier-Grade Servers from Oracle	www.oracle.com/us/products/servers-storage/servers/netra-carrier-grade/
UltraSPARC T2 Datasheet	www.oracle.com/processors/UltraSPARC-T2/datasheet.pdf
PAPERS	
"Unleashing 10 Gigabit Networks"	www.oracle.com/products/networking/ethernet/10gigethernet/unleashing_10gb_wp.pdf
"Virtualization with Logical Domains and Sun CoolThreads Servers"	www.oracle.com/servers/coolthreads/ldoms/wp.pdf
Sun Netra Data Plane Software Suite	www.oracle.com/ocom/groups/public/@ocom/documents/webcontent/045170.pdf
OpenSPARC T2 Core Microarchitecture Specification	opensparc-t2.sunsource.net/specs/`OpenSPARCT2_Core_ Micro_Arch.pdf
OpenSPARC T2 System-on-Chip (SoC) Microarchitecture Specification	opensparc-t2.sunsource.net/specs/OpenSPARCT2_SoC_ Micro_Arch.pdf
UltraSPARC T2 Supplement to UltraSPARC Architecture 2007 Specification	opensparc-t2.sunsource.net/specs/UST2-UASuppl-current-draft-HP-EXT.pdf

Appendix A

Power Planning

Planning the installation of a Sun Netra ATCA CT900 Blade Server from Oracle must encompass the task of power planning. This chapter provides background and general guidelines to assist in this task, although specific site and equipment configurations more precisely determine overall power requirements.

ATCA Node Board Power

The Sun Netra ATCA CT900 Blade Server is based on PICMG 3.0 and 3.1 specifications. As such, each node board is designed to meet the general requirement of 200 watts per board. The addition of an Advanced Rear Transition Module (Advanced RTM) extends this power requirement by 25 watts, resulting in a total maximum node power of 225 watts per slot. In specific power calculations, node board power is affected by a number of factors, including CPU type, CPU workload, memory configuration, mezzanine cards (PMC or AMC), and Advanced RTMs.

The number and type of processor cores impacts power use, and CPU workload also represents a significant variable. In Sun ATCA blades, UltraSPARC T1 processors are available in 6- or 8-core configurations, UltraSPARC T2 processors in 8-core configurations, SPARC T3 in 12-core configurations, quad-core Intel Xeon processors, and dual- or quad-core AMD Opteron processors. A significant advantage of multi-core architectures is high memory utilization. In the past, single-core CPUs running memory stress programs exhibited memory power data that was 30-50% higher than testing performed with actual application loads. This is not the case with the UltraSPARC processor's Chip Multi Threading (CMT) architecture. Actual applications yield peak power usage that is nearly the same as that of memory stress tests. These results may occur with UltraSPARC processors because memory clocking is gated when accesses are idle, creating a significant power delta between idle and full memory access loads.

Memory power is also affected by specific DIMM configurations. As memory technologies mature, higher and higher density Integrated Circuits (ICs) become available. As density increases, fewer ICs are needed to produce a given size DIMM, and the number of ICs used generally determines power requirements. (As an example, a 2 GB ECC DIMM can be made with either thirty-six 512 MB DRAM ICs or eighteen 1 GB DRAM ICs. A 1 GB DIMM made with 512 MB will have a similar power demand as a 2 GB DIMM made with 1 GB ICs.) Since node board power is limited, higher density DIMM configurations are possible when memory technology progresses to the point where the power consumption is the same as that of the lower density DIMMs.

Mezzanine cards are another variable in node board power and most new node board architectures have moved to Advanced Mezzanine Card (AMC) designs. AMCs can support a wide variety of configurations from hard drives to Sonet interfaces. The challenge to power planners is that these cards are hot pluggable from the front and can be swapped out.

Advanced RTMs provide similar functionality as an AMC but in a different form factor. IPMC EKeying software verifies AMC and ARTM power and will not power on a node board if the sum of the components exceeds the maximum allowed. For the purposes of power distribution planning, worst case numbers for a given configuration are provided here, but these values should be used as guidelines only — specific configurations will likely differ.

Shelf Power

The ATCA specification requires the support of dual power feeds to all front boards (PICMG 3.0 section 4.1.1). Because of power density and typical central office wiring practices, input power is divided into four power groups, or "zones". Note that the word "zone" in this context is not to be confused with the grouping of midplane signals that are referred to as "zones" in the ATCA PICMG specifications. The term "zone" also can refer to the grouping of node boards, fan trays, and ACBs (see Figure 25). Central Office or Data Center power distribution provides DC power feeds that can vary from 15 to 50 Amps per feed.

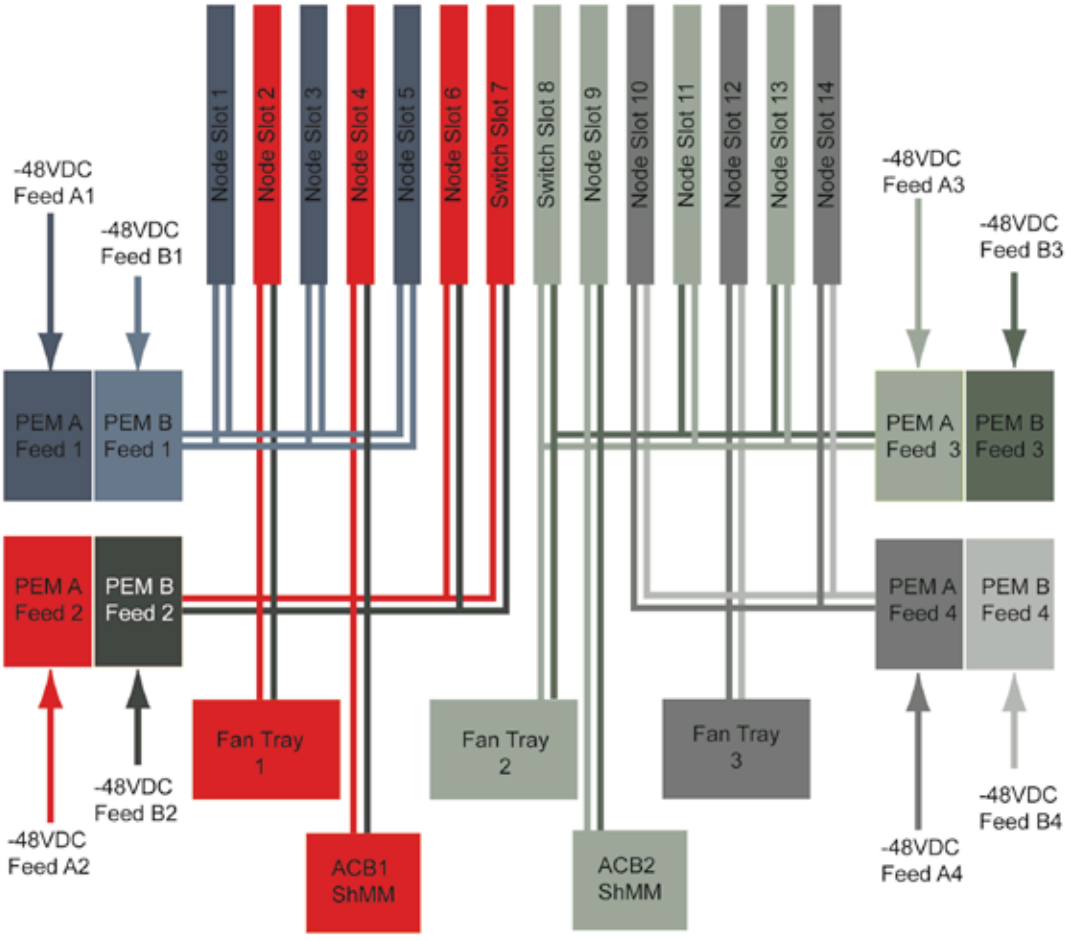

Figure 25. Power distribution varies across zones in the Sun Netra ATCA CT900 Blade Server.

Oracle's Sun Netra ATCA CT900 Blade Server has four 30 Amp input feeds (derated to 28 Amps for safety) to the redundant Power Entry Modules (PEMs A and B). There is an additional loss through the PEMs so maximum current is 27.6 Amps per feed. The Sun Netra ATCA CT900 Blade Server is a 14-slot shelf. As shown in Figure 25, all power zones are not equal (for example, zones 2 and 3 have three node boards, a switch, a fan tray, and an ACB).

One other variable to consider is input voltage. ATCA PICMG 3.0 specifies a minimum input voltage to the shelf of -40.5VDC (PICMG 3.0 section 4.2.2). The actual shutdown limit for node boards is below -39.5VDC to provide for a drop across the midplane. For power planning, -40.5 VDC is used.

Therefore, the worse case power available per feed (or "zone") is as follows:

Power per feed = -40.5volts x 27.6 amps = 1,117 watts / feed

Watts per CT900 shelf = 4 x 1,117 = 4,471 watts / shelf

The maximum power available under worse case conditions (at minimum input voltage) is as follows:

Node Board	200 Watts
ARTM	25 Watts
Switch	200 Watts
Fan Tray	200 Watts
ACB ShMM	10 Watts
3 Node Boards	600 Watts
3 ARTMs	75 Watts
1 Switch	200 Watts
1 Fan Tray	200 Watts
1 ACB/ShMM	14 Watts
Total Power	1,089 Watts

This leaves a 50-watt margin for zones 2 and 3, which have the most loads.

The Sun Netra CP3240 switch can consume 200 Watts of power (it has three AMC slots and can support an Advanced RTM with up to six 10G SFP+ transceivers). Using these numbers to calculate a fully loaded shelf, maximum power use can be estimated as follows:

12 Node Boards	2,400 Watts
12 ARTMs	300 Watts
2 Switches	400 Watts
3 Fan Trays	600 Watts
2 ACB/ShMM	28 Watts
Total Power	3,728 Watts

This is under the total power feed capacity of 4,471 Watts.

For power distribution planning it may not be practical to use different power inputs for different feeds. One approach would be to take the worse case zone and multiply by four. For example, using the worst-case values above for zones 2 and 3, the power available is:

1,089 X 4 = 4,356 Watts

Of course, these estimates are worst case numbers and other node boards can either be configured to be less than 200 Watts or will be less based on the running applications. Estimates presented here are guidelines only and actual load values can vary depending on precise configurations (including memory and peripheral options).

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

Enabling End-to-End 10 Gigabit Ethernet in Oracle's Sun Netra ATCA Product Family
February 2011

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

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

SOFTWARE. HARDWARE. COMPLETE.