



## **SPARC ENTERPRISE T5120, T5220, T5140, AND T5240 SERVER ARCHITECTURE**

Unleashing the UltraSPARC®T2 and UltraSPARC T2 Plus Processors  
with Innovative Multi-core Multi-thread Technology

White Paper  
July 2009

## **TABLE OF CONTENTS**

### **EXECUTIVE SUMMARY 1**

### **THE EVOLUTION OF MULTI-CORE MULTI-THREAD TECHNOLOGY 2**

**BUSINESS CHALLENGES FOR WEB 2.0 2**

**RULE-CHANGING MULTI-CORE MULTI-THREAD TECHNOLOGY 3**

**SPARC ENTERPRISE T5120, T5220, T5140, AND T5240 SERVERS 7**

### **THE ULTRASPARC T2 AND T2 PLUS PROCESSORS 13**

**THE WORLD'S FIRST MASSIVELY THREADED SYSTEM ON A CHIP (SoC) 13**

**TAKING CHIP MULTITHREADED DESIGN TO THE NEXT LEVEL 14**

**ULTRASPARC T2 ULTRASPARC T2 PLUS PROCESSOR ARCHITECTURE 16**

### **SERVER ARCHITECTURE 23**

**SYSTEM-LEVEL ARCHITECTURE 23**

**CHASSIS DESIGN INNOVATIONS 26**

**SPARC ENTERPRISE T5120 SERVER OVERVIEW 28**

**SPARC ENTERPRISE T5220 SERVER OVERVIEW 29**

**SPARC ENTERPRISE T5140 SERVER OVERVIEW 31**

**SPARC ENTERPRISE T5240 SERVER OVERVIEW 33**

**PCI EXPRESS EXPANSION UNIT 36**

### **ENTERPRISE-CLASS MANAGEMENT AND SOFTWARE 38**

**SYSTEM MANAGEMENT TECHNOLOGY 38**

**SCALABILITY AND SUPPORT FOR INNOVATIVE MULTITHREADING TECHNOLOGY 40**

### **CONCLUSION 47**

**FOR MORE INFORMATION 47**

## EXECUTIVE SUMMARY

Use of the Web is changing in fundamental ways, driven by Web 2.0 applications and the thousands of people who join the global Internet every day through a proliferation of new interactive devices. The character of applications and services is changing too. Increasingly, users don't need to install anything, upgrade anything, license anything, subscribe to anything, or even buy anything in order to participate and transact. Web users can even interact directly with content, changing it and improving it. Intellectual property is shared, rather than locked away, and the most popular services are available free of charge. Even very small transactions are now encouraged, becoming large in aggregate. Social networking and other collaborative sites let like-minded people from around the world share information on enormous range of topics and issues. Business transactions too are now predominantly Web based.

Serving this dynamic and growing space is becoming very challenging for datacenter operations. Services need to be able to start small and scale very rapidly, often doubling capacity every three months even as they remain highly available. Infrastructure must keep up with these enormous scalability demands, without generating additional administrative burden. Unfortunately, most datacenters are already severely constrained by both real estate and power — and energy costs are rising. There is also a new appreciation for the role that the datacenter plays in reducing energy consumption and pollution. Virtualization has emerged as an extremely important tool as organizations seek to consolidate redundant infrastructure, simplify administration, and leverage under-utilized systems. Security too has never been more important, with increasing price of data loss and corruption. In addressing these challenges, organizations can ill afford proprietary infrastructure that imposes arbitrary limitations.

Employing the UltraSPARC T2 and UltraSPARC T2 Plus processors — the industry's first massively threaded system on a Chip (SoC) — SPARC Enterprise T5120/T5220 and T5140/T5240 servers offer breakthrough performance and energy efficiency to drive Web 2.0 infrastructure and address other demanding datacenter challenges. Next-generation multi-core multi-thread technology supports up to 128 threads in as little as one rack unit (RU) — providing increased computational density while staying within variously 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. Uniformity of management interfaces and adoption of standards helps reduce administrative costs, while an innovative chassis design shared across SPARC Enterprise T1000 and T2000 servers provides density, efficiency, and economy for modern datacenters. With both the processor and the Solaris™ Operating System (Solaris OS) available under open source licensing, organizations are free to innovate and join with a world-wide technical community.

## Chapter 1

# The Evolution of Multi-core Multi-thread Technology

By any measure, first-generation multi-core multi-thread processors were an unprecedented success. SPARC Enterprise T1000 and T2000 servers based on the UltraSPARC T1 processor with innovative multi-core multi-thread technology won enthusiastic praise, and generated the fastest product ramp in server history. Delivering up to five times the throughput in a quarter of the space and power, these systems even garnered the first ever rebate from a power utility<sup>1</sup>— a trend that is being repeated across the world. Now multi-core multi-thread technology is evolving rapidly to meet the constantly changing demands of a wide range of Web and other applications.

*1. In August of 2006, Pacific Gas and Electric (PG&E) began offering a substantial energy rebate for purchasing and deploying SPARC Enterprise T1000 and T2000 servers*

## Business Challenges for Web 2.0

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

### • *Building out for Web Scale*

Web scale applications engender a new pace and urgency to infrastructure deployment. Organizations must accelerate time to market and time to service, while delivering scalable high-quality and high-performance applications and services. Many need to be able to start small with the ability to scale very quickly, with new customers and innovative new Web services often implying a doubling of capacity in months rather than years.

At the same time, organizations must reduce their environmental impact by working within the power, cooling, and space available in their current datacenters.

Operational costs too are receiving new scrutiny, along with system administrative costs that can account for up to 40 percent of an IT budget.

Simplicity and speed are paramount, giving organizations the ability to respond quickly to dynamic business conditions. Organizations are also striving to eliminate vendor lock-in as they look to preserve previous, current, and future investments.

Open platforms built around open standards help provide maximum flexibility while reducing costs of both entry and exit.

- ***Driving Datacenter Virtualization and Eco-Efficiency***

Coincident with the need to scale services, many datacenters are recognizing the advantages of deploying fewer standard platforms to run a mixture of commercial and technical workloads. This process involves consolidating under-utilized and often sprawling server infrastructures with effective virtualization solutions that serve to enhance business agility, improve disaster recovery, and reduce operating costs. This focus can help reduce energy costs and break through datacenter capacity constraints by improving the amount of realized performance for each watt of power the datacenter consumes.

Eco-efficiency provides tangible benefits, improving *ecology* by reducing carbon footprint to meet legislative and corporate social responsibility goals, even as it improves the *economy* of the organization paying the electric bill. As systems are consolidated onto more dense and capable computing infrastructure, demand for datacenter real estate is also reduced. With careful planning, this approach can also improve service uptime and reliability by reducing hardware failures resulting from excess heat load. Servers with high levels of standard reliability, availability, and serviceability (RAS) are now considered a requirement.

- ***Securing the Enterprise at Speed:***

Organizations are increasingly interested in securing all communications with their customers and partners. Given the risks, end-to-end encryption is essential in order to inspire confidence in security and confidentiality. Encryption is also increasingly important for storage, helping to secure stored and archived data even as it provides a mechanism to detect tampering and data corruption.

Unfortunately, the computational costs of increased encryption can increase the burden on already over-taxed computational resources. Security also needs to take place at line speed, without introducing bottlenecks that can impact the customer experience or slow transactions. Solutions must help to ensure security and privacy for clients and bring business compliance for the organization, all without impacting performance or increasing costs.

## **Rule-Changing Multi-core Multi-thread Technology**

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

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

The oft-quoted tenant of Moore's Law states that the number of transistors that will fit in a square inch of integrated circuitry will approximately double every two years.

For over three decades the pace of Moore's law has held, driving processor performance to new heights. Processor manufacturers have long exploited these gains in chip real estate to build increasingly complex processors, with instruction-level parallelism (ILP) as a goal. Today these traditional processors employ very high frequencies along with a variety of sophisticated tactics to accelerate a single instruction pipeline, including:

- Large caches
- Superscalar designs
- Out-of-order execution
- Very high clock rates
- Deep pipelines
- Speculative pre-fetches

While these techniques have produced faster processors with impressive-sounding multiple-gigahertz frequencies, they have largely resulted in complex, hot, and power-hungry processors that are not well suited to the types of workloads often found in modern datacenters. In fact, many datacenter workloads are simply unable to take advantage of the hard-won ILP provided by these processors. Applications with high shared memory and high simultaneous user or transaction counts are typically more focused on processing a large number of simultaneous threads (thread-level parallelism, TLP) rather than running a single thread as quickly as possible (ILP).

Making matters worse, the majority of ILP in existing applications has already been extracted and further gains promise to be small. In addition, microprocessor frequency scaling itself has leveled off because of microprocessor power issues. With higher clock speeds, each successive processor generation has seemingly demanded more power than the last, and microprocessor frequency scaling has leveled off in the 2-3 GHz range as a result. Deploying pipelined Superscalar processors requires more power, limiting this approach by the fundamental ability to cool the processors.

### **Multiprocessing with Multicore Processors**

To address these issues, many in the microprocessor industry have used the transistor budget provided by Moore's Law to group two or even four conventional processor cores on a single physical die — creating multicore processors. The individual processor cores introduced by many multiprocessing designs have no greater performance than previous single-processor chips, and in fact, have been observed to run single-threaded applications more slowly than single-core processor versions. However, the aggregate chip performance increases since multiple programs (or multiple threads) can be accommodated in parallel (thread level parallelism).

Unfortunately, most currently-available (or soon to be available) chip multiprocessors simply replicate cores from existing (single-threaded) processor designs. This approach typically yields only slight improvements in aggregate performance since it

ignores key performance issues such as memory speed and hardware thread context switching. As a result, while these designs provide some additional throughput and scalability, they can consume considerable power and generate significant heat — without a commensurate increase in overall performance.

### Innovative Multi-core Multi-thread Technology

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

Unfortunately, this relative gap between processor and memory speeds leaves ultra-fast processors idle as much as 85 percent of the time, waiting for memory transactions to complete. Ironically, as traditional processor execution pipelines get faster and more complex, the effect of memory latency grows — fast, expensive processors spend more cycles doing nothing. Worse still, idle processors continue to draw power and generate heat. It is easy to see that frequency (gigahertz) is truly a misleading indicator of real performance.

First introduced with the UltraSPARC T1 processor, multi-core multi-thread technology takes advantage of multiprocessing advances, but adds a critical capability —the ability to scale with threads rather than frequency. Unlike traditional single-threaded processors and even most current multicore processors, hardware multithreaded processor cores allow rapid switching between active threads as other threads stall for memory. Figure 1 illustrates the difference between multiprocessing, fine-grained hardware Multithreading (FG-MT), and multi-core multi-thread processor. The key to this approach is that each core in a multi-core multi-thread processor is designed to switch between multiple threads on each clock cycle. As a result, the processor's execution pipeline remains active doing real useful work, even as memory operations for stalled threads continue in parallel.

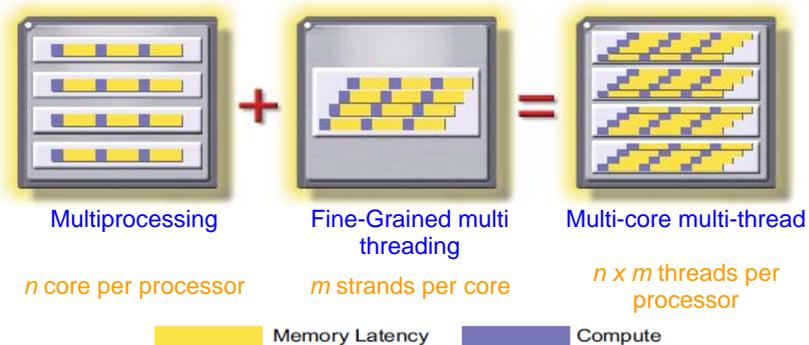


Figure 1. Multi-core Multi-thread technology combines Multiprocessing and fine-grained hardware multithreading

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

### **The UltraSPARC T2 and UltraSPARC T2 Plus Processors**

Unlike complex single-threaded processors, multithreaded processors utilize the available transistor budget to implement multiple hardware multithreaded processor cores on a chip die. The UltraSPARC T2 and UltraSPARC T2 Plus processors takes the multi-core multi-thread model to the next level, providing up to eight cores per processor with each core supporting up to eight threads via two independent pipelines — effectively doubling the throughput of the UltraSPARC T1 processor without raising the clock frequency. In addition, these processors use the increased transistor budget resulting from the use of a 65 nm silicon technology to implement the industry's first massively threaded System on a Chip (SoC), with a single processor die hosting:

- Up to 64 threads per processor (up to eight cores supporting eight threads each)
- On-chip Level-1 and Level-2 caches
- Per-core floating point capabilities
- Per-core cryptographic acceleration
- Two on-chip 10 Gb Ethernet interfaces(UltraSPARC T2 processor only)
- On-chip PCI Express interface
- On-chip cache coherency logic and links (UltraSPARC T2 Plus processor only)

Through this SoC design, the UltraSPARC T2 processor significantly enhances the general purpose nature of the CPU by building in eight floating point units (1 per core). Enhanced floating point capabilities open the UltraSPARC T2 to the world of compute-intensive applications as well as the traditionally multi-core multi-thread friendly datacenter throughput applications. No-cost security and cryptographic acceleration is provided by the on-chip, per-core streaming accelerators. In addition, the ability to move data in and out of the UltraSPARC T2 processor is significantly aided by an integrated PCI Express interface and dual 10 Gigabit Ethernet interfaces on the UltraSPARC T2 processor. The UltraSPARC T2 Plus processor replaces the

dual 10 Gigabit Ethernet interfaces with cache coherency logic and links that facilitate multisocket system designs.

## **SPARC Enterprise T5120, T5220, T5140, and T5240 Servers**

SPARC Enterprise T5120/T5220 and T5140/T5240 servers are designed to leverage the considerable resources of the UltraSPARC T2 and UltraSPARC T2 Plus processor in the form of cost-effective general-purpose platforms. SPARC Enterprise T5120 and T5220 servers deliver up to twice the throughput of their predecessors, while leading competitors in terms of performance, performance per watt, and per space. SPARC T5140 and T5240 servers extend this scalability by adding dual-sockets for UltraSPARC T2 Plus processors and considerable large memory support. All of these systems also extend the benefits of multi-core multi-thread from multithreaded commercial workloads into technical workloads rich in floating point operations.



SPARC Enterprise T5120 server



SPARC Enterprise T5140 server



SPARC Enterprise T5220 server



SPARC Enterprise T5240 server

*Figure 2. SPARC Enterprise T5120, T5220, T5140, and T5240 servers (top to bottom)*

## Overview

With support for 128 threads, large memory, cryptographic acceleration, and integrated on-chip I/O technology, these servers represent a departure from traditional system design. The 1U SPARC Enterprise T5120 and T5140 servers are ideal for providing high throughput within significant power, cooling, and space constraints, delivering either one or two 64-thread UltraSPARC T2 processors in a space- and power-efficient 1U rackmount package. As a compute node with massively horizontally scaled environments, the SPARC Enterprise T5120 or T5140 servers can help provide a substantial building block for application tier or Web services infrastructure. Network infrastructure applications such as portal, directory, network identity, file service, and backup are all a good fit for this server.

The SPARC Enterprise T5220 and T5240 servers provide both throughput as well as expandability, with extra I/O and internal disk options afforded by the 2U rackmount form factor. With greater I/O and internal disk, typical workloads include demanding mid-tier application server deployments or Web- and application-tier consolidation, virtualization, and consolidation projects requiring maximum uptime with future growth and integration into diverse environments. The SPARC Enterprise T5220 and T5240 servers are also ideal for OLTP database deployments.

Designed to complement each other, as well as the rest of server product line, the SPARC Enterprise T5120/T5220 and T5140/T5240 servers address the dynamic needs of the modern datacenter.

- ***Efficient and Predictable Scalability***

With support for 64 threads and large memories, SPARC Enterprise T5120 and T5220 servers are the first to utilize the 10 Gb Ethernet, I/O, and cryptographic acceleration provided directly by the UltraSPARC T2 processor chip. This approach provides leading levels of performance and scalability with extremely high levels of power, heat, and space efficiency.

SPARC Enterprise T5140 and T5240 servers extend this breakthrough compute and memory density, delivering up to 128 threads in as little as 1 rack unit, while typically consuming less than 700 watts of power. These systems also deliver twice the I/O bandwidth of the T5120 and T5220 servers by providing a PCI Express root complex associated with each socket.

- ***Accelerated Time to Market***

SPARC Enterprise T5120/T5220 and T5140/T5240 servers running the Solaris OS provide full binary compatibility with earlier UltraSPARC systems, preserving investments and speeding time to market.

- ***Simplified Management***

Each SPARC Enterprise T5120/T5220 and T5140/T5240 servers all provide an Integrated Lights Out Management (ILOM) service processor, compatible with

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

- ***Industry-Leading Tools for Virtualization and Consolidation***

Multi-core multi-thread technology is ideal for consolidation, providing low-level multi-core multi-thread support for virtualization at every layer of the technology stack. Logical Domains exploit the UltraSPARC T2 and UltraSPARC T2 Plus processor's 64 threads per socket to support multiple guest operating system instances, while Solaris Containers provide virtualization within a single Solaris OS instance. The advanced Solaris ZFS file system provides storage virtualization for storage and considerable scalability.

- ***A Tradition of Leading Eco Efficiency***

SPARC Enterprise T1000 and T2000 servers were the industry's first eco-responsible servers. SPARC Enterprise T5120/T5220 and T5140/T5240 servers continue this tradition by offering the best performance and performance-per-watt across a wide range of commercial and technical workloads. In addition, the UltraSPARC T2 processor was the first and only processor to incorporate unique power management features at both core and memory levels of the processor.

- ***System and Datacenter Reliability***

Reliability is key to keeping applications available and costs down. With the greater levels of integration provided by an SoC design, the SPARC Enterprise T5120/T5220 and T5140/T5240 servers offer greatly reduced part counts, and provide commensurately higher levels of reliability, availability, and serviceability (RAS). Lower power consumption and higher performance per watt greatly reduce generated heat loads and the associated issues they cause. Technologies such as Solaris Predictive Self Healing are integrated with the hardware, and help keep systems available.

- ***Zero-Cost Security***

Providing secure communications and data protection has never been more important, with attempted electronic intrusion and theft at an all-time high. With up to eight integrated cryptographic accelerators on each UltraSPARC T2 and UltraSPARC T2 Plus processor, there is simply no need to send plain text on the network or store plain text in storage systems. SPARC Enterprise T5120, T5220, T5140, and T5240 servers support many more crypto operations per second than competitive processors and dedicated crypto accelerator cards, with minimal impact to system overhead.

## **Innovative System Design**

Beyond the capabilities of individual systems, Fujitsu understands that datacenters have unique and pressing needs that require attention on the part of system designers. Density, performance, and scalability are all essential considerations, but systems must also be serviceable and fit in with modern datacenter strategies that consider power, cooling, and serviceability. SPARC Enterprise T5120/T5220 and T5140/T5240 servers share an innovative design philosophy. Principles of this philosophy include:

- **Maximum Compute Density** — SPARC Enterprise T5120/T5220 and T5140/T5240 servers provide leading density in terms of CPU cores, memory, storage and I/O. This focus on density often lets 1 rack unit (1 RU) rackmount servers replace competitive 2U rackmount servers, for a 50-percent space savings.
- **Leading Storage Capacity** — SPARC Enterprise T5120/T5220 and T5140/T5240 servers provide leading density. Smaller disk drives and innovations in structure, airway, and carrier design allow more disk capacity in smaller spaces, while enhancing system airflow.
- **Continued Investment Protection** — Fujitsu designs for maximum investment protection. Even with breakthrough technology such as multi-core multi-thread technology, Solaris Binary Compatibility means that applications simply run without modification.

Table 1 compares the features of SPARC Enterprise T5120, T5220, T5140, and T5240 servers.

Table 1. SPARC Enterprise T5120, T5220, T5140, T5240 server features

Feature	SPARC Enterprise T5120 Server	SPARC Enterprise T5220 Server	SPARC Enterprise T5140 Server	SPARC Enterprise T5240 Server
CPU	Four-, or eight-core 1.2 GHz, eight-core 1.4 GHz, or eight-core 1.6 GHz. UltraSPARC T2 processor	Four-, or eight-core 1.2 GHz, eight-core 1.4 GHz, , or eight-core 1.6 GHz. UltraSPARC T2 processor	Dual four-, six-, or eight-core 1.2 GHz., or eight-core 1.4 GHz. UltraSPARC T2 Plus processors	Dual four-, six-, or eight-core 1.2 GHz, eight-core 1.4 GHz or eight-core 1.6 GHz. UltraSPARC T2 Plus
Threads	Up to 64	Up to 64	Up to 128	Up to 128
Memory capacity	Up to 128 GB (2, 4, or 8 GB FBDIMM)	Up to 128 GB (2, 4, or 8 GB FBDIMM)	Up to 128 GB (2, 4, or 8 GB FBDIMM)	Up to 256 GB (2, 4, or 8 GB FBDIMM) <sup>d</sup>
Maximum internal disk drives	Up to eight SFF 2.5-inch SAS 73, 146, or 300 GB disk drives	Up to sixteen SFF SAS 2.5-inch SAS 73, 146, or 300 GB disk drives	Up to eight SFF 2.5-inch SAS 73, 146, or 300 GB disk drives	Up to sixteen SFF 2.5-inch SAS 73, 146, or 300 GB disk drives <sup>b</sup>
Removable and pluggable I/O	One CD-RW, DVD+/-RW Drive, Four USB 2.0 ports	One CD-RW, DVD+/-RW Drive, Four USB 2.0 ports	One CD-RW, DVD+/-RW Drive, Four USB 2.0 ports	One CD-RW, DVD+/-RW Drive, Four USB 2.0 ports
PCI	One x8 PCI Express slot, Two x4 PCI Express or XAUI combo slots <sup>b</sup>	Two x8 PCI Express Two x4 PCI Express Two x4 PCI Express or XAUI combo slots <sup>b</sup>	One x8 PCI Express slot, Two x8 PCI Express or XAUI combo slots <sup>b</sup>	Four x8 PCI Express slot, Two x8 PCI Express or XAUI combo slots <sup>b</sup>
Ethernet	Four on-board Gigabit Ethernet ports (10/100/1000) Two 10 Gb Ethernet ports via XAUI combo slots	Four on-board Gigabit Ethernet ports (10/100/1000) Two 10 Gb Ethernet ports via XAUI combo slots	Four on-board Gigabit Ethernet ports (10/100/1000) Two 10 Gb Ethernet ports via XAUI combo slots <sup>c</sup>	Four on-board Gigabit Ethernet ports (10/100/1000) Two 10 Gb Ethernet ports via XAUI combo slots <sup>c</sup>
Power supplies	Two hot-swappable AC 720W or DC 660W power supply units (N+1 redundancy)	Two hot-swappable AC 750W or 1,100W or DC 1,200W power supply units (N+1 redundancy)	Two hot-swappable AC 720W or DC 660W power supply units(N+1 redundancy)	Two hot-swappable AC 1,100W or 1,200W power supply units (N+1 redundancy)
Fans	Four hot-swappable fan trays, with 2 fans per tray, N+1 redundancy	Three hot-swappable fan trays, with 2 fans per tray, N+1 redundancy	Six hot-swappable fan trays, with two fans per tray, N+1 redundancy	Five hot-swappable fan trays, with two fans per tray, N+1 redundancy
Form factor	1 rack unit (1U)	2 rack units (2U)	1 rack unit (1U)	2 rack units (2U)
Operating System	Solaris 10 8/07 and 10/08 OS	Solaris 10 8/07 and 10/08 OS	Solaris 10 8/07 and 10/08 OS	Solaris 10 8/07 and 10/08 OS

a. Eight-disk backplane is not supported with 1.6 GHz CPU (Fujitsu SPARC Enterprise T5120/T5240 server) or with 1.4 GHz CPU (Fujitsu SPARC Enterprise T5140 server)

b. Optional XAUI adapter cards required for access to dual 10 Gb Ethernet ports on all systems. Each XAUI consumes a PCI Express slot

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

d. 256 GB maximum memory configuration requires the optional Memory Mezzanine Kit

## **Leading Reliability, Availability, and Serviceability (RAS)**

The SPARC Enterprise T5120/T5220 and T5140/T5240 servers provide excellent reliability, availability, and serviceability (RAS) characteristics. Highly-reliable parts and a relatively low total component count minimize the opportunity for system errors. Dual PCI Express root complexes and the ability to configure multiple processors on SPARC Enterprise T5140/T5240 servers add to resiliency. In addition, these servers include core and thread offlining capabilities, integrated disk RAID functions, and extensive ECC hardware protection — along with redundant hot-swap disks, power supplies, and fans. The following key design elements in the SPARC Enterprise T5120/T5220 and T5140/T5240 servers are key to improving the dependability of IT services:

- Reduced parts count
- Processor thread and core offlining and built-in RAID capabilities
- Redundancy and hot-swap components
- Parity protection and error correction capabilities
- System monitoring
- Integrated Lights Out Management (ILOM) service processor
- Superior energy efficiency
- Robust virtualization technology
- Comprehensive fault management

## Chapter 2

# The UltraSPARC T2 and T2 Plus Processors

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

## The World's First Massively Threaded System on a Chip (SoC)

The UltraSPARC T2 and UltraSPARC T2 Plus processors eliminate the need for expensive custom hardware and software development by integrating computing, networking, security, and I/O on to a single chip. Binary compatible with earlier UltraSPARC processors, no other processor delivers so much performance in so little space and with such small power requirements — letting organizations rapidly scale the delivery of new network services with maximum efficiency and predictability. The UltraSPARC T2 and UltraSPARC T2 Plus processors are shown in Figure 3, to the left of the previous-generation UltraSPARC T1 processor. Even with twice the computational throughput and significantly higher levels of integration, the UltraSPARC T2 and UltraSPARC T2 Plus processors are physically smaller than the UltraSPARC T1 processor.



Figure 3. The and UltraSPARC T2 Plus, UltraSPARC T2) and UltraSPARC T1 Processors with Innovative Multithreading Technology(left to right respectively)

Table 2 provides a comparison between the UltraSPARC T1, UltraSPARC T2 and UltraSPARC T2 Plus processor.

Table 2. UltraSPARC T1 ,T2 and T2 Plus processor features

Feature	UltraSPARC T1 Processor	UltraSPARC T2 processor	UltraSPARC T2 Plus processor
Cores per processor	Up to 8	Up to 8	Up to 8
Threads per core	4	8	8
Threads per processor	32	64	64
Hypervisor	Yes	Yes	Yes
Sockets Supported	1	1	2 or 4 <sup>a</sup>
Memory	4 memory controllers, 4 DIMMs per controller	4 memory controllers, up to 16 FB-DIMMs	2 memory controllers, up to 16 or 32 FB-DIMMs
Caches	16 KB instruction cache, 8 KB data cache, 3 MB L2 cache (4 banks, 12-way associative)	16 KB instruction cache, 8 KB data cache, 4 MB L2 cache (8 banks, 16-way associative)	16 KB instruction cache, 8 KB data cache, 4 MB L2 cache (8 banks, 16-way associative)
Technology	9-layer Cu metal, CMOS process, 90 nm technology	65 nm technology	65 nm technology
Floating point	1 FPU per chip	1 FPU per core, 8 FPUs per chip	1 FPU per core, 8 FPUs per chip
Integer resources	Single execution unit	2 integer execution units per core	2 integer execution units per core
Cryptography	Accelerated modular arithmetic operations (RSA)	Stream processing unit per core, support for the 10 most popular ciphers	Stream processing unit per core, support for the 10 most popular ciphers
Additional on-chip resources		Dual 10 Gb Ethernet interfaces, PCI Express interface (x8)	PCI Express interface (x8), Coherency logic and links (4.8 Gb/second)

a. Two-socket implementations include the SPARC Enterprise T5140 and T5240 servers

## Taking Chip Multithreaded Design to the Next Level

When the next-generation of multi-core multi-thread processors set to be designed, they started with key goals in mind:

- Increasing computational capabilities to meet the growing demand from Web applications by providing twice the throughput of the UltraSPARC T1 processor
- Supporting larger and more diverse workloads with greater floating point performance
- Powering faster networking to serve new network-intensive content
- Providing end-to-end datacenter encryption
- Increasing service levels and reducing downtime
- Improving datacenter capacities while reducing costs

Multi-core multi-thread architecture is ultimately very flexible, allowing different modular combinations of processors, cores, and integrated components. The considerations listed above drove an internal engineering effort that compared different approaches with regard to making improvements on the successful

UltraSPARC T1 architecture. For example, simply increasing the number of cores would have gained additional throughput, but would have resulted in consuming extra die area, leaving no room for integrated components such as floating point processors.

The final UltraSPARC T2 and UltraSPARC T2 Plus processor designs recognize that memory latency is truly *the* bottleneck to improving performance. By increasing the number of threads supported by each core, and by further increasing network bandwidth, these processors are able provide approximately twice the throughput of the UltraSPARC T1 processor.

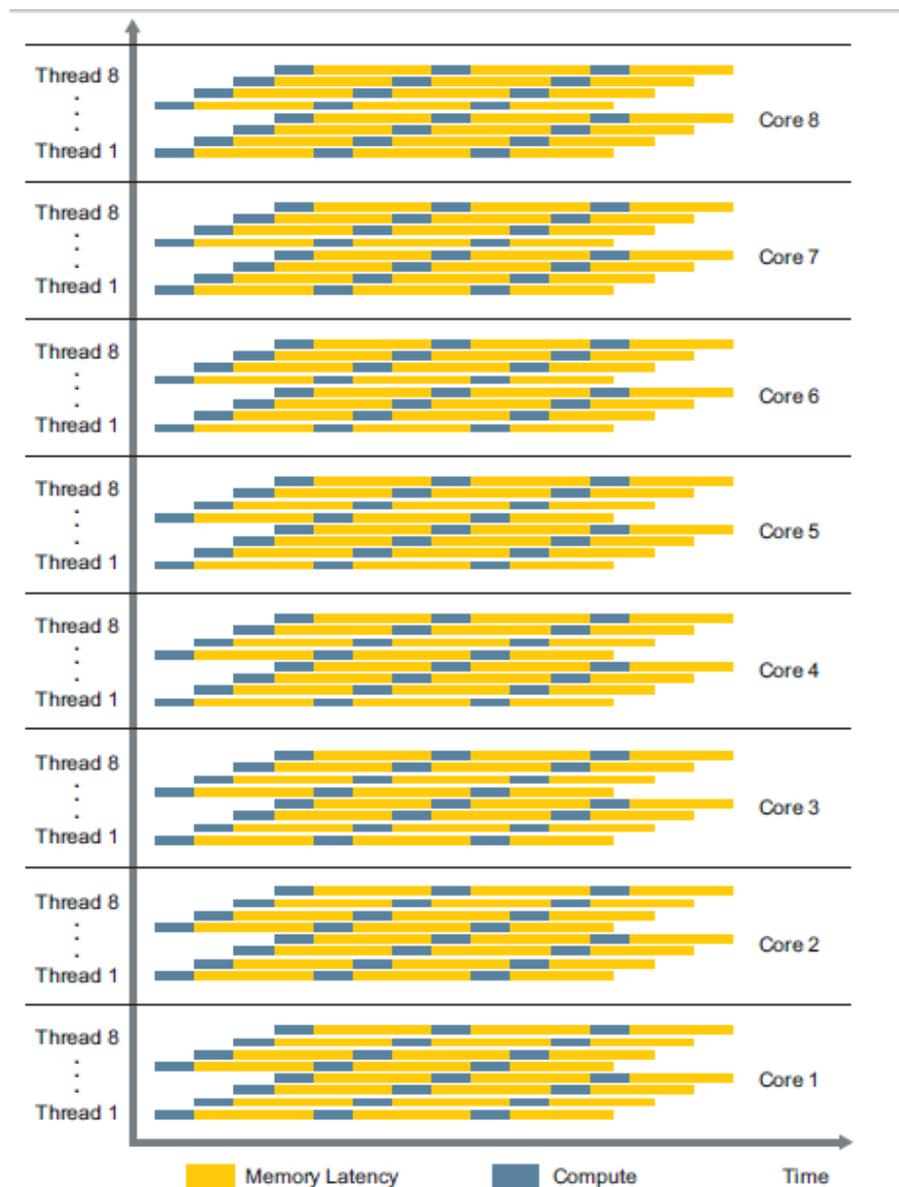


Figure 4. A single eight-core UltraSPARC T2 or UltraSPARC T2 Plus processor supports up to 64 threads, with up to two threads running in each core simultaneously

Each UltraSPARC T2 and UltraSPARC T2 Plus processor provides up to eight cores, with each core able to switch between up to eight threads (64 threads per processor). In addition, each core provides two integer execution units, so that a single UltraSPARC core is capable of executing two threads at a time. Figure 4 provides a simplified high-level illustration of the thread model supported by an eight-core UltraSPARC T2 or UltraSPARC T2 Plus processor.

## UltraSPARC T2 UltraSPARC T2 Plus Processor Architecture

The UltraSPARC T2 processor and the UltraSPARC T2 Plus processor extend Throughput Computing initiative with an elegant and robust architecture that delivers real performance to applications.

### UltraSPARC T2 Processor Architecture

A high level block diagram of the UltraSPARC T2 processor is shown in Figure 5.

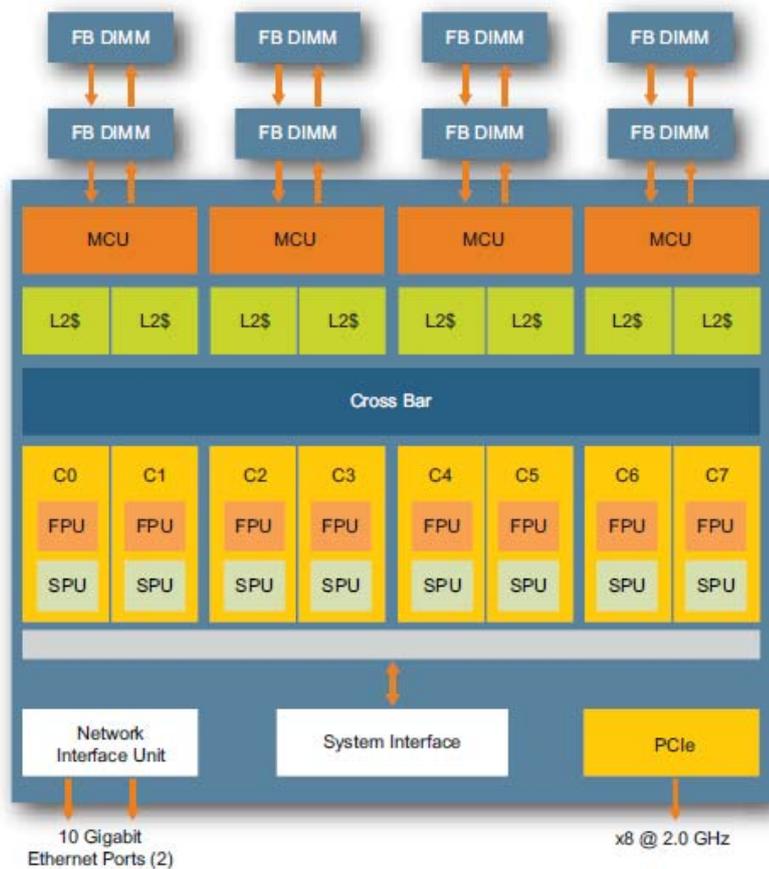


Figure 5. The UltraSPARC T2 processor combines eight cores, memory management, cryptographic support, 10 Gb Ethernet, and PCI Express on a single chip

The eight cores on the UltraSPARC T2 processor are interconnected with a full on-chip non-blocking 8 x 9 crossbar switch. The crossbar connects each core to the eight banks of L2 cache, and to the system interface unit for IO. The crossbar provides approximately 300 GB/second of bandwidth and supports 8-byte writes from a core to a bank and 16-byte reads from a bank to a core. The system interface unit connects networking and I/O directly to memory through the individual cache banks. Using FB-DIMM memory supports dedicated northbound and southbound lanes to and from the caches to accelerate performance and reduce latency. This approach provides higher bandwidth than with DDR2 memory, with up to 42.4 GB/second of read bandwidth and 21 GB/second of write bandwidth.

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

### UltraSPARC T2 Plus Processor Architecture

Figure 6 provides a block-level diagram of the UltraSPARC T2 Plus processor.

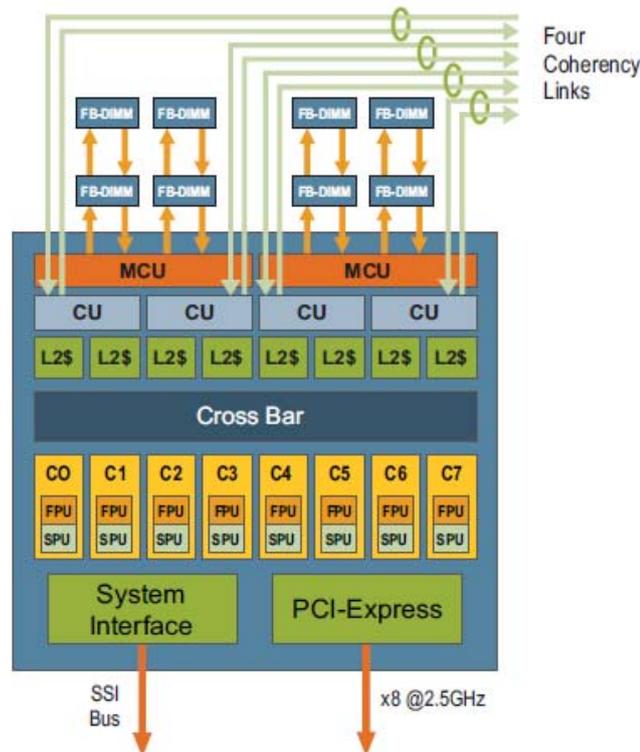


Figure 6. The UltraSPARC T2 Plus processor provides four Coherency Links to connect to up to four other processors

The UltraSPARC T2 Plus architecture omits the dual on-chip 10 Gb Ethernet interfaces and uses the on-chip real estate to provide four on-chip Coherency Units (CUs). The processor also replaces two memory channels with four Coherence Channels (or coherence links) — one provided by each Coherency Unit. These links run a cache coherence (snoopy) protocol over an FB-DIMM-like physical interface to provide up to 4.8 Gigatransfers per port, providing over 200 GB per second in each direction. The memory link speed of the UltraSPARC T2 Plus processor was also increased to 4.8 Gbps over the 4.0 Gbps of the UltraSPARC T2 processor.

The UltraSPARC T2 Plus processor can support both two- and four-socket implementations. A typical two-socket implementation is shown in Figure 7. Dual-socket UltraSPARC T2 Plus implementations interconnect the processors' four coherence links, and require no additional circuitry.

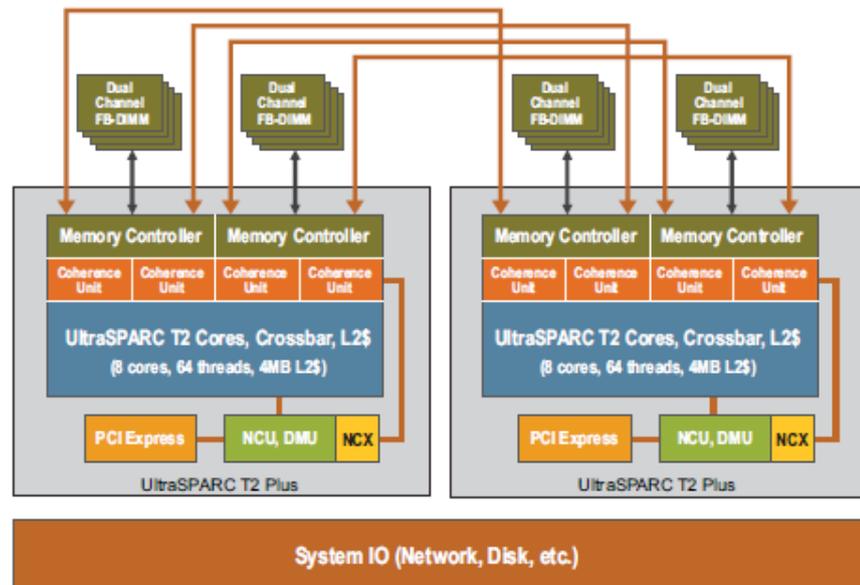


Figure 7. A typical dual-socket UltraSPARC T2 Plus configuration

## Core Architecture and Pipelines

Both the UltraSPARC T2 and UltraSPARC T2 Plus processors share the same core design. Figure 6 provides a block-level diagram representing a single UltraSPARC core on the UltraSPARC T2 processor (up to eight are supported per processor).

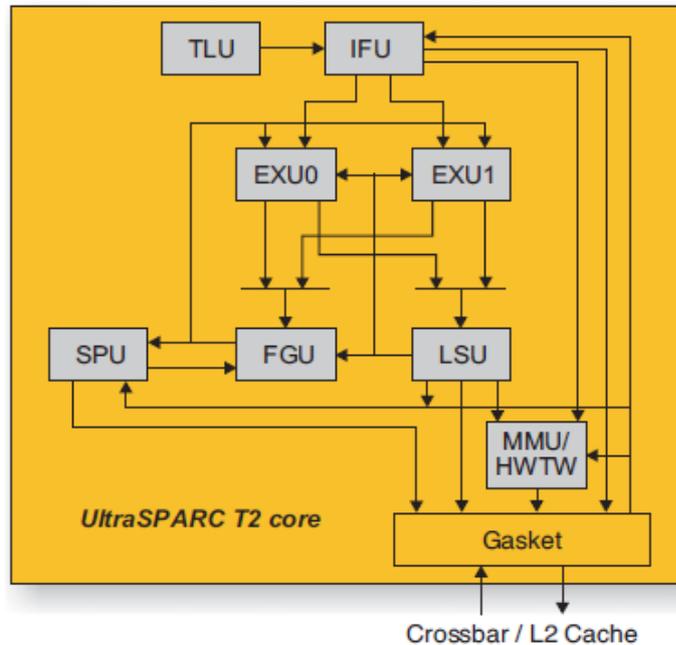


Figure 6. UltraSPARC T2 and UltraSPARC T2 Plus core block diagram

Components implemented in each core include:

- Trap Logic Unit (TLU)— The trap logic unit updates the machine state as well as handling exceptions and interrupts.
- Instruction Fetch Unit (IFU)— The instruction fetch unit includes the 16KB instruction cache (32-byte lines, 8-way set associative) and a 64-entry fully-associative instruction translation lookup buffer (ITLB).
- Integer Execution Units (EXU)— Dual integer execution units are provided per core with four threads sharing each unit. Eight register windows are provided per thread, with 160 integer register file (IRF) entries per thread.
- Floating Point/Graphics Unit (FGU)— A floating point/graphics unit is provided within each core and it is shared by all eight threads assigned to the core. 32 floating point register file entries are provided per thread.
- Stream Processing Unit (SPU)— Each core contains a stream processing unit that provides cryptographic coprocessing.
- Memory Management Unit (MMU) — The memory management unit provides a hardware table walk (HWTW) and supports 8 KB, 64 KB, 4 MB, and 256 MB pages.

An eight-stage integer pipeline and a 12-stage floating-point pipeline are provided by each UltraSPARC T2 and UltraSPARC T2 Plus processor core (Figure 7). A new “pick” pipeline stage has been added to choose two threads (out of the eight possible per core) to execute each cycle.

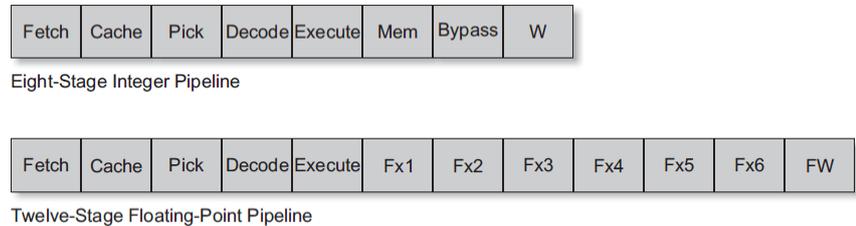


Figure 7. UltraSPARC T2 per-core integer and floating-point pipelines

To illustrate how the dual pipelines function, Figure 8 depicts the integer pipeline with the load store unit (LSU). The instruction cache is shared by all eight threads within the core. A least-recently-fetched algorithm is used to select the next thread to fetch. Each thread is written into a thread-specific instruction buffer (IB) and each of the eight threads is statically assigned to one of two thread groups within the core.

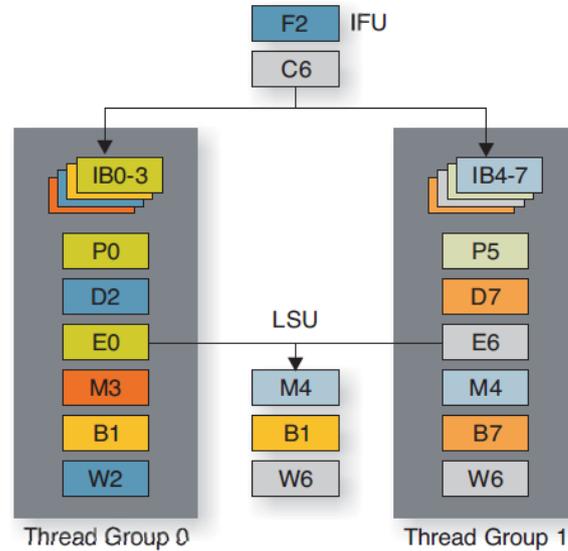


Figure 8. Threads are interleaved between pipeline stages with very few restrictions (integer pipelines shown, letters depict pipeline stages, numbers depict different scheduled threads)

The “pick” stage chooses one thread each cycle within each thread group. Picking within each thread group is independent of the other, and a least-recently-picked algorithm is used to select the next thread to execute. The decode state resolves resource conflicts that are not handled during the pick stage. As shown in the illustration, threads are interleaved between pipeline stages with very few restrictions.

Any thread can be at the fetch or cache stage, before being split into either of the two thread groups. Load/store and floating point units are shared between all eight threads. Only one thread from either thread group can be scheduled on such a shared unit.

### **Integrated Networking**

By providing integrated on-chip networking, the UltraSPARC T2 processor is able to provide better networking performance. All network data is supplied directly from and to main memory. Placing networking so close to memory reduces latency, provides higher memory bandwidth, and eliminates inherent inefficiencies of I/O protocol translation.

The UltraSPARC T2 processor provides two 10 Gb Ethernet ports with integrated SerDes, offering line-rate packet classification at up to 30 million packets/second (based on layers 1-4 of the protocol stack). Multiple DMA engines (16 transmit and 16 receive DMA channels) match DMAs to individual threads, providing binding flexibility between ports and threads. Virtualization support includes provisions for eight partitions, and interrupts may be bound to different hardware threads.

### **Stream Processing Unit**

The stream processing unit on each UltraSPARC T2 core runs in parallel with the core at the same frequency. Two independent sub-units are provided along with a DMA engine that shares the core's crossbar port:

- A Modular Arithmetic Unit (MAU) shares the FGU multiplier, providing RSA encryption/decryption, binary and integer polynomial functions, as well as elliptic curve cryptography (ECC)<sup>1</sup>
- The cipher/hash unit provides support for popular RC4, DES/3DES, AES-128/192/256, MD5, SHA-1, and SHA-256 ciphers

The SPU is designed to achieve wire-speed encryption and decryption on both of the processor's 10 GB Ethernet ports.

*1. Supported in a future Solaris OS release*

### **Integral PCI Express Support**

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

### **Power Management**

Beyond the inherent efficiencies of multi-core multi-thread design, the UltraSPARC T2 and UltraSPARC T2 Plus processors are the first processor to incorporate unique power management features at both the core and memory levels of the processor.

These features include reduced instruction rates, parking of idle threads and cores, and ability to turn off clocks in both cores and memory to reduce power consumption.

Substantial innovation is present in the areas of:

- Limiting speculation such as conditional branches not taken
- Extensive clock gating in the data path, control blocks, and arrays
- Power throttling that allows extra stall cycles to be injected into the decode stage

## Chapter 3

# Server Architecture

SPARC Enterprise T5120/T5220 and T5140/T5240 servers have been designed to provide breakthrough performance while maximizing reliability and minimizing power consumption and complexity. This section details the physical and architectural aspects of these systems.

## System-Level Architecture

A unified motherboard design is common to both the SPARC Enterprise T5120 and T5220 servers (Figure 9). The motherboard is a 20-layer printed circuit board (PCB) containing the UltraSPARC T2 processor, FB-DIMM sockets for main memory, ILOM service processor, disk controller, and I/O subsystems. I/O options include USB, DVD control, quad Gigabit Ethernet, and two levels of PLX PCI Express branching out into sockets for a wide variety of third-party PCI Express expansion options. Shaded regions indicate features that are only available on the SPARC Enterprise T5220 server.

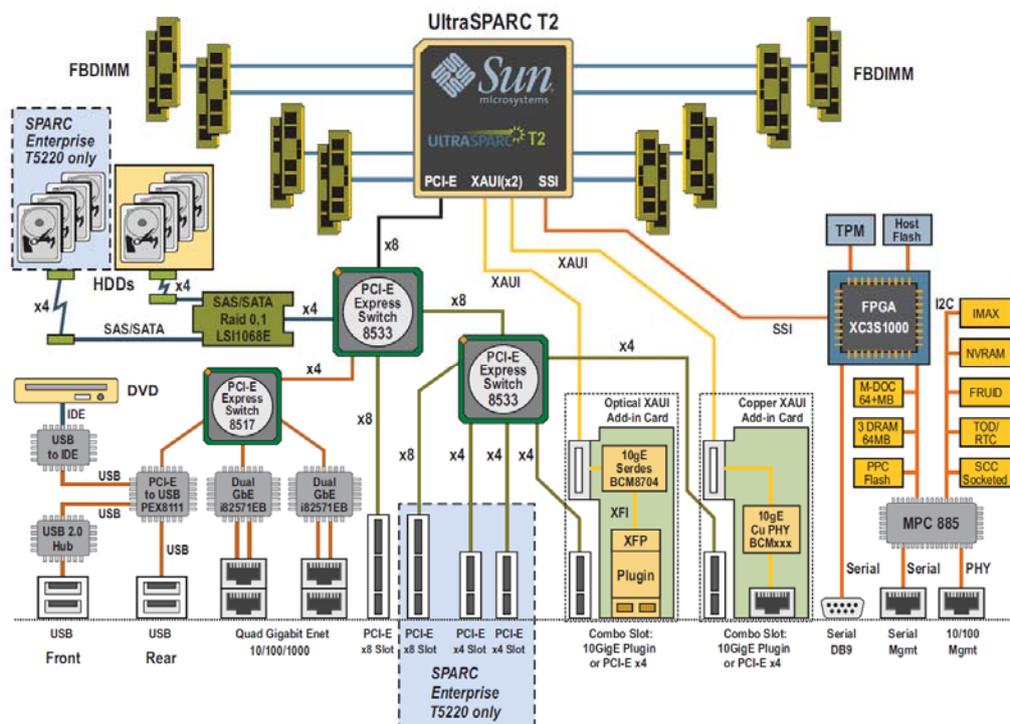


Figure 9. Block-level diagram of the common SPARC Enterprise T5120/T5220 server motherboard

SPARC Enterprise T5140 and T5240 servers share a common motherboard design (Figure 10)

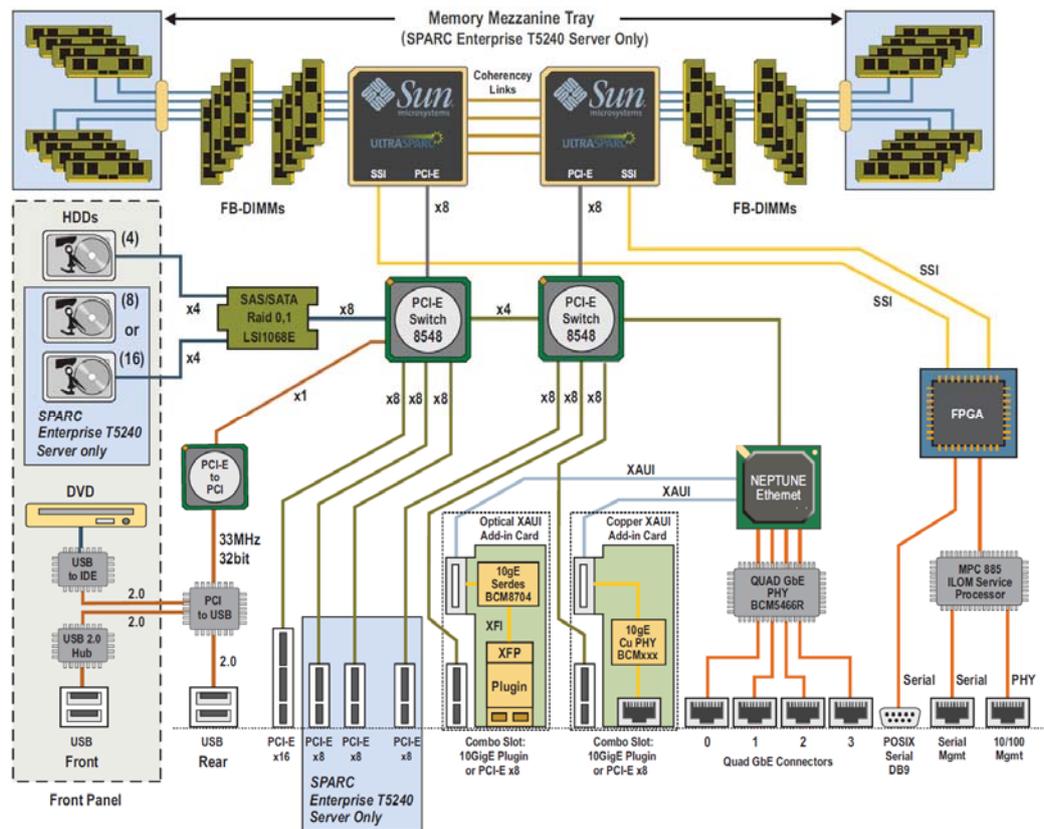


Figure 10. Block-level diagram of the common SPARC Enterprise T5140 and T5240 servers

Key features of the SPARC Enterprise T5140 and T5240 motherboards include:

- Dual sockets for UltraSPARC T2 Plus processors, connected by four coherency links
- A memory mezzanine tray to supply additional memory to the SPARC Enterprise T5240 server (up to 256 GB system maximum with 8 GB FB-DIMMs)
- Integration of Neptune chip to provide 10 gigabit Ethernet functionality as well as standard quad-gigabit Ethernet ports (10/100/1000-BaseT)

The motherboard interconnect for these systems has been greatly simplified over previous-generation systems. 12-volt power is distributed to the motherboard through a pair of metal bus bars, connected to a Power Distribution Board (PDB). A single flex-circuit connector routes all critical power control and DVD drive signaling over to the PDB. One or two mini-SAS cables connect the motherboard to the disk drive backplane, providing data access to the system hard drives.

## Memory Subsystem

In SPARC Enterprise T5120/T5220 and T5140/T5240 servers, the UltraSPARC T2 or UltraSPARC T2 Plus processor provides on-chip memory controllers that communicate directly to FB-DIMM memory through high-speed serial links. Four dual-channel FBDIMM memory controller units (MCUs) are provided on the UltraSPARC T2 processor while the UltraSPARC T2 Plus processor provides two MCUs. Each MCU can transfer data at an aggregate rate of 4.0 Gbps (UltraSPARC T2) or 4.8 Gbps (UltraSPARC T2 Plus).

Sixteen motherboard memory socket locations provide sufficient board space for two rows of 667 MHz FB-DIMMs per channel. SPARC Enterprise T5240 servers support an optional Memory Mezzanine Tray that can be added to support an additional 16 FBDIMM slots. The Memory Mezzanine Tray (Figure 11) allows the SPARC Enterprise T5240 server to support up to 256 GB of RAM using 8 GB FB-DIMMs.



*Figure 11. An optional Memory Mezzanine Tray doubles the memory capacity of SPARC Enterprise T5240 servers to up to 256 GB.*

## I/O Subsystem

Each UltraSPARC T2 and UltraSPARC T2 Plus processor incorporates a single, 8-lane (x8) PCI Express port capable of operating at 4 GB/second bidirectionally. In each server, this port natively interfaces to the I/O devices through a series of PLX technology PCI Express expander chips, connecting either to PCI Express card slots, or to bridge devices that interface with PCI Express, such as those listed below.

- Disk Controller — Disk control is managed by an LSI Logic SAS1068E SAS controller chip that interfaces to a four-lane (x4) PCI Express port. RAID levels 0 and 1 are provided as standard.
- Modular disk backplanes — Depending on the system, a four-, eight-, or sixteen-disk backplane is attached to the LSI disk controller by one or more x4 SAS links. The 16-disk backplane provides a 28-port LSI Logic SAS Expander to support the additional disk drives.
- Quad gigabit Ethernet — On SPARC Enterprise T5120 and T5220 servers, two x4 PCI Express ports connect to two Intel Ophir dual Gb Ethernet chips, providing four 10/100/1000 Mbps Ethernet interfaces on the rear of each chassis. On SPARC Enterprise T5140 and T5240 servers, Neptune Ethernet chip provides two 10/

100/1000-BaseT ports and two 10/100/1000/10000-BaseT interfaces, exposed as four RJ-45 connectors on the rear panel.

- **Dual 10 Gb Ethernet** — SPARC Enterprise T5120, T5220, T5140, and T5240 servers all provide dual 10Gb XAUI connections, expressed through shared XAUI/PCI Express slots. On SPARC Enterprise T5120 and T5220 servers, these ports are provided by the dual 10Gb Ethernet ports integrated into the UltraSPARC T2 processor. On SPARC Enterprise T5140 and T5240 servers, the 10 Gb Ethernet interfaces are provided by Neptune Ethernet chip. When the 10 Gb Ethernet ports are connected, two of the gigabit Ethernet ports become unavailable for use.
- **USB and DVD** — On all servers, a single-lane PCI Express port connects to a PCI bridge device. A second bridge chip converts the 32-bit 33MHz PCI bus into multiple USB 2.0 ports. The system's USB interconnect is driven from those ports. In addition, the DVD is driven from a further bridge chip that interfaces one of the USB ports to IDE format.

## Chassis Design Innovations

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

- **Enhanced System and Component Serviceability**

Finding and identifying servers and components in a modern datacenter can be challenge. T5120/T5220 and T5140/T5240 servers are optimized for lights-out datacenter configurations with easy to identify servers and modules. Color-coded operator panels provide easy-to-understand diagnostics and systems are designed for deployment in hot-isle / cold-isle multi-racked deployments with both front and rear diagnostic LEDs to pinpoint faulty components. Fault Remind features identify failed components.

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

- **Robust Chassis, Component, and Subassembly Design**

SPARC Enterprise T5120/T5220 and T5140/T5240 servers share chassis that are carefully designed to provide reliability and cool operation. Even features such as the hexagonal chassis ventilation holes are designed to provide the best compromise for high strength, maximum air flow, and maximum electronic attenuation.

Next-generation hard disk drive carriers leverage the hexagonal ventilation of the chassis and provide a seven-percent smaller front plate for greater storage density while increasing airflow to the system.

A removable disk cage in each system plugs directly in front of the fan tray assemblies, allowing airflow to be directed both above and below disk drives, and then above and below memory DIMMs and mezzanine boards to efficiently cool the system. Dual cooling fan modules are isolated from chassis to avoid transfer of rotational vibration to other system components. Also, integration of the fan power board into the Fan Tray assembly protects users from electrical shock during fan removal/insertion.

In spite of their computational, I/O, and storage density, Fujitsu's servers are able to maintain adequate cooling using conventional technologies. In fact, while the fan trays can hold a large number of fans, only the number of fans actually needed to cool the various systems are configured. Minimized DC-DC power conversions also contribute to overall system efficiency. By providing 12 volt power to the motherboard, power conversion stages are eliminated. This approach reduces generated heat, and introduces further efficiencies to the system.

• **Minimized Cabling for Maximized Airflow**

To minimize cabling and increase reliability, a variety of smaller boards and riser cards are employed, appropriate to each chassis. These infrastructure boards serve various functions in the SPARC Enterprise T5120/T5220 and T5140/T5240 servers.

- Power distribution boards distribute system power from the dual power supplies to the motherboard and to the disk backplane (via a connector board)
- Connector boards eliminate the need for many discrete cables, providing a direct card plug-in interconnect to distribute control and most data signals to the disk backplane, fan boards, and the PDB.
- Fan boards provide connections for power and control for both the primary and secondary fans in the front of the chassis. No cables are required since every dual fan module plugs directly into one of these PCBs which, in turn, plugs into the Connector Board.
- PCI Express riser cards plug directly into the motherboard, allowing PCI Express cards to be installed.
- Two XAUI riser cards provide slots that access to the on-chip 10 Gb Ethernet interfaces on the UltraSPARC T2 Processor or Neptune Ethernet chip. Alternately, these slots can provide access to PCI Express interfaces. Each slot can either accept an optical/copper XAUI card, or an industry standard low-profile PCI Express card with up to an x8 form factor edge connector. Cards are installed in a horizontal orientation.
- The disk backplane mounts to the disk cages in the two chassis, delivering disk data through one or two 4-channel discrete mini-SAS cables from the motherboard.

4, or 8-disk backplane is offered for the SPARC Enterprise T5120 and T5140 servers. SPARC Enterprise T5220 and T5240 server support either an 8-disk or 16-disk backplane.

- Also provided via the disk backplane, are two USB connections to the front of the system.

## SPARC Enterprise T5120 Server Overview

The compact SPARC Enterprise T5120 server provides significant computational power in a space-efficient low-power 1U rackmount package. With high levels of price/performance and a low acquisition cost together with tightly-integrated high-performance 10 gigabit Ethernet, this server is ideally suited to the delivery of horizontally-scaled transaction and Web services which require extreme network performance. The server is designed to address the challenges of modern datacenters with greatly reduced power consumption and a small physical footprint. Depending on the model selected, the SPARC Enterprise T5120 server features a single four-, or eight-core UltraSPARC T2 processor.

### Enclosure

The 1U SPARC Enterprise T5120 server enclosure is designed for use in a standard 19-inch rack (Table 3).

*Table 3. Dimensions and weight of the SPARC Enterprise T5120 server*

<b>Dimension</b>	<b>U.S.</b>	<b>International</b>
Height	1.746 inches (1 RU)	44 millimeters
Width	16.75 inches	425 millimeters
Depth	28.125 inches	714 millimeters
Weight (approximate, without PCI cards or rackmounts)	40 pounds	18 kilograms

The SPARC Enterprise T5120 server includes the following major components:

- An UltraSPARC T2 processor with four, or eight cores at speeds of 1.2, 1.4, or 1.6 GHz
- Up to 128 GB of memory in 16 Fully Buffered Dual Inline Memory Module (FB-DIMM) slots (2 GB, and 4 GB, and 8 GB FB-DIMMs supported)
- Four on-board 10/100/1000 Mbps Ethernet ports
- Dedicated low-profile PCI Express slot (x8)
- Two combination XAUI or low-profile PCI Express x4 slots
- Four USB 2.0 ports (2 forward, 2 rear facing)
- Four or eight available disk drives supporting SAS disk drives (note The eight disk backplane is not supported with 1.6GHz CPUs)
- Integrated Lights out Management (ILOM) system controller
- Two (N+1) hot-swappable high-efficiency 720 watts AC or 660 watts DC power

supply units

- Four fan assemblies (each with two fans) populated of a possible eight, under environmental monitoring and control, N+1 redundancy. Fans are accessed through a dedicated top panel door.

## Front and Rear Perspectives

Figure 12 illustrates the front and rear panels of the SPARC Enterprise T5120 server.

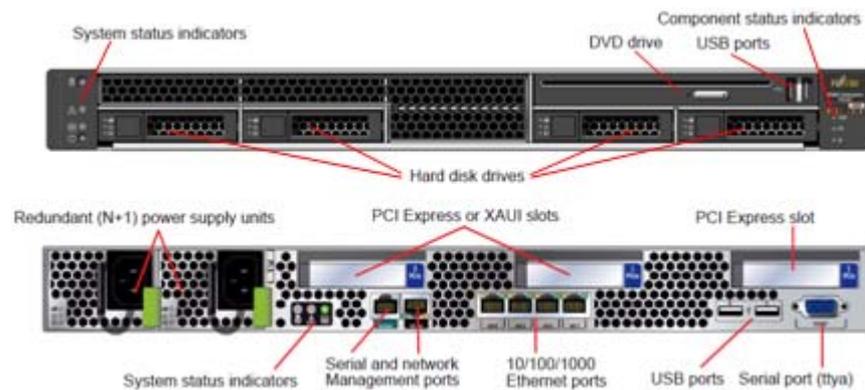


Figure 12. SPARC Enterprise T5120 server, front and rear panels

External features of the SPARC Enterprise T5120 server include:

- Front and rear system and component status indicator lights provide locator (white), service required (amber), and activity status (green) for the system.
- Four or eight hot-plug SAS disk drives insert through the front panel of the system.
- One DVD+/-RW is accessed through the front panel.
- Four USB 2.0 ports are provided, two on the front panel, and two on the rear.
- Two hot-plug/hot-swap (N+1) power supplies with integral fans insert from the rear.
- Rear power-supply indicator lights convey the status of each power supply.
- A single AC plug is provided on each hot-plug/hot-swap power supply.
- Four 10/100/1000Base-T autosensing Ethernet ports are provided.
- A DB-9 TTYA serial port is provided for serial devices (not connected to the ILOM system controller serial port).
- A total of three PCI Express card slots are provided, two of which can alternately support XAUI cards connected to the UltraSPARC T2 10 Gb Ethernet interfaces.
- Two management ports are provided for use with the ILOM system controller. The RJ-45 serial management port provides the default connection to the ILOM controller.

## SPARC Enterprise T5220 Server Overview

The expandable SPARC Enterprise T5220 server is optimized to deliver transaction and Web services, including Java 2 Platform, Enterprise Edition (J2EE<sup>®</sup> platform)

technology application services, enterprise application services (ERP, CRM, and SCM), and distributed databases. With considerable expansion capabilities and integrated virtualization technologies, the SPARC Enterprise T5220 server is also an ideal platform for consolidated Tier-1 and Tier-2 workloads.

## Enclosure

The SPARC Enterprise T5220 server features a compact, yet expandable 2U rackmount chassis (Table 4), giving organizations the flexibility to scale their processing and I/O needs without wasting precious space.

*Table 4. Dimensions and weight of the SPARC Enterprise T5220 server*

<b>Server/Dimension</b>	<b>U.S.</b>	<b>International</b>
Height	3.49 inches (2 RU)	88 millimeters
Width	16.75 inches	425 millimeters
Depth	28.125 inches	714 millimeters
Weight (without PCI cards or rack mounts)	55 pounds	25 kilograms

The SPARC Enterprise T5220 server includes the following major components:

- An UltraSPARC T2 processor with four, or eight cores
- Up to 128 GB of memory in 16 Fully Buffered Dual Inline Memory Module (FB-DIMM) slots (2 GB, 4 GB, and 8GB FB-DIMMs supported)
- Four on-board 10/100/1000 Mbps Ethernet ports
- Four dedicated low-profile PCI Express slots
- Two combination XAUI or low-profile PCI Express x4 slots
- Four USB 2.0 ports (2 forward, 2 rear facing)
- Up to eight, or sixteen available disk drives supporting SAS disk drives
- Integrated Lights out Management (ILOM) system controller
- Two (N+1) hot-plug/hot-swap high-efficiency 750 or 1100 AC Watts power supplies or 1200 Watts DC power supplies
- Three fan assemblies (each with two fans) populated of a possible six, under environmental monitoring and control, N+1 redundancy

## Front and Rear Perspectives

Figure 13 illustrates the front and back panels of the SPARC Enterprise T5220 server.

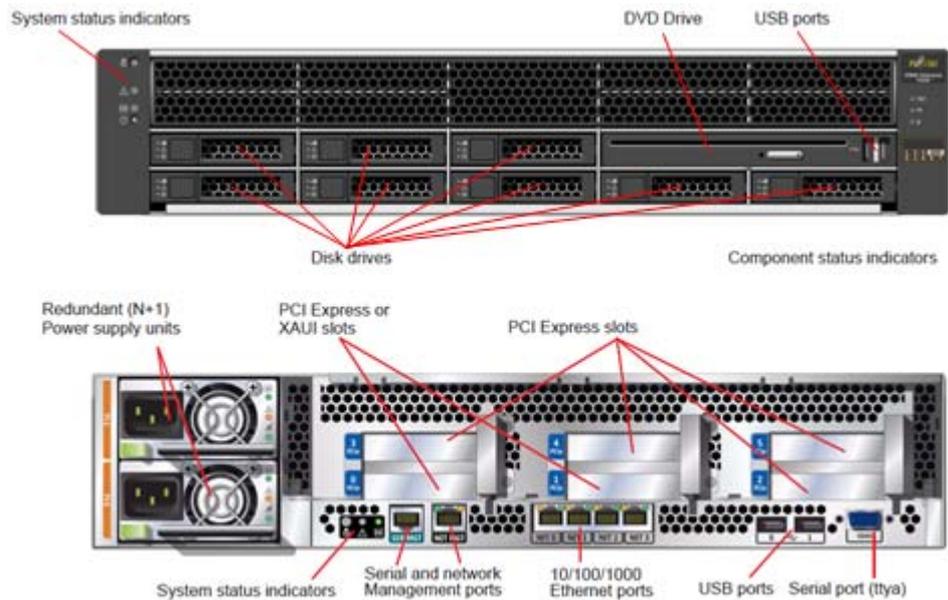


Figure 13. SPARC Enterprise T5220 server, front and rear panels

External features of the SPARC Enterprise T5220 server include:

- Front and rear system and component status indicator lights provide locator (white), service required (amber), and activity status (green) for the system
- Hot-plug SAS disk drives insert through the front panel of the system
- One DVD+/-RW drive is accessed through the front panel
- Four USB 2.0 ports are provided, two on the front panel, and two on the rear
- Two hot-plug/hot-swap N+1 power supplies with integral plugs and fans insert from the rear (rear power-supply indicator lights convey the status of each power supply)
- Four 10/100/1000Base-T autosensing Ethernet ports are provided
- A DB-9 TTYA serial port is provided for serial devices (not connect to the ILOM system controller serial port)
- A total of six PCI Express card slots are provided, two of which can support XAUI cards connected to the UltraSPARC T2 10 Gb Ethernet interfaces
- Two management ports are provided for use with the ILOM system controller. The RJ-45 serial management port provides the default connection to the ILOM controller. The network management port supports an optional RJ-45 10/100Base-T connection to the ILOM system controller.

## SPARC Enterprise T5140 Server Overview

With support for up to two UltraSPARC T2 Plus processors, and up to 128 threads, the compact SPARC Enterprise T5140 server provides breakthrough computational power in a space-efficient low-power 1U rackmount package. With high levels of price/performance and a low acquisition cost, this server is ideally suited to the

delivery of horizontally-scaled transaction and Web services, and presents many opportunities as a consolidation and virtualization server. The server is designed to address the challenges of modern datacenters with greatly reduced power consumption and a small physical footprint. Depending on the model selected, the SPARC Enterprise T5140 server features dual four-, six-, or eight-core UltraSPARC T2 Plus processors.

## Enclosure

The 1U SPARC Enterprise T5140 server enclosure is designed for use in a standard 19-inch rack (Table 5).

Table 5. Dimensions and weight of the SPARC Enterprise T5140 server

<b>Server/Dimension</b>	<b>U.S.</b>	<b>International</b>
Height	1.746 inches (1 RU)	44 millimeters
Width	16.75 inches	425 millimeters
Depth	28.125 inches	714 millimeters
Weight (without PCI cards or rack mounts)	42 pounds	19 kilograms

The SPARC Enterprise T5140 server includes the following major components:

- Two UltraSPARC T2 Plus processors with four, six, or eight cores at 1.2 GHz
- Up to 128 GB of memory in 16 Fully Buffered Dual Inline Memory Module (FB-DIMM) slots (2 GB, 4 GB, and 8 GB FB-DIMMs currently supported)
- Four on-board 10/100/1000 Mbps Ethernet ports
- Dedicated low-profile PCI Express slot (x8 with x16 connector)
- Two combination XAUI or low-profile PCI Express x8 slots
- Four USB 2.0 ports (2 forward, 2 rear facing)
- Four, eight available disk drives slots supporting SAS commodity disk drives
- Integrated Lights out Management (ILOM) system controller
- Two (N+1) hot-swappable high-efficiency 720 watts AC or 660 watts power supply units
- Six fan assemblies (each with two fans), under environmental monitoring and control, N+1 redundancy. Fans are accessed through a dedicated top panel door.

## Front and Rear Perspectives

Figure 14 illustrates the front and rear panels of the SPARC Enterprise T5140 server.

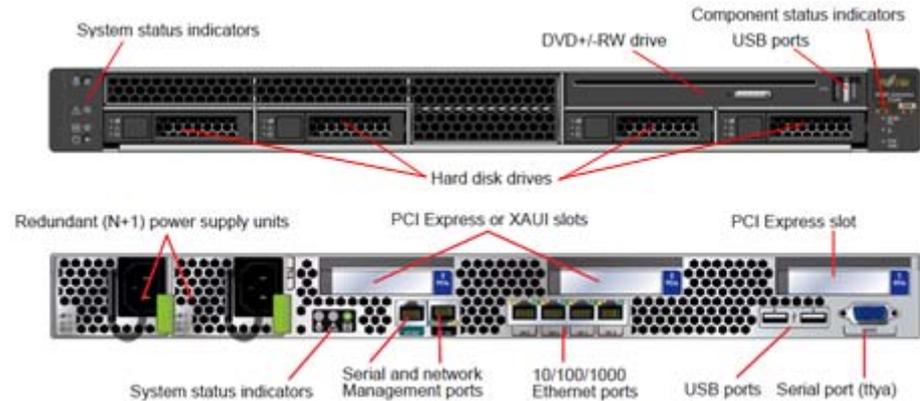


Figure 14. SPARC Enterprise T5140 server, front and rear panels

External features of the SPARC Enterprise T5140 server include:

- Front and rear system and component status indicator lights provide locator (white), service required (amber), and activity status (green) for the system.
- Hot-plug SAS disk drives insert through the front panel of the system.
- One slimline, slot-accessible DVD+/-RW is accessed through the front panel.
- Four USB 2.0 ports are provided, two on the front panel, and two on the rear.
- Two hot-plug/hot-swap (N+1) power supplies with integral fans insert from the rear.
- Rear power-supply indicator lights convey the status of each power supply.
- A single AC plug is provided on each hot-plug/hot-swap power supply.
- Four 10/100/1000Base-T autosensing Ethernet ports are provided.
- A DB-9 TTYA serial port is provided for serial devices (not connected to the ILOM system controller serial port).
- A total of three PCI Express card slots are provided, two of which can alternately support XAUI cards connected to the Neptune Ethernet chip XAUI interfaces.
- Two management ports are provided for use with the ILOM system controller. The RJ-45 serial management port provides the default connection to the ILOM controller.

## SPARC Enterprise T5240 Server Overview

The expandable SPARC Enterprise T5240 server is optimized to deliver highlyscalable transaction and Web services, including Java 2 Platform, Enterprise Edition (J2EE™ platform) technology application services, enterprise application services (ERP, CRM, and SCM), and distributed databases. With support for two UltraSPARC T2 Plus processors, considerable expansion capabilities, and integrated virtualization technologies, the SPARC Enterprise T5240 server is also an ideal platform for consolidated Tier-1 and Tier-2 workloads.

## Enclosure

The SPARC Enterprise T5240 server features a compact, yet expandable 2U rackmount chassis (Table 6), giving organizations the flexibility to scale their processing and I/O needs without wasting precious space.

Table 6. Dimensions and weight of the SPARC Enterprise T5240 server

<b>Server/Dimension</b>	<b>U.S.</b>	<b>International</b>
Height	3.49 inches (2 RU)	88 millimeters
Width	16.75 inches	425 millimeters
Depth	28.125 inches	714 millimeters
Weight (without PCI cards or rack mounts)	57 pounds	26 kilograms

The SPARC Enterprise T5240 server includes the following major components:

- Two UltraSPARC T2 Plus processors with four, six, or eight cores at 1.2 GHz, or eight cores at 1.4 GHz, eight core at 1.6 GHz
- Up to 256 GB of memory in 16 or 32 Fully Buffered Dual Inline Memory Module (FBDIMM) slots (2 GB, 4 GB, and 8GB FB-DIMMs supported) with optional Memory Mezzanine Tray
- Four on-board 10/100/1000 Mbps Ethernet ports
- Four dedicated low-profile x8 PCI Express slots (all x8 electrically with one x16 physical connector)
- Two combination XAUI or low-profile PCI Express x8 slots
- Four USB 2.0 ports (2 forward, 2 rear facing)
- Up to eight or sixteen available disk drive slots supporting commodity SAS disk drives  
(note The sixteen disk backplane is not supported with 1.6GHz CPUs)
- Integrated Lights out Management (ILOM) system controller
- Two (N+1) hot-plug/hot-swap high-efficiency 1,100 watt AC or 1,200 DC power supplies
- Five fan assemblies (each with two fans) under environmental monitoring and control, N+1 redundancy

## Front and Rear Perspectives

Figure 15 illustrates the front and back panels of the SPARC Enterprise T5240 server.

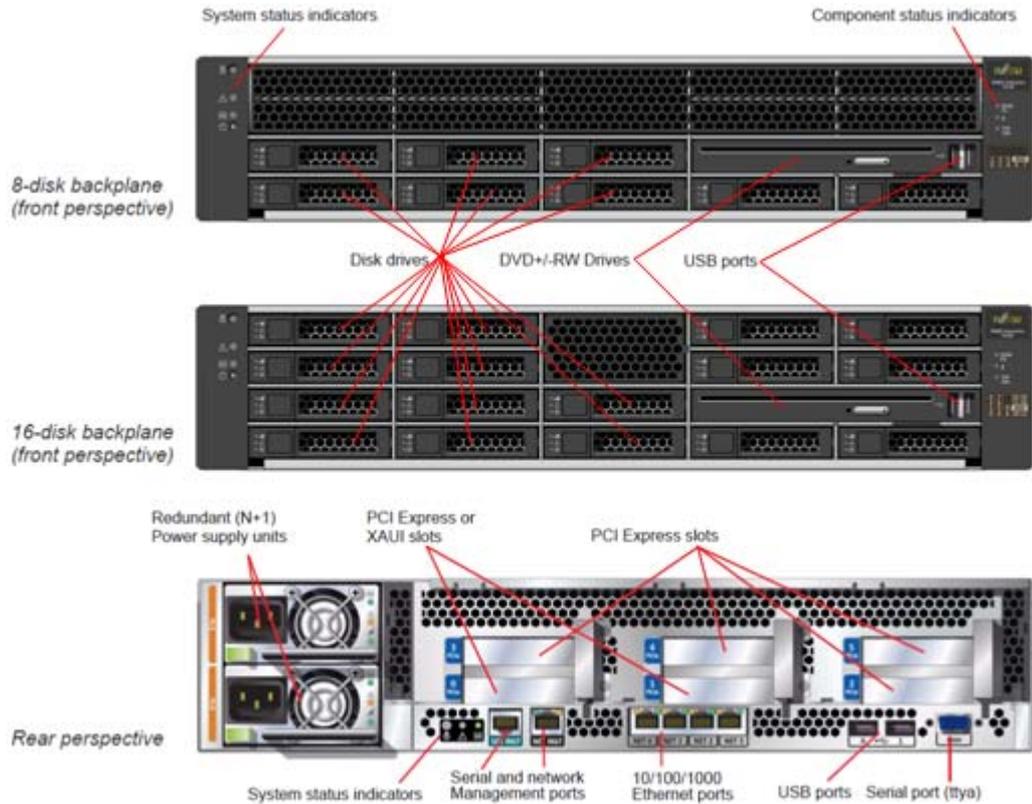


Figure 15. SPARC Enterprise T5240 server, front and rear panels

External features of the SPARC Enterprise T5240 server include:

- Front and rear system and component status indicator lights provide locator (white), service required (amber), and activity status (green) for the system
- Eight or 16 hot-plug SAS disk drives insert through the front panel of the system
- One slimline DVD+/-RW drive is accessed through the front panel
- Four USB 2.0 ports are provided, two on the front panel, and two on the rear
- Two hot-plug/hot-swap N+1 power supplies with integral plugs and fans insert from the rear (rear power-supply indicator lights convey the status of each power supply)
- Four 10/100/1000Base-T autosensing Ethernet ports are provided
- A DB-9 TTYA serial port is provided for serial devices (not connected through the ILOM system controller serial port)
- A total of six PCI Express card slots are provided, two of which can support XAUI cards connected to the Neptune Ethernet chip XAUI interfaces
- Two management ports are provided for use with the ILOM system controller. The RJ-45 serial management port provides the default connection to the ILOM controller. The network management port supports an optional RJ-45 10/100Base-T connection to the ILOM system controller.

## PCI Express Expansion Unit

With the strength of UltraSPARC T2 and T2 Plus processors, Fujitsu SPARC Enterprise T5120/T5220 and T5140/T5240 servers are ideally suited for mission-critical applications and databases that often require fast access to considerable near-line storage. These systems also have considerable I/O capacity and bandwidth available for external expansion. To scale I/O expansion beyond the constraints of the individual system chassis, Fujitsu SPARC Enterprise T5120/T5220 and T5140/T5240 servers support the attachment of an optional Sun External I/O Expansion Unit to provide additional I/O connectivity.

The External I/O Expansion Unit is a 4U rack-mountable device which accommodates up to two 12 additional PCI Express slots – connected to, and managed by, the Fujitsu SPARC Enterprise server. By using cassettes, the external I/O chassis supports active replacement of hot-plug cards. An I/O Link card mounted in the host provides connectivity to each of two “I/O boats” the External I/O Expansion Unit and facilitates host management control via sideband signals. The I/O link card is available as a low height copper card, and includes a single 8-lane PCI Express bus with 4 GB/sec bandwidth. The architecture of the External Expansion unit provides high-throughput I/O performance, supporting maximum data rates for many types of PCI Express cards and burst traffic from additional PCI Express cards.

Front and rear perspectives of the each External I/O Expansion Unit contains either one or two I/O boats, with each boat providing six external x8 PCI Express slots. Individual I/O boats connect to the Fujitsu SPARC Enterprise T5120/T5220 or T5140/T5240 server via a link card that is installed in one of the system’s PCI Express slots. As a result, a fully-configured External I/O Expansion Unit requires two PCI Express slots in the server chassis and provides an addition 10 PCI Express slots. The Fujitsu External I/O Expansion Unit includes several key technologies, including:

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

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

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

## Chapter 4

# Enterprise-Class Management and Software

While new technology often requires time for tools and applications to arrive, delivering agile and highly-available services that take advantage of available resources requires stable development tools, operating systems, middleware and management software. Fortunately, in spite of the breakthrough UltraSPARC T2 and UltraSPARC T2 Plus processor technology, SPARC Enterprise T5120/T5220 and T5140/T5140 servers provide full binary compatibility with earlier SPARC systems and are delivered ready to run with pre-loaded tools and the solid foundation of the Solaris OS. Moreover, these systems are provided with a wealth of sophisticated tools that let organizations develop and tune applications as they consolidate and manage workloads while effectively utilizing the resources of UltraSPARC T2 and UltraSPARC T2 Plus processors.

## System Management Technology

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

### Integrated Lights-Out Management (ILOM) System Controller

Provided across SPARC Enterprise T1000 and T2000, the Integrated Lights Out Management (ILOM) service processor acts as a system controller, facilitating remote management and administration of SPARC Enterprise T5120/T5220 and T5140/T5240 servers. As a result, these servers integrate easily with existing management infrastructure.

Critical to effective system management, the ILOM service processor:

- Implements an IPMI 2.0 compliant services processor, providing IPMI management functions to the server's firmware, OS and applications, and to IPMI-based management tools accessing the service processor via the ILOM Ethernet management interface, providing visibility to the environmental sensors (both on the server module, and elsewhere in the chassis)
- Manages inventory and environmental controls for the server, including CPUs, DIMMs, and power supplies, and provides HTTPS/CLI/SNMP access to this data
- Supplies remote textual console interfaces
- Provides a means to download upgrades to all system firmware

The ILOM service processor also allows the administrator to remotely manage the server, independent of the operating system running on the platform and without interfering with any system activity. ILOM can also send e-mail alerts of hardware failures and warnings, as well as other events related to each server. The ILOM circuitry runs independently from the server, using the server's standby power. As a result, ILOM firmware and software continue to function when the server operating system goes offline, or when the server is powered off. ILOM monitors the following SPARC Enterprise T5120/T5220 and T5140/T5240 server conditions:

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

## **Server Management Software**

Monitoring and control software is essential in managing today's server infrastructure and the complications of IT device distribution. Fujitsu's Server System Manager (SSM) enables management of all Solaris, Linux, and Windows servers. Plus, in combination with Systemwalker Fujitsu's integrated systems management software, SSM provides autonomic operation and ensures you maintain business continuity as your IT environment changes to match your business.

### **Server System Manager (SSM)**

SMM software provides a common server management environment for PRIMEQUEST mission-critical open servers, industry standard PRIMERGY servers, and SPARC Enterprise UNIX servers. Graphical server management views and functions for hardware configuration and hardware monitoring make SSM the easy option in a heterogeneous server environment.

Not only are you able to monitor the status of multiple servers you can also control server power from the same single consistent display. Ease of use continues as all server configurations are viewed using a single tree structure and common formats. This lets you check server component status a glance. Racked servers and related items are shown relative to their rack positions using life-like server images. This further simplifies status monitoring of server resources such as CPU and memory. Importantly, the networked nature of SSM means you can perform power on/off and other operations from any location with a management display.

With SSM server operation management is greatly simplified, letting you concentrate on hardware problem resolution and other critical work. As a result your work and the

work of other administrators will be both reduced and more productive.

### **Enhanced Support Facility**

Enhanced Support Facility is specific software that improves the operation management and the maintainability of SPARC Enterprise servers. Working in combination with ILOM, server configuration, status and error messages can all be displayed. If a problem occurs, the information reported to ILOM ensures the status, of disks, power, PCI cards and OS, is always monitored. It also enables you to display other system information including batch collections, /etc/system file settings, server power on/off scheduling and disk hot swap procedures.

### **Systemwalker Centric Manager**

Centric Manager lets you follow the system operation lifecycle (installation/setup, monitoring, fault recovery, assessment), making it possible for you to create highly-reliable systems. It reduces the workload required for operations management and provides high-value functions for life-cycle tasks. These include the remote distribution of software resources, central monitoring of systems and networks, and prompt resolution of problems from any location. It performs integrated management, operational process standardization (ITIL), while enabling security control of the latest business IT technology such as multi-platform and intranet/Internet environments

## **Scalability and Support for Innovative Multithreading Technology**

The Solaris 10 Operating System is specifically designed to deliver the considerable resources of UltraSPARC T2 and UltraSPARC T2 Plus processor based systems. In fact, the Solaris 10 OS provides key functionality for virtualization, optimal utilization, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. The Solaris 10 OS runs on a broad range of SPARC systems and compatibility with existing applications is guaranteed.

One of the most attractive features of systems based on the UltraSPARC T2 and UltraSPARC T2 Plus processor is that they appear as a familiar SMP system to the Solaris OS and the applications it supports. In addition, the Solaris 10 OS has incorporated many features to improve application performance on multithreading architectures:

- ***System Awareness***

The Solaris 10 OS is aware of the UltraSPARC T2 and UltraSPARC T2 Plus processor hierarchy so that the scheduler can effectively balance the load across all the available pipelines. Even though it exposes each of these processors as 64 logical processors, the Solaris OS understands the correlation between cores and the threads they support, and provides a fast and efficient thread implementation.

- ***Fine-Granularity Manageability***

For the UltraSPARC T2 and UltraSPARC T2 Plus processor, the Solaris 10 OS has the ability to enable or disable individual cores and threads (logical processors). In addition, standard Solaris OS features such as processor sets provide the ability to define a group of logical processors and schedule processes or threads on them.

- ***Binding Interfaces***

The Solaris OS allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set, if required or desired.

- ***Support for Virtualized Networking and I/O, and Accelerated Cryptography***

The Solaris OS contains technology to support and virtualize components and subsystems on the UltraSPARC T2 processor, including support for the on-chip 10 Gb Ethernet ports and PCI Express interface. As a part of a high-performance network architecture, system-aware device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices. Accelerated cryptography is supported through the Solaris Cryptographic framework.

### ***NUMA Optimization in the Solaris OS***

With memory managed by each UltraSPARC T2 Plus processor on SPARC Enterprise T5140 and T5240 servers, the implementation represents a non-uniform memory access (NUMA) architecture. In NUMA architectures, the speed needed for a processor to access its own memory is slightly different than that required to access memory managed by another processor. The Solaris OS provides technology that can specifically help applications improve performance on NUMA architectures.

- Memory Placement Optimization (MPO) — The Solaris 10 OS uses MPO to improve the placement of memory across the physical memory of a server, resulting in increased performance. Through MPO, the Solaris 10 OS works to help ensure that memory is as close as possible to the processors that access it, while still maintaining enough balance within the system. As a result, many database applications are able to run considerably faster with MPO.
- Hierarchical lgroup support (HLS) — HLS improves the MPO feature in the Solaris OS. HLS helps the Solaris OS optimize performance for systems with more complex memory latency hierarchies. HLS lets the Solaris OS distinguish between the degrees of memory remoteness, allocating resources with the lowest possible latency for applications. If local resources are not available by default for a given application, HLS helps the Solaris OS allocate the nearest remote resources.

- ***Solaris ZFS File System***

Solaris ZFS offers a dramatic advance in data management, automating and

consolidating complicated storage administration concepts and providing unlimited scalability with the world's first 128-bit file system. Solaris ZFS is based on a transactional object model that removes most of the traditional constraints on I/O issue order, resulting in dramatic performance gains. Solaris ZFS also provides data integrity, protecting all data with 64-bit checksums that detect and correct silent data corruption.

- ***A Secure and Robust Enterprise-Class Environment***

Best of all, the Solaris OS doesn't require arbitrary sacrifices. The Solaris Binary Compatibility Guarantee helps ensure that existing SPARC applications continue to run unchanged on UltraSPARC T2 and UltraSPARC T2 Plus platforms, protecting investments. Certified multi-level security protects Solaris environments from intrusion. Comprehensive Fault Management Architecture means that elements such as Solaris Predictive Self Healing can communicate directly with the hardware to help reduce both planned and unplanned downtime. Effective tools such as DTrace help organizations tune their applications to get the most of the system's resources.

## **End-to-End Virtualization Technology**

Virtualization technology is increasingly popular as organizations strive to consolidate disparate workloads onto fewer more powerful systems, while increasing utilization. SPARC Enterprise T5120/T5220 and T5140/T5240 servers are specifically designed for virtualization, providing very fine-grained division of multiple resources — from processing to virtualized networking and I/O. Most importantly, virtualization technology is provided as a part of the system, not an expensive add-on.

## **A Multithreaded Hypervisor**

Like the UltraSPARC T1 processor, the UltraSPARC T2 and UltraSPARC T2 Plus processors offer a multithreaded hypervisor — a small firmware layer that provides a stable virtual machine architecture that is tightly integrated with the processor. Multi-core Multi-thread is crucial, since the hypervisor interacts directly with the underlying chip-multithreaded UltraSPARC T2 processor. This architecture is able to context switch between multiple threads in a single core, a task that would require additional software and considerable overhead in competing architecture.

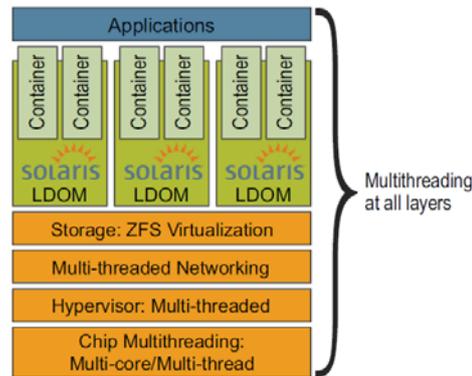


Figure 16. Fujitsu provides parallelization and virtualization at every level of the technology stack

Corresponding layers of virtualization technology are built on top of the hypervisor as shown in Figure 16. The strength of Fujitsu’s approach is that all of the layers of the architecture are fully multithreaded, from the processor up through applications that use the fully threaded Java application model. Far from new technology, the Solaris OS has provided multi-core multi-thread support since 1992. This experience has helped to inform technology decisions at other levels, ultimately resulting in a system that parallelizes and virtualizes at every level. In addition to the processor and hypervisor, we provides fully multithreaded networking and the fully multithreaded Solaris ZFS file system. Logical Domains (LDMs), Solaris Containers, and multithreaded applications are able to receive exactly the resources they need.

## Logical Domains

Supported in all servers utilizing multi-core multi-thread technology, Logical Domains provide full virtual machines that run an independent operating system instance, and contain virtualized CPU, memory, storage, console, and cryptographic devices. Within the Logical Domains architecture, operating systems such as the Solaris 10 OS are written to the hypervisor, which provides a stable, idealized, and virtualizable representation of the underlying server hardware to the operating system in each Logical Domain. Each Logical Domain is completely isolated, and the maximum number of virtual machines created on a single platform relies upon the capabilities of the hypervisor, rather than the number of physical hardware devices installed in the system. For example, the SPARC Enterprise T5220 server with a single UltraSPARC T2 processor supports up to 64 logical domains<sup>1</sup>, and each individual logical domain can run a unique OS instance.

*1. Though possible, this practice is not a generally recommended.*

By taking advantage of Logical Domains, organizations gain the flexibility to deploy multiple operating systems simultaneously on a single platform. In addition, administrators can leverage virtual device capabilities to transport an entire software stack hosted on a Logical Domain from one physical machine to another. Logical

Domains can also host Solaris Containers to capture the isolation, flexibility, and manageability features of both technologies. Deeply integrating Logical Domains with both the UltraSPARC T2 and UltraSPARC T2 Plus processors and the Solaris 10 OS increases flexibility, isolates workload processing, and improves the potential for maximum server utilization.

The Logical Domains architecture includes underlying server hardware, hypervisor firmware, virtualized devices, and guest, control, and service domains. The hypervisor firmware provides an interface between each hosted operating system and the server hardware. An operating system instance controlled and supported by the hypervisor is called a *guest domain*. Communication to the hypervisor, hardware platform, and other domains for creation and control of guest domains is handled by the *control domain*. Guest domains are granted virtual device access via a *service domain* which controls both the system and hypervisor, and also assigns I/O.

To support virtualized networking, Logical Domains implement a virtual Layer 2 switch, to which guest domains can be connected. Each guest domain can be connected to multiple vswitches and multiple guest domains can also be connected to the same vswitch. Vswitches can either be associated with a real physical network port, or they may exist without an associated port, in which case the vswitch provides only communications between domains within the same server. This approach also gives guest domains a direct communication channel to the network (Figure 17). Each guest domain believes it owns the entire NIC and the bandwidth it provides, yet in practice only a portion of the total bandwidth is allotted to the domain. As a result, every NIC can be configured as demand dictates, with each domain receiving bandwidth on an as-needed basis. Dedicated bandwidth can be made available by tying a vswitch device to a dedicated physical Ethernet port.

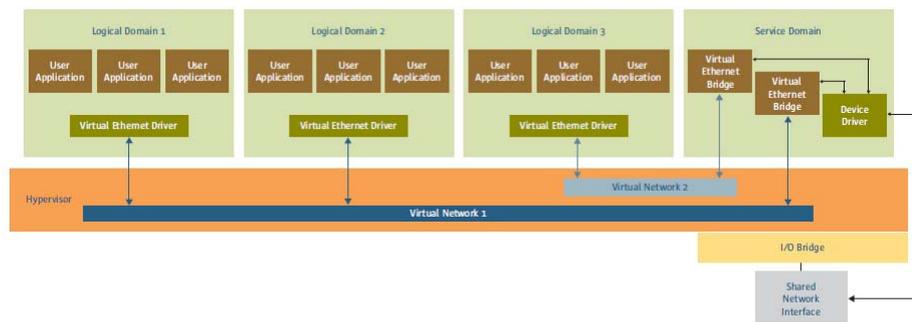


Figure 17. Data moves directly between a Logical Domain and a virtualized device

## Solaris™ Containers

Providing virtualization at the OS level, Solaris Containers consist of a group of technologies that work together to efficiently manage system resources, virtualize the environment, and provide a complete, isolated, and secure runtime environment for applications. Solaris containers include important technologies that work together with the fair-share scheduler:

- ***Solaris Zones***

The Solaris 10 OS provides a unique partitioning technology called Solaris Zones that can be used to create an isolated and secure environment for running applications. A zone is a virtualized operating system environment created within a single instance of the Solaris OS. Zones can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability since processes in one zone are prevented from interfering with processes running in another zone.

- ***Resource Management***

Resource management tools provided with the Solaris OS help allocate resources such as CPUs to specific applications. CPUs in a multiprocessor system (or threads in the UltraSPARC T2 and UltraSPARC T2 Plus processors) can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to a Solaris zone. Resource pools provide the capability to separate workloads so that consumption of CPU resources do not overlap, and also provide a persistent configuration mechanism for processor sets and scheduling class assignment. In addition, the dynamic features of resource pools enable administrators to adjust system resources in response to changing workload demands.

## **Fault Management and Predictive Self Healing**

The Solaris 10 OS introduced a new architecture for building and deploying systems and services capable of fault management and predictive self-healing. Predictive Self Healing is an innovative capability in the Solaris 10 OS that automatically diagnoses, isolates, and recovers from many hardware and application faults. As a result, business critical applications and essential system services can continue uninterrupted in the event of software failures, major hardware component failures, and even software misconfiguration problems.

- ***Solaris Fault Manager***

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

- ***Solaris Service Manager***

The Solaris Service Manager facility creates a standardized control mechanism for

application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services can then be automatically restarted if they are accidentally terminated by an administrator, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem. In addition, the Solaris Service Manager software reduces system boot time by as much as 75 percent by starting services in parallel according to their dependencies. An “undo” feature helps safeguard against human errors by permitting easy change rollback. The Solaris Service Manager is also simple to deploy; developers can convert most existing applications to take full advantage of Solaris Service Manager features by simply adding a simple XML file to each application.

Predictive self healing and fault management provide the following specific capabilities on SPARC Enterprise T5120/T5220 servers:

- *CPU Offlining* takes a core or threads offline that has been deemed faulty. Offlined CPUs are stored in the resource cache and stay offline on reboot unless the processor has been replaced, in which case the CPU is cleared from the resource cache.
- *Memory Page Retirement* retires pages of memory that have been marked as faulty. Pages are stored in the resource cache and stay retired on reboot unless the offending DIMM has been replaced, in which case affected pages are cleared from the resource cache.
- *I/O Retirement* logs errors and faults.
- *fmlog* logs faults detected by the system.

## Chapter 5

# Conclusion

Delivering on the demands of Web 2.0 applications and virtualized, eco-efficient data centers requires a comprehensive approach that includes innovative processors, system platforms, and operating systems, along with leading application, middleware, and management technology. With its strong technology positions and R&D investments in all of these areas, Fujitsu is in a unique position to deliver on this vision. Far from futuristic, Fujitsu has effective solutions today that can help organizations cope with the need for performance and capacity while effectively managing space, power and heat.

Building on the successful UltraSPARC T1 processor, the UltraSPARC T2 and UltraSPARC T2 Plus processors deliver approximately twice the throughput and efficiency, and serves as the industry's first massively-threaded system on a chip. With 64 threads per processor, on-chip memory management, two 10 Gb Ethernet interfaces, PCI express, and on-chip cryptographic acceleration, the UltraSPARC T2 processor fundamentally redefine the capabilities of a modern processor. By incorporating cache coherency for multiprocessor support, UltraSPARC T2 Plus processors allow these capabilities to be multiplied incrementally. SPARC Enterprise T5120/T5220 and T5140/T5240 servers leverage these strengths to provide powerful and highly-scalable server platforms while delivering new levels of performance and performance-per-watt in a compact rackmount chassis. The result is datacenter infrastructure that can truly scale to meet new challenges with a very small footprint.

SPARC Enterprise T5120/T5220 and T5140/T5240 servers provide the computational, networking, and I/O resources needed by the most demanding Web, application, and database applications, and they facilitate highly-effective consolidation efforts. With end-to-end support for multi-core multi-thread technology and virtualization, these systems can consolidate workloads and effectively utilize system resources even as they preserve investments in SPARC/Solaris technology and provide tools for open-source software environments. With innovations such as Logical Domains, Solaris Containers, and Java technology, organizations can adopt these radical new systems for their most important projects —acting responsibly toward the environment and the bottom line.

## For More Information

To learn more about products and the benefits of SPARC Enterprise T5120/T5220 and T5140/T5240 servers, contact a Fujitsu sales representative, or consult the related documents and Web site.

<http://www.fujitsu.com/sparcenterprise/>

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company Limited.

SPARC is a trademark used under license from SPARC International, Inc. of the U.S.A.

Sun, Sun Microsystems, the Sun Logo, Solaris and all Solaris based marks and logos are trademarks or registered trade marks of Sun Microsystems, Inc, in the U.S. and other countries, and are used under license.

SPARC64 is a trademark of Fujitsu Limited under license from SPARC International, Inc. of the U.S.A.

All other company/product names mentioned may be trademarks or registered trademarks of their respective holders and are used for identification purposes only.

**FUJITSU**