ORACLE

SUN SERVER

An Oracle White Paper
June 2011; Revised July 2012

Oracle's Sun Server X2-4 System Architecture: Virtualization Platform for Mission-Critical Applications

ORACLE

Introduction

Software-based virtualization products enable most all industry-standard server models to support workload virtualization. However, not all systems make the best virtualization platforms. In order to support the execution of multiple workloads without starving applications of vital compute resources, a virtualization platform needs to offer high compute performance, a large memory footprint, and superior I/O expansion capabilities. The performance, memory capacity, I/O, and high reliability characteristics of Oracle's Sun Server X2-4 system make it an ideal mission-critical virtualization platform.

Oracle's Sun Server X2-4 system

By offering high-performance processors, a large memory footprint, and superior I/O bandwidth and expandability, the architecture of Oracle's Sun Server X2-4 system (Figure 1) can help IT managers maximize virtualization results and simplify the creation a high-performance and efficient virtualized business application infrastructure.

Figure 1. The high-performance, compact Sun Server X2-4 system is ideal for virtualized business application infrastructure.

The Sun Server X2-4 system is available with two or four Intel® Xeon® Processor E7-4800 product family CPUs and combines Oracle's system design expertise with the latest Intel core microarchitecture to offer exceptional performance and I/O expansion in a compact three rack unit (3U) chassis. The Intel Xeon Processor E7-4800 product family CPU includes Intel QuickPath Interconnects (QPIs) to support high-speed, low latency communication between processors and to I/O Controller Hubs (IOHs) within the Sun Server X2-4 system. Four Scalable Memory Interfaces on the die of the Intel Xeon Processor E7-4800 product family CPU support fast data transfer to and from memory DIMMs. In addition, the Sun Server X2-4 system utilizes two Intel 7500 IOHs and one ICH10 controller hub to establish connectivity to ten PCI Express 2.0 slots, five USB 2.0 ports, and SATA-2 storage devices.

To meet the large memory requirements of virtualized workloads, Sun Server X2-4 systems support up to 1TB of memory using 16 GB DIMMs. Four built-in Gigabit Ethernet ports and ten PCI Express 2.0 slots also provide extensive I/O expandability. To further increase I/O responsiveness, the Sun Server X2-4 system provides the high-speed throughput of SAS-2 disk devices and a large number of PCI Express 2.0 data transfer buses. The Sun Server X2-4 system also supports SSD Flash technology to accelerate data transfer rates well beyond conventional approaches.

The Sun Server X2-4 system also adds significant value to consolidation and virtualization strategies by offering advanced approaches to energy conservation and system management. The use of two cooling compartments and independent fan speed control promotes targeted airflow. Internal sensors monitor power draw levels and advanced power management capabilities help optimize the efficiency of internal power supplies. To simplify system management, the Sun Server X2-4 system includes an integrated ILOM service processor that offers access to management functions using an Ethernet, serial, or on-board video connection. As with all of Oracle's Sun Fire servers, the Sun Server X2-4 system supports simultaneous execution of 32-bit and 64-bit applications and offers a choice of

operating systems including, Oracle® Solaris, Oracle Enterprise Linux, Oracle VM, Linux, and Microsoft Windows operating environments, providing extensive flexibility and investment protection.

Table 1 details the technical specifications and expansion capabilities of the Sun Server X2-4 system.

TABLE 1. CHARACTERISTICS OF THE SUN SERVER X2-4 SYSTEM

	SUN SERVER X2-4 SYSTEM TECHNICAL SPECIFICATIONS
Enclosure	3 rack unit (3U)
Processor	Two or four Intel Xeon Processor E7-4800 product family CPUs • Intel Xeon processor E7-4870 (10-core, 2.4 GHz, 30 MB L3 Cache, 6.4 GT/sec. QPI, 130 Watts) • Intel Xeon processor E7-4860 (10-core, 2.26 GHz, 24 MB L3 Cache, 6.4 GT/sec. QPI, 130 Watts) • Intel Xeon processor E7-4820 (8-core, 2 GHz, 18 MB L3 Cache, 5.86 GT/sec. QPI, 105 Watts)
Memory	64 DIMM slots (1066 MHz) Up to 1 TB of memory using 16 GB DIMMs 4 GB, or 8 GB Low Voltage DDR3-1333 ECC Registered DIMMs or 16 GB Low Voltage DDR3-1066 ECC Registered DIMMs
Internal Storage	Up to six hot-plug SAS HDDs or SATA SSDs • 300 GB or 600 GB 2.5-inch SAS-2 drive (10,000 RPM) • 32 GB 2.5-inch SATA-1 SLC SSD drive One DVD+/-RW drive (SATA device connected to SATA-USB bridge)
System I/O Interfaces	Two VGA ports, DB-15 (1 Front and 1 Rear) Five USB 2.0 ports, Type A (2 Front, 2 Rear, 1 Internal) Four Gigabit Ethernet ports, RJ-45 (Rear) One Service Processor 10/100 Ethernet port, RJ-45 (Rear) One Service Processor RS-232 Serial port, RJ-45 (Rear)
Expansion Slots	Ten internal low-profile PCI Express 2.0 slots • Eight x8-lane slots OR two x16-lane slots plus four x8 lane slots (Use of x16-lane slot disables adjacent x8 lane-slot) • Two x4-lane slots
Power	Dual redundant, hot-swappable power supply with power sharing Two AC Power inlets120/240V, IEC-320 C-14 (Lower Left Rear)

System Management	Oracle Integrated Lights Out Manager (ILOM) Service Processor • Remote management with full Keyboard, Mouse, Video, Storage (KVMS) • Remote media capability (floppy, CD etc.) • Full DMTF CLI • Browser UI for control of the system through a graphical interface. • IPMI 2.0 compliant for management and control • SNMP v1, V2c, V3 for system monitoring • Monitor and report system and component status on all FRUs

The Sun Server X2-4 system Architecture

The Sun Server X2-4 system design is based on a highly integrated system architecture that emphasizes processor performance, memory footprint, and I/O expansion.

The server implements the Next Generation Intel Microarchitecture utilizing Intel Xeon Processor E7-4800 product family (formerly Westmere-EX) CPUs and the Intel 7500 Chipset (formerly Boxboro). The Intel Xeon Processor E7-4800 product family microarchitecture offers up to ten cores per die, two integrated memory controllers, and a large Level-3 cache. Embedded virtualization technologies provide hardware-based assistance for I/O device virtualization, improved virtualization efficiency, and enhanced connectivity within a virtualized server.

The sections that follow detail the physical and architectural aspects of this system. Details about the Intel Xeon Processor E7-4800 product family and its advantages can be found in Appendix A.

Motherboard

The Sun Server X2-4 system is available in a two- or four-processor configuration. The motherboard layout of a Sun Server X2-4 system with two processors is shown in Figure 2 and a four-processor configuration is illustrated in Figure 3. The design of the Sun Server X2-4 system supports the following system architecture features:

- Two-processor or four-processor Intel Xeon Processor E7-4800 product family CPUs
- Dual integrated memory controller on each processor
- The Intel QuickPath InterConnect architecture, providing 6.4 GT/sec links, delivering up to 25.6 GB/sec of total bandwidth.
- The Intel 82801JB I/O Controller Hub, supporting PCIe, SATA, and USB connectivity
- Ten high-speed PCI Express 2.0 slots for high-performance I/O expansion
- Four Gigabit Ethernet ports with RJ-45 connectors to provide fast network connectivity
- Support for six SAS-2/SATA HDDs/SSDs connected to an embedded SATA Gen2 six disk controller or optional connectivity to a 6 Gb/sec. SAS (SAS-2) RAID Controller

- ILOM service processor with an ASPEED AST2100 integrated communications processor, providing system management, monitoring, and remote Keyboard, Video, Mouse, Storage (KVMS) capabilities

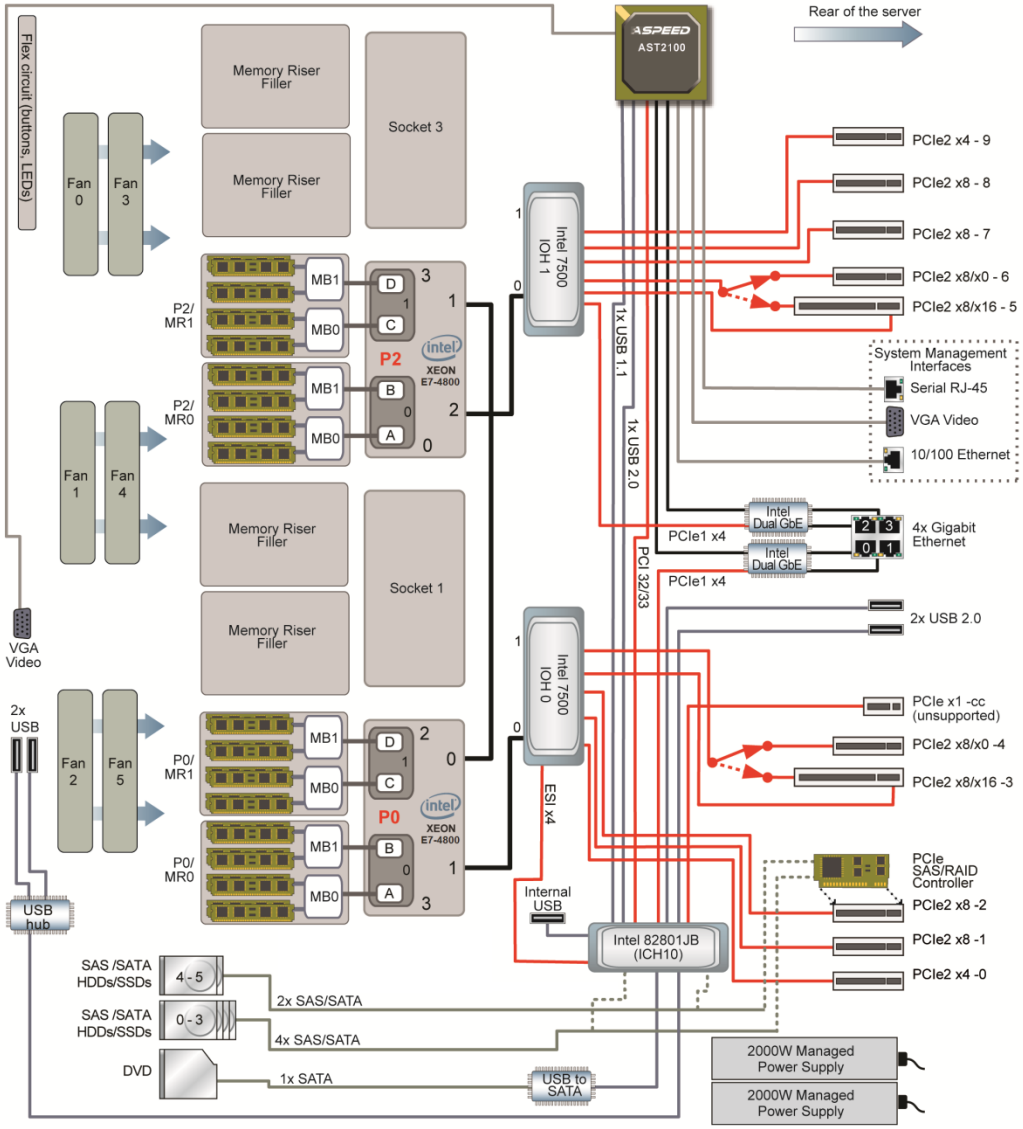

Figure 2. System architecture diagram for a Sun Server X2-4 system two-processor configuration.

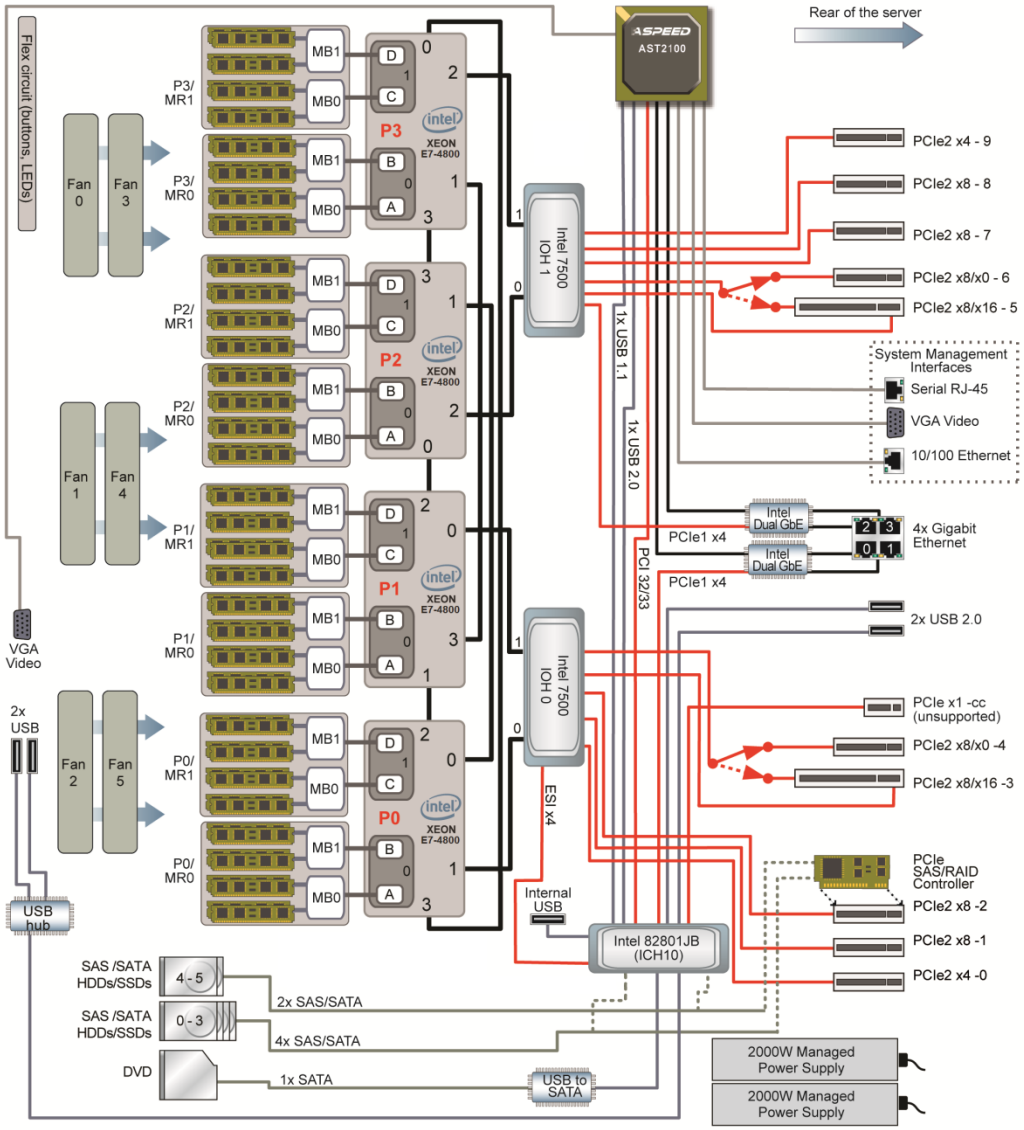

Figure 3. System architecture diagram for a Sun Server X2-4 system four-processor configuration.

Memory

This design of the Sun Server X2-4 system is considered a Non-Uniform Memory Access (NUMA)-style memory architecture, since each processor in multi-socketed systems can access local memory (connected to the local memory controller) as well as remote memory that is connected to another processor. The memory architecture of the Sun Server X2-4 system relies on the use of memory risers that each include eight DIMM slots. Each processor socket in the Sun Server X2-4 system is connected to two memory risers. Configurations with four processors include eight memory riser cards and a total of 64 DIMM slots, while two-processor configurations provide four memory riser cards and a total of

32 DIMM slots. The memory architecture of the Sun Server X2-4 system supports a maximum of 1 TB of memory.

With the introduction of the new Intel Xeon E7-4800 product family, support for low voltage DIMMs is now available. With the support of low voltage DIMMs and integrated reduced-power memory buffers, operating costs can be reduced further by decreasing energy demand.

Memory DIMM slots can be populated with 4 GB, or 8 GB low voltage DDR3-1333 ECC Registered DIMMs or 16 GB low voltage DDR3-1066 ECC Registered DIMMs. DIMM fillers are available to fill empty DIMM slots, helping optimize system cooling. The main guidelines for memory population include the following:

- Populate DIMMs on every memory riser (two, four, six, or eight DIMMs per riser).
- Whenever possible, populate DIMMs on every DDR3 channel (four, six, or eight DIMMs per riser).
- Using a larger number of smaller DIMMs offers the potential to provide better performance than a smaller number of larger DIMMs.
- Decrease memory latency by enabling Hemisphere mode—install equal DRAM capacity between the two risers of each CPU.

The following DIMM population rules and best practices can help organizations create configurations with the best possible performance.

- Use DIMMs with identical part numbers within each SMI lock-step pair.
- Populate DIMMs across the memory controllers of a single processor socket identically — use the same quantity and size DIMMs on both memory risers.
- For best performance, populate two DIMMs per channel using identical DIMMs.
- For configurations that are not fully populated, install one DIMM onto all eight DRAM channels of each populated CPU socket using identical DIMMs.
- Using higher rank DIMMs results in better performance — quad rank DIMMs are better than dual rank DIMMS, and dual rank DIMMS are better than single rank DIMMs.
- For optimal performance, there should be at least one pair of same-capacity DIMMs per memory riser. If there is more than one pair of same capacity DIMMs per riser, start populating using the largest capacity DIMMs first.
- If there are fewer pairs than the number of memory risers, spread pairs across risers, starting with MR0 on the lowest numbered CPU.
- If low voltage DIMMs are mixed with standard 1.5V DIMMs, the low voltage DIMMs will operate at 1.5V instead of 1.35V.

PCI Express

The Sun Server X2-4 system supports 10 low-profile PCI Express 2.0 slots. To maximize bandwidth and speed data transfer, bus paths for these slots are divided among two Intel IOH ASICs. As shown in Figure 4, slots 0, 1, 2, 3, and 4 are assigned to IOH0 and slots 5, 6, 7, 8, and 9 are assigned to IOH1.

Slots 0 and 9 are connected to the IOHs using 4 lanes (x4, 8 lane mechanical). All other slots are connected to the IOHs using 8 lanes (x8). As such, slot 3 and slot 5 generally operate as x8. However, an automatic switching mechanism promotes these slots to x16, based on the following criteria at boot time:

- Slot 3 can operate as a 16-lane slot if it contains an x16 card and slot 4 is unpopulated. If not, both slots 3 and 4 operate as 8-lane slots.
- Slot 5 can operate as a 16-lane slot if it contains an x16 card and slot 6 is unpopulated. If not, both slots 5 and 6 operate at 8-lane slots.

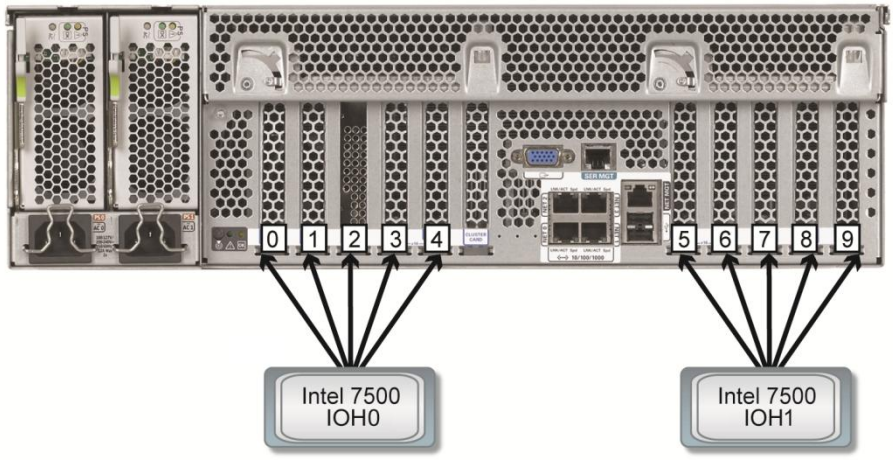

Figure 4. To optimize I/O bandwidth, PCI Express slots are divided among two Intel 7500 I/O Controller Hubs.

Using the following guidelines can help optimize performance by balancing the I/O load across the two IOHs:

- Since the majority of the on-board devices are connected to IOH0 (via the ICH10), install the first PCI Express card that generates the highest IO traffic behind IOH1 by using one of slots 5, 6, 7, 8, or 9.
- For best performance, balance high traffic cards — such as Quad Data Rate InfiniBand, Dual 10 Gigabit Ethernet, and Dual x4 SAS-2 adapters — between the two IOHs, starting with IOH1.
- Populate the x4 slots (slots 0 and 9) with adapter cards that offer the ability to switch to Gen2 speed when inserted in an x4 slot (running an adapter card at x4 Gen2 provides bandwidth that is equivalent to x8 Gen1).
- The most thermally sensitive cards, such as some HBAs with on-board batteries, should be placed in outermost slots (slots 8, 1, 7, and 2). These slots receive slightly cooler airflow than the slots in the center of the chassis.

A PCI Express slot labeled "Cluster Card" exists on the rear of the Sun Server X2-4 system between slot 4 and the Gigabit Ethernet ports. This slot is not supported for use in the Sun Server X2-4 system. Table 2 summarizes the PCI Express slot characteristics and recommended population order, assuming all cards are considered equal.

TABLE 2. PCI SLOT CHARACTERISTICS AND RECCOMENDED POPULATION ORDER

RECCOMENDED POPULATION ORDER	SLOT NUMBER	DESCRIPTION
First	Slot 8	IOH 1
Second	Slot 1	IOH 0
Third	Slot 7	IOH 1, 8-lane: Populating this slot prevents slot 5 from operating with 16-lanes
Fourth	Slot 2	IOH 0, 8-lane: Default slot for PCI Express SAS-2/RAID Controller Card
Fifth	Slot 6	IOH 1, 8-lane
Sixth	Slot 3	IOH 0, 8-lane or 16-lane
Seventh	Slot 5	IOH 1, 8-lane or 16-lane
Eighth	Slot 4	IOH 0, 8-lane: Populating this slot prevents slot 3 from operating with 16-lanes
Ninth	Slot 9	IOH 1, 4-lane (8-lane mechanical)
Tenth	Slot 0	IOH 0, 4-lane (8-lane mechanical)
N/A	Cluster Card slot	Not supported in the Sun Server X2-4 system

Internal Drives

The Sun Server X2-4 system provides six internal disk drive slots. These slots can be populated with a choice of the following devices:

- 300 GB 10000 rpm 2.5-inch SAS-2 hard disk drive
- 600 GB 10000 rpm 2.5-inch SAS-2 hard disk drive
- 32 GB 2.5-inch SATA-2 SLC SSD drive
- 100 GB 2.5-inch eMLC SSD
- 300 GB 2.5-inch eMLC SSD

Using the six internal hot-swappable drive bays, the Sun Server X2-4 system can support up to a maximum of 3.6 TB of disk capacity (using 600 GB drives). Two mini-SAS cables connect the disk

backplane to either the onboard SATA controller or an optional SAS-2 RAID Controller HBA that is seated in PCI Express slot 2 or slot 4.

The following two RAID controllers are compatible with the Sun Server X2-4 system:

- Sun StorageTek SAS-2 HBA — Provides basic HW RAID 0, 1, 1E.
- Sun StorageTek SAS-2 RAID HBA — Offers HW RAID 0, 1, 1E, 5, 5EE, 6, 10, 50, 60.

In addition to disk drives, the Sun Server X2-4 system supports the installation of an optional DVD+/-RW drive. Although the Sun Server X2-4 system DVD drive is technically a SATA device, this drive connects to a USB-SATA bridge and appears to the system software as a USB storage device.

Oracle Sun Flash Accelerator F20 PCIe Card

Based on Oracle's Oracle's Enterprise Flash technology, the Sun Flash Accelerator F20 PCIe card is able to help servers and applications run faster and more efficiently by improving response times and reducing I/O latency. It is capable of delivering the I/O performance of over 100K IOPS, equivalent to over 300 15K rpm disk drives, to eliminate storage I/O bottlenecks. The Sun Fi is capable of supporting up to four Sun Flash Accelerator F20 PCIe cards, each providing an additional 96 GB of enterprise flash memory. The F20 also provides superior reliability and data protection with advanced wear-leveling and write endurance, 25% extra capacity for durability and bad block mapping, and an integrated super capacitor for data integrity and write through persistence. Additional information on the Sun Flash Accelerator F20 can be found at: *http://www.oracle.com/us/products/servers-storage/storage/disk-storage/043966.html*.

Oracle Solaris Zettabyte File System (ZFS) for Oracle's Enterprise Flash technology

Taking best advantage of the performance and cost characteristics of Oracle's Enterprise Flash technology requires an enabling technology that can transparently combine the strengths of HDDs and Flash accelerated storage systems. The Oracle Solaris ZFS file system can harness these technologies within Hybrid Storage Pools to create a solution that offers a balance between the speed of Flash technology and the economy and capacity of HDDs. A Hybrid Storage Pool automatically places data on the most appropriate storage media to optimize performance and manage costs. In fact, the ZFS file system can transparently cache data on SSDs without any need to modify applications. For more information, please see the Sun ZFS Storage Software data sheet at *http://www.oracle.com/us/products/servers-storage/sun-zfs-storage-sw-ds-173241.pdf*.

Chassis

The Sun Server X2-4 system is built within a 3 rack unit (3U) chassis with the following dimensions:

- Height: 5.1 inches (129.85 mm)
- Width: 17.2 inches (436.5 mm)
- Depth Chassis alone: 28.8 inches (732 mm)
- Depth with PSUs Ejector Handles: 29.6 inches (752.35 mm)

- Min Weight = 29.5 kg (65 lbs)
- Max Weight = 38.5 kg (85 lbs)

External Interfaces

External interfaces provided by Sun Server X2-4 systems include the following:

- Two VGA ports, DB-15 (1 front and 1 rear) Both the front and rear video ports are active at all times. However, only the rear video port provides monitor identification.
- Five USB 2.0 ports, Type A (2 front, 2 rear, and 1 internal)
- Four Gigabit Ethernet ports, RJ-45 (rear)
- One Service Processor 10/100 Ethernet port, RJ-45 (rear)
- One Service Processor RS-232 serial port, RJ-45 (rear)
- Ten PCI Express expansion slots
- RFID tag for asset tracking — containing a serial number of the RFID tag also known as the Electronic Product Code (EPC)

The four Gigabit Ethernet ports located on the rear of the server provide 10/100/1000 Mb/sec. full- and half-duplex operation, IEEE 802.3ab Auto-Negotiation for speed, duplex, and flow control, and NIC teaming and failover. In addition, PXE Boot is enabled on all four Gigabit Ethernet interfaces. Gigabit Ethernet NIC0 and NIC1 are connected to IOH0 via the ICH10 while Gigabit Ethernet NIC2 and NIC3 are directly connected to IOH1. If the on-board SATA controller (embedded in ICH10) is used heavily with SATA HDDs, or SSDs, using NIC2 and NIC3 may provide better network performance and lower latency. Figures 5 and 6 provide views of the front and rear of the Sun Server X2-4 system chassis.

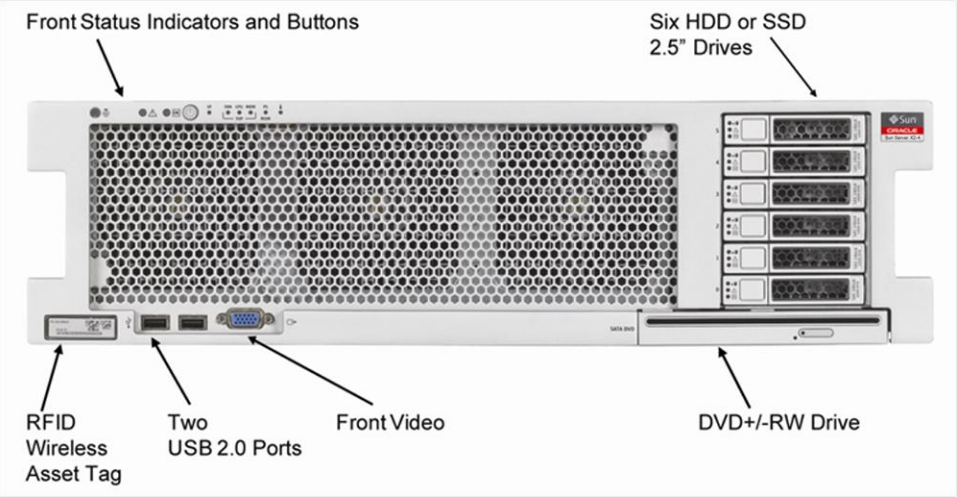

Figure 5. The front panel of the Sun Server X2-4 system features a VGA interface as well as USB ports and access to the internal storage devices.

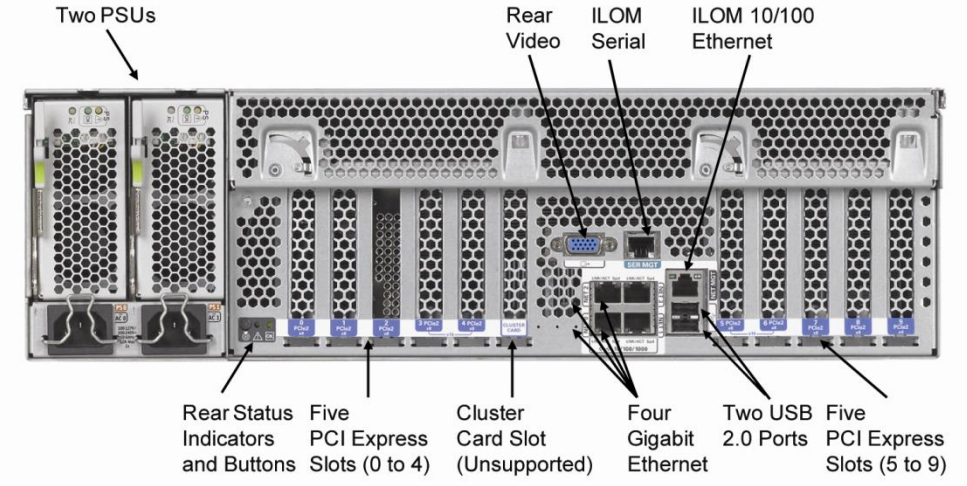

Figure 6. The rear panel of the Sun Server X2-4 system provides access to the PCI Express card interfaces, Gigabit Ethernet and USB ports, and system management connection options.

Indicators

Panel fault indicators include the following:

- Front and rear Locate LEDs (White)
- Front and rear System Fault LEDs (Amber)
- Front and rear Power/OK LEDs (Green)
- SP OK/Fault LED (bi-color Amber/Green)
- Six Front facing Amber Fault LEDs (All Amber)
  - Chassis Alert !
  - Fan Fault !
  - Processor Fault !
  - Memory Fault (DIMM) !
  - Power Supply Fault !
  - Over Temperature Warning !

The Sun Server X2-4 system features a chassis interlock intrusion magnetic switch that is activated by removal or replacement of the top cover. In addition, fault-remind buttons help identify or locate a faulty component even after the system has been disconnected from the power source. The processor, and memory riser components provide an amber fault LED that activates when the corresponding fault-remind button is pressed. For a DIMM fault, pressing the system fault-remind button lights the

fault LED on the module where the faulty component resides. Once the designated module is removed, the fault-remind button on the module can be pressed to identify the specific faulty component.

Cooling

The Sun Server X2-4 system chassis implements front to back cooling with two independent cooling compartments as shown in Figure 7. Within compartment 1, drives bays are cooled by airflow suction from the AC-DC power supply unit (PSU) fans. Compartment 2 covers the bulk of the motherboard and is subdivided into three independently controlled zones. Temperature sensors monitor each zone and the ILOM service processor controls the fan speeds to keep cooling levels optimized. Six 92mm high performance, hot-swap fans are used to cool compartment 2. These fans are arranged in two rows of three fans for redundancy. Air is pressurized in the main compartment for front-to-back cooling. Airflow is also unrestricted by this design, resulting in less noise. The three fans in the rear row must be populated for the system to run. In addition, there is a 30-second replacement time limit for replacing rear row fans while the system is powered on.

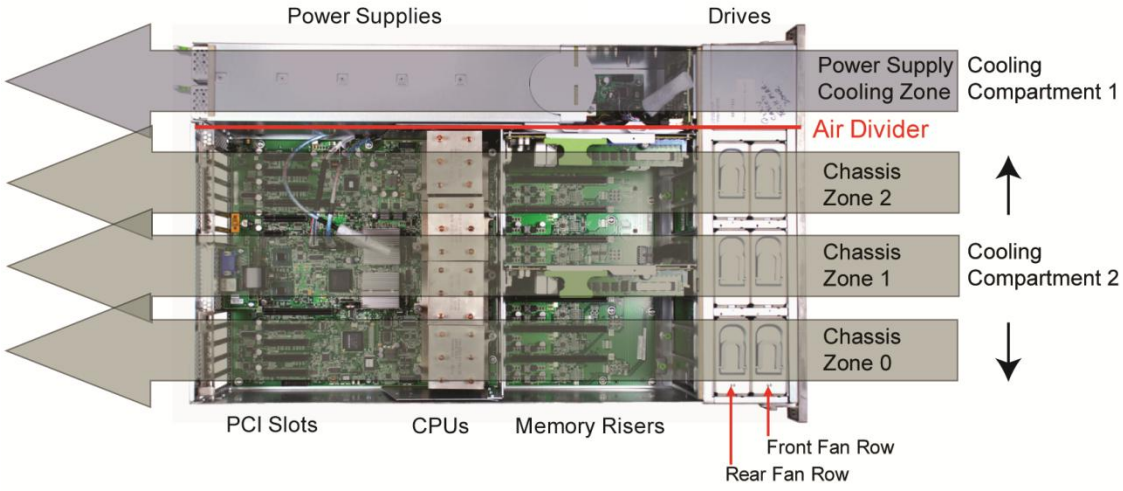

Figure 7. The use of multiple cooling compartments helps optimize the cooling efficiency of the Sun Server X2-4 system.

Power

The Sun Server X2-4 system features redundant power supplies that are rated 12V at 167Amps. Power output ratings are 2060W at 200-240V and 1030W at 100-127V. Power input ratings are 2200W at 200-240V and 1300W at 100-127V. Sun Server X2-4 system configurations with four processors require that both PSUs are connected into a high-line AC (200-240 Volt) power source. If this condition is not met, the server will not boot. Two-processor Sun Server X2-4 systems can run on either high-line (200-240 Volt) or low-line AC (100-127 Volt) power sources.

If needed, the Sun Server X2-4 system PSUs can support power over-subscription. This means that the PSUs in the Sun Server X2-4 system are designed to support 50% power overload for at least 10 milliseconds. In order to return to normal operation, the ILOM service processor invokes software

power capping algorithms to reduce system power consumption until a level below the intended maximum rating for the PSU is reached. As an example, an over-subscription situation may occur when a two-processor configuration connected to low-line power has a single power supply failure. In that scenario the system power consumption may exceed the 1000W max output rating of a single power supply on low-line power.

System Management

To address the need for remote system monitoring and management, the Sun Server X2-4 system provides an integrated ILOM Service Processor. The Sun Server X2-4 system also supports open standards to simplify integration with add-on Oracle and third-party enterprise management software tools.

ILOM Service Processor

Lights-out management features provided by the Sun Server X2-4 system ILOM service processor can help organizations simplify system management tasks. Powered by an ASPEED AST2100 integrated communications processor, the service processor runs independent of the host platform, executing a robust, security-hardened operating system.

Capabilities of the ILOM service processor include the following:

- Full local and remote keyboard, video, mouse, and storage (RKVMS) access via redirection over IP, eliminating the need for KVM switches
- Monitoring and reporting of environmental, power, hardware, BIOS, and operating system events
- Remote power control, diagnostics, media attachment and upgrades of the system BIOS and service processor software
- System configuration information retrieval
- User-configurable serial console access through a physical port or redirected through the management network
- Java™ technology-enabled remote console access across a secure Web connection
- Multi-level role based access with support for RADIUS, LDAP, LDAP/SSL, and Microsoft Active Directory Service lookup of authentication data
- User SSH key authentication and predefined root and user accounts
- Ability to disable the network management port, as well as individual services, such as IPMI, SSH, and KVMS
- Simple Network Management Protocol (SNMP) V1, V2c, and V3 support

For system management operations, the AST2100 uses the following connections:

- Two USB ports for virtual devices (both ports are routed directly on the motherboard between the AST2100 and the ICH10)

- One RJ45 RS-232 serial interface for console redirection
- Two RJ45 10/100 MB/sec Ethernet network interfaces for IP-based management connections (one for dedicated out-of-band management, and one shared host-port for side-band management)
- One HD-15 SVGA video port for local video output (the video signal is routed on the motherboard between the AST2100 and ICH10 using a PCI 32-bit/33 MHz link)

Management functions provided by the service processor are implemented by ILOM 3.0 system management software. This software provides an Intelligent Platform Management Interface (IPMI) Baseboard Management Controller, platform control agents, diagnostics software, and Remote Keyboard, Video, Mouse, and Storage (RKVMS) drivers. Many other Sun servers from Oracle incorporate this same firmware, providing organizations with a single, consistent, and standards-based management interface.

Secure access to the service processor and associated ILOM software functions is available in several ways, including:

- Intuitive browser-based user interface (BUI) over SSL
- Distributed Management Task Force (DMTF) command line interface over Secure Shell (SSH)
- Redirection of the platform console, keyboard, mouse, and video to the Sun ILOM Remote Console application
- SNMP v3 interfaces, providing easy integration with Oracle Enterprise Manger Ops Center, or third-party applications from companies such as Hewlett-Packard and IBM
- IPMI 2.0 command interface, for remote management using IPMI-based tools, such as IPMItool

Maximizing Management Flexibility

While system management tools can play an important role in streamlining operations, organizations must consider the best approach for each environment. Executing management software directly on the host with or without use of a service processor is known as in-band management. Using a dedicated Ethernet or serial port to execute administrative tasks independent of the host is known as out-of-band management. Sharing a single Ethernet port for host and service processor network connectivity is known as side-band management. The Sun Server X2-4 system offers extensive flexibility by supporting in-band, out-of-band, and side-band management. Table 3 offers a comparison of each of these management strategies.

Organizations can choose a single management method or utilize in-band management in tandem with out-of-band or side-band management approaches. The Sun Server X2-4 system provides out-of-band management across a dedicated 10/100 Mb Ethernet port or a RS232 serial port on the ILOM service processor. Side-band management is supported over one of the Gigabit Ethernet interfaces that is shared by both the host and the service processor. When utilizing a side-band management approach, the single, shared physical Ethernet port provides unique MAC addresses and IP addresses for both the platform and the service processor. Side-band management can provide all the benefits of out-of-

band management at considerable cost savings, by eliminating the need to consume a switch port for both the management and platform connection.

In-band Server Management

In-band server management offers the opportunity for organizations to take advantage of industry-standard protocols and applications across all datacenter platforms. The Sun Server X2-4 system facilitates in-band server management by supporting the IPMI 2.0 and SNMP v1, v2c, and v3 standards. Operating system-resident platform management functions for the Sun Server X2-4 system are accomplished using one of the following two options:

- IPMI with a Keyboard Controller Style (KCS) interface and an IPMI kernel driver
- SNMP agents

Out-of-band and Side-band Server Management

Out-of-band and side-band management approaches support the completion of administrative tasks without placing unnecessary burden on the host. This administrative strategy is desirable for performance-intensive environments. While in-band management only works as long as the host operating system is up and running, out-of-band and side-band management are fully functional even while the host is powered off. In order for side-band management to work while the host is shut down, the Gigabit Ethernet network controller device operates on stand-by power (similar to the service processor). Although out-of-band and side-band approaches differ in the approach to network connectivity, these methods offer similar capabilities and benefits.

- Based on serial port redirection (serial-over-LAN), serial port connectivity to the ILOM service processor provides direct console access to the command line interface (CLI) and to the system console stream. The CLI is designed to follow the Distributed Management Task Force (DMTF) Command Line Protocol (CLP).
- Utilizing Ethernet connectivity, administrators can access a Web interface or command-line interface (CLI). A secure shell (SSH) session is required for access to the CLI and the Web interface supports both secure (https) and non-secure (http) access. Secure access is the default configuration for Web-based access.

TABLE 3. COMPARISON OF IN-BAND, OUT-OF-BAND, AND SIDE-BAND MANAGEMENT

	CHARACTERISTICS	BENEFITS	IDEAL USE CASES
IN-BAND	Relies upon operating system-resident software Management tasks utilize platform compute resources	Use of a single network connection and switch port minimizes cost and complexity.	Heterogeneous environments with requirements to provide a common administrative tool across all platforms
OUT-OF-BAND	Utilizes a dedicated Ethernet or serial port for administrative traffic Management tasks execute on an independent service processor	Continuous access to management capabilities even when host is disabled Management tasks do not consume host resources Increased security by physically separating management traffic and server data	Environments with compute or bandwidth-intensive applications Projects with complex management requirements or high levels of administrative burden
SIDE-BAND	The host and service processor share an Ethernet port and are each assigned an independent MAC and IP address Management tasks execute on an independent service processor	Continuous access to management capabilities even when host is disabled Management traffic uses part of the host network bandwidth Requires only one switch port, lowering implementation costs Minimizes processing overhead on the host	Cost-sensitive environment

Remote Keyboard, Video, Mouse, and Storage (RKVMS)

The ILOM Service Processor provides access to Keyboard, Video, Mouse, and Storage (RKVMS) remotely over IP. Remote video display is accomplished through the Java Web Start software known as Sun ILOM Remote Console. To set up a system as a remote console, ILOM Remote Console software is downloaded from the ILOM service processor to the target machine[1]. From this point onward, the ILOM Remote Console executes locally. Since ILOM Remote Console does not run on the server, it does not put overhead on the host.

ILOM Remote Console software can be used to redirect the BIOS and setup screens as well as all other platform video output. A true remote video console to the management console is provided by handling the input and output to and from virtual devices and the Sun Server X2-4 system. With 8-bit and 16-bit support and 8 MB of video memory, the SVGA display provides resolutions up to

[1] ILOM Remote Console requires the installation of Java Runtime Environment 5.0 on the management console

1280 x 1024. A single instance of ILOM Remote Console can open multiple sessions to manage several remote servers simultaneously.

The two USB 2.0 ports connected to the ILOM service processor enable remote keyboard, mouse, and storage functions. The ILOM Remote Console software captures keyboard, mouse, floppy, CD, and DVD input on the management console and redirects it over IP to the ILOM service processor. Keyboard, mouse, and storage inputs are then transmitted over the USB ports to the server. The Sun Server X2-4 system interprets these inputs as originating from locally connected USB devices. These devices are called Virtual devices.

ILOM Remote Console can also be used to boot the remote server from a local device. The virtual device can be a local physical device or an image file. Several types of devices can be redirected as virtual devices by ILOM Remote Console, including the following:

- CD/DVD-ROM
- Floppy
- CD/DVD-ROM image (.iso files)
- Floppy image (.img files)

Intelligent Platform Management Interface

Intelligent Platform Management Interface (IPMI) refers to the autonomous monitoring, logging, recovery, and inventory control features that are implemented in hardware and firmware. The key differentiation of Intelligent Platform Management is that these functions are independent of the main CPU, BIOS, and OS. There are two major components of platform management: the system management controller (or BMC) and System Management Software (SMS). Intelligent Platform Management provides key capabilities that allow enterprise-class management for high-availability systems.

The ILOM service processor provides autonomous sensor monitoring and event logging. Typical sensor-related events are out-of-range temperature or voltage and fan failure. When an event occurs, it is noted in the system event log and made available to the system management controller. The system management controller is powered by the power supply stand-by voltage and will function even when the server is powered down or the operating system has crashed. As a result, platform status can be obtained and recovery initiated under situations where in-band delivery mechanisms are unavailable.

In modern systems, the Intelligent Platform Management Interface (IPMI) provides a hardware-level interface specification for monitoring and control functions. It defines a standard, abstract, message-based interface between the BMC and SMS and a common set of commands for operations such as accessing sensor values, setting thresholds, logging events, and controlling a watchdog timer. IPMI messages can be used to communicate with the BMC over serial and LAN interfaces, so software designed for in-band (local) management can be re-used for out-of-band (remote) management simply by changing the low-level communications layer.

IPMItool

IPMItool is a simple command-line interface to systems that support the IPMI v2.0 specification. IPMItool provides the ability to remotely read the sensor data repository and print sensor values, display the contents of the system event log, print field-replaceable unit information, read and set LAN configuration parameters, and perform remote chassis power control. IPMItool was originally written to take advantage of IPMI-over-LAN interfaces. This tool is also capable of interfacing with the system through a Linux kernel device driver such as OpenIPMI or the Oracle Solaris BMC driver within the Oracle Solaris 10 IPMItool is available under a BSD-compatible license.

System Management Software (SMS) is generally complex and is only part of a much larger management picture. However, system administrators and developers can use command-line tools and scripting to manage these complex systems. IPMItool takes a different approach to SMS and provides a completely command line-oriented tool.

IPMItool is not designed to replace the OpenIPMI library. Where possible, IPMItool supports printing comma-separated values for output to facilitate parsing by other scripts or programs. IPMItool is designed to run quick command-response functions that can be as simple as turning the system on or off, or as complex as reading in the sensor data records and extracting and printing detailed sensor information for each record.

SNMP Support

The Simple Network Management Protocol (SNMP) provides remote access to monitor and control network devices and to manage configurations, statistics collection, performance, and security on a network. SNMP is a network management protocol used almost exclusively in TCP/IP networks. The Sun Server X2-4 system provides SNMP MIBs (Management Information Base) to manage and monitor the servers using any SNMP-capable network management system, such as HP OpenView Network Node Manager (NNM), Tivoli, CA Unicenter, or IBM Director. The MIB data describes the information being managed, reflects current and recent server status, and provides server statistics.

The ILOM service processor supports SNMP v1, v2c, and v3. SNMP v3 is enabled by default; v1 and v2c are disabled by default. SNMP sets may be enabled and disabled and are disabled by default. SNMP traps can be generated from within the service processor. An IPMI-specific trap, called a Platform Event Trap (PET), may also be generated. The following SNMP MIBs are supported:

- The system group and SNMP group from the RFC1213 MIB
- SNMP-FRAMEWORK-MIB
- SNMP-USER-BASED-SM-MIB
- SNMP-MPD-MIB
- ENTITY-MIB
- SUN-PLATFORM-MIB

Oracle® Enterprise Manager Ops Center Software

Oracle Enterprise Manager Ops Center software is a highly scalable datacenter management platform that provides organizations with systems lifecycle management and process automation capabilities to help simplify consolidated platform management, compliance reporting, and system provisioning tasks. Oracle Enterprise Manager Ops Center provides a single console to help discover, provision, update, and manage globally dispersed heterogeneous IT environments, which may include Oracle and non-Oracle hardware running Windows, Linux, and Oracle Solaris operating systems. The capabilities of Oracle Enterprise Manager Ops Center can help enterprises to provision and administer both physical and virtual datacenter assets.

When used in conjunction with Sun Server X2-4 systems, Oracle Enterprise Manager Ops Center can automate patch lifecycle management and maintenance. Oracle Enterprise Manager Ops Center can help system administrators automate software installations, simulation, rollback, compliance checking, reporting, and many other related activities. The framework of Solaris JumpStart software is utilized by Oracle Enterprise Manager Ops Center to effectively provision Oracle Solaris onto individual domains. Oracle Enterprise Manager Ops Center also helps facilitate and control administrative actions from a central location to insure accountability and auditing. These automation capabilities can be used for knowledge-based change management in conjunction with existing configuration management investments. Taking advantage of Oracle Enterprise Manager Ops Center can help organizations create a more reliable environment that offers considerable cost savings through maintenance reduction and the ability to rapidly rebuild systems as needed.

A Choice of Operating Systems

Support for a variety of 64-bit operating systems can help protect investments in server hardware. The Sun Server X2-4 system supports multiple operating systems, letting organizations deploy a choice of application environments without having to shift hardware platforms when software requirements change. This added flexibility can reduce cost and complexity and help organizations to increase return on investment while lowering risk.

The Sun Server X2-4 system supports the following operating systems and versions (at the time of publication of this paper):

- Oracle Solaris 10 09/10 (Update 9)
- Oracle Solaris 11 Express
- Oracle VM 2.2.1, 2.2.2, 3.0.2, 3.0.3
- Oracle Linux 5.5, 5.6, 5.7, 6.0, 6.1, 5.8, 6.2 / 5.5, 5.6, 6.0, 6.1, Unbreakable Enterprise Kernel
- Red Hat Enterprise Linux 5.5, 5.6, 5.7, 5.8, 6.0, 6.1, 6.2
- SUSE Linux Enterprise Server 11 SP 1, 10 SP 4, 11 SP 2
- Microsoft Windows 2008 R2 SP1 Standard/Enterprise/Data Center Edition
- Microsoft Windows 2008 SP2 Standard/Enterprise/Data Center Edition

The latest operating system support information can be found at
https://wikis.oracle.com/display/SystemsComm/Sun+Server+X2-4

Oracle Solaris Operating System

As a part of a broad strategic alliance, Oracle and Intel work together — from design and architecture through implementation — to optimize Oracle Solaris and unleash the power and capabilities of current and future Intel Xeon processors. Oracle Solaris and the Intel Xeon processor E7 family based platform work together to maximize the performance and efficiency of the Sun Server X2-4 system, providing:

- **Improved performance.** Oracle Solaris takes advantage of unique Intel Xeon Processor E7-4800 product family features, including Intel Hyper-Threading Technology, Intel Turbo Boost Technology, and Intel QuickPath Technology, to drive exceptional performance. Solaris provides an outstanding threading model for commercial workloads, outperforming the competition on customer applications as well as industry-standard benchmarks. With specific optimizations for Intel Xeon Processor E7-4800 product family CPUs, Oracle Solaris enables new levels of performance as applications incorporate multi-threaded design, increasing throughput, responsiveness, efficiency, scalability, and overall performance.
- **Automated power efficiency and utilization.** Oracle Solaris has been optimized to take advantage of Intel's power management functions, improving energy efficiency and performance-per-watt. Intel Intelligent Power Technology enables policy-based control that allows processors to operate at optimal frequency and power. Oracle Solaris can make this determination automatically, or administrators can designate which applications require high-frequency processing and which should be executed at lower frequencies to conserve power. Automated low-power states automatically put processor and memory into the lowest available power states that will meet the requirement of the current workload, greatly enhancing power efficiency.
- **Increased reliability, availability, and serviceability (RAS).** The Oracle Solaris Fault Management Architecture (FMA) infrastructure is enhanced to take advantage of the Intel Xeon Processor E7-4800 product family RAS features to provide a dependable enterprise compute platform. For example, Intel MCA recovery enables the system to detect and correct errors in memory and cache that were previously "uncorrectable" through ECC or other means. MCA accomplishes this by first detecting and containing errors before the data is consumed by an application, then works in conjunction with Oracle Solaris to determine the best course of action to keep the system and application running. This advanced recovery capability means that systems equipped with Intel Xeon Processor E7-4800 product family CPUs and Oracle Solaris can recover and keep running in situations where other x86-based systems may fail to continue operating.
- **Virtualization enhancements.** Oracle Solaris takes advantage of Intel Virtualization Technology features to offer cost-effective virtualization options. In fact, many of the unique, built-in features of Oracle Solaris contribute to the value of virtualization running on Intel Xeon Processor E7-4800 product family CPUs. Oracle Solaris offers a top-to-bottom engineered approach to virtualization where the hardware, the hypervisor, the operating system, networking, and the ZFS file system are

all designed to deliver optimal performance and manageability. In addition, Oracle Solaris offers features that help improve overall IT efficiency. For example, Oracle Solaris Containers safely isolate resources, allowing enterprise administrators to maintain the one-application per-server deployment model when consolidating many single servers into a larger system equipped with Intel Xeon Processor E7-4800 product family CPUs.

Additional details on Oracle Solaris optimizations for Intel architectures can be found at: *http://www.oracle.com/technetwork/server-storage/solaris/overview/index-jsp-137397.html*. Additional details about the Intel Xeon Processor E7-4800 product family can be found in Appendix A.

Linux and Microsoft Windows Environments

Sun Server X2-4 systems support Oracle Linux, Red Hat Enterprise Linux and SUSE Linux operating environments. Recognizing the demand for true enterprise-quality Linux support and seeing an opportunity to significantly reduce IT infrastructure costs, Oracle offers Linux operating system support contracts. With the Oracle Unbreakable Linux support program, Oracle is committed to delivering high-quality, comprehensive, and integrated support solutions to help ensure organizations realize success with the Linux operating system.

Oracle also certifies and supports the Microsoft Windows operating environment on Sun Server X2-4 systems. Support for Microsoft Windows increases the flexibility of the Sun Server X2-4 systems and provides organizations the freedom to choose solutions that best meet business needs.

Virtualization Platforms

Oracle VM is a no-cost, next-generation server virtualization and management solution that makes enterprise applications easier to deploy, manage, and support. Backed worldwide by affordable enterprise-quality support, Oracle VM facilitates the virtualization of enterprise application workloads on Sun Server X2-4 systems. Adopting Oracle VM for Sun Server X2-4 system deployments can help reduce operations and support costs while simultaneously increasing IT efficiency and agility. As an alternative, virtualization solutions from VMware are also compatible with Sun Server X2-4 systems. By taking advantage of technology from VMware, enterprises can further capitalize on the high-performance, scalability, availability, and energy efficiency of the Sun Server X2-4 system.

Conclusion

The Sun Server X2-4 system is a powerful and affordable system that can help optimize business application infrastructure performance and flexibility. Supporting powerful Intel Xeon Processor E7-4800 product family CPUs and extensive memory and I/O expansion, the Sun Server X2-4 system is ideal for virtualized business application infrastructure projects, as well as compute and memory intensive applications. Built for rack-dense system environments, the Sun Server X2-4 system also offers a compact 3U design and features to enhance power and cooling efficiency.

For More Information

To learn more about the capabilities and benefits of Oracle's Sun Server X2-4 systems please contact an Oracle sales representative, or consult the related Web sites listed in Table 4.

TABLE 4. WEB SITES FOR MORE INFORMATION

DESCRIPTION	URL
Sun x86 Systems from Oracle	http://oracle.com/goto/x86
Oracle Solaris	http://www.oracle.com/solaris
Oracle VM	http://www.oracle.com/us/technologies/virtualization/oraclevm/index.html
Oracle Linux	http://www.oracle.com/us/technologies/linux/index.html
Oracle Enterprise Manager Ops Center Software	http://www.oracle.com/us/products/enterprise-manager/opscenter

Appendix A—Intel® Technology

The Sun Server X2-4 system supports two or four Intel Xeon Processor E7-4800 product family CPUs interfacing to each other and to two Intel 7500 I/O Hubs (IOH) over Intel QPI links. Systems with two processors installed offer a single connection between the two processors, while systems populated with four processors provide redundant connections between CPUs. Both types of interconnects are shown in Figure 8. The two Intel 7500 chipsets (IOHs) primarily serve as QPI-to-PCI Express bridges, each one providing 36 PCI Express 2.0 lanes for a peak bandwidth of up to 18 GB/sec. in each direction per IOH. In addition, IOH0 provides 4 PCI Express 1.0 lanes, which are used to interconnect with the 82801JB (ICH10) Controller Hub. The Intel ICH10 chip provides bus support for many types of devices, including those with SATA or USB connections.

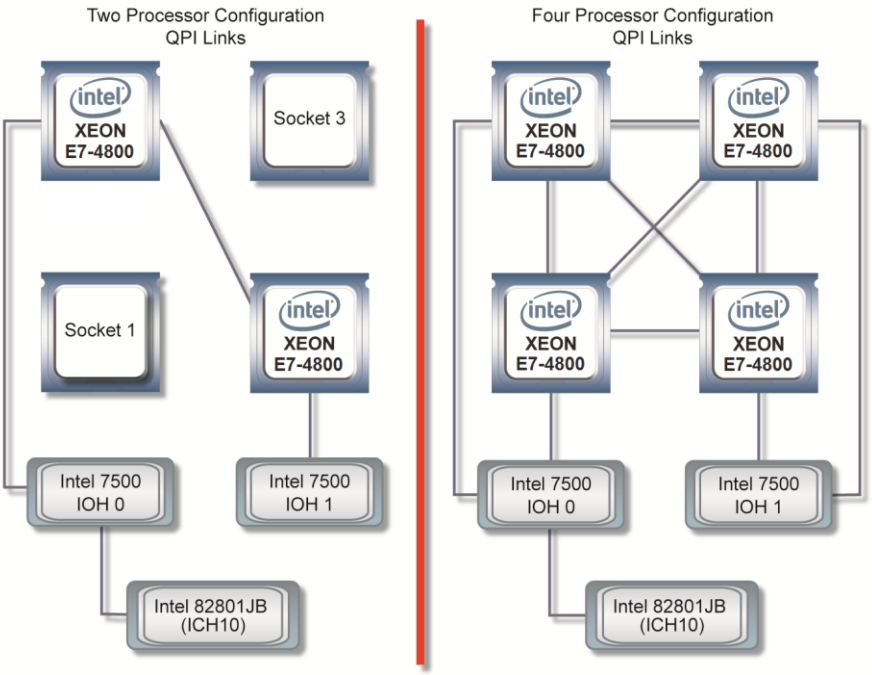

Figure 8. Sun Server X2-4 system interconnect diagrams for two- and four- processor configurations.

Intel Xeon Processor E7-4800 Product Family

The Intel Xeon Processor E7-4800 product family CPU illustrated in Figure 9 features the following:

- Up to ten cores per die
- Up to 30 MB Level-3 shared inclusive cache
- Two Integrated Memory Controllers (IMCs) with two Intel Scalable Memory Interfaces (SMIs) each
- Four full-width, bidirectional Intel QPI buses

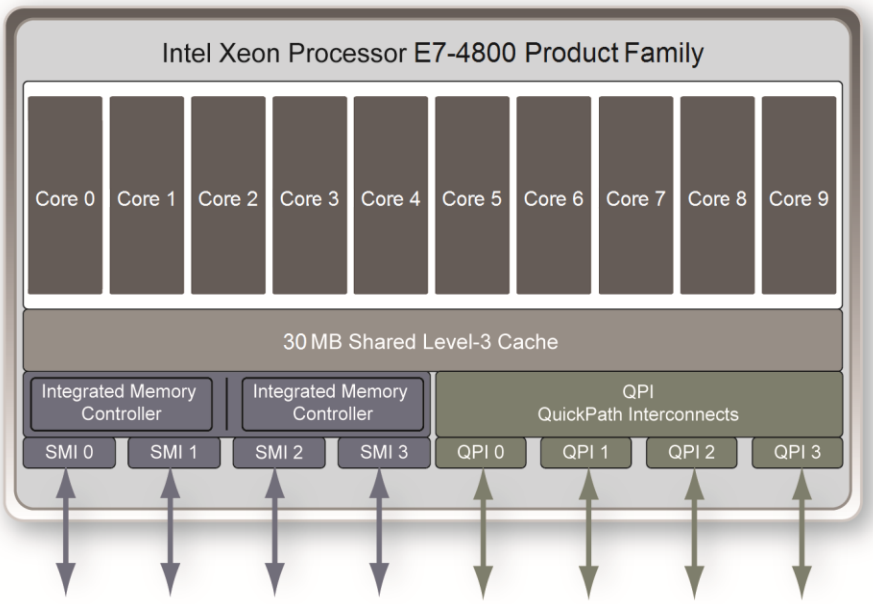

Figure 9. The Intel Xeon Processor E7-4800 product family CPU supports high performance for virtualized workloads.

Within the Sun Server X2-4 system, Intel QPI technology provides high-speed, point-to-point interconnects between processors and I/O. Because QPI provides a point-to-point interconnect, processors do not contend for a single bus when accessing I/O, and do not compete for bus bandwidth — enhancing scalability. Each QPI port includes two unidirectional links that support up to 6.4 GT/sec per link, offering up to 12.8 GB/sec bandwidth in each direction for a total bandwidth of 25.6 GB/sec per port — significantly higher than previous bus designs.

Additional features of the Intel Xeon Processor E7-4800 product family that help drive optimal performance for bandwidth-intensive, threaded applications include the following:

- **Intel HyperThreading (HT) technology.** This technology provides two threads per core, increasing performance for highly threaded applications.
- **Intel Turbo Boost Technology.** For both multithreaded and single-threaded workloads, this technology increases performance by taking advantage of processor and system power and thermal headroom. At times when a system is operating below maximum power and thermal ratings, Turbo Boost pushes the processors to operate cores at a higher frequency than normal. The frequency increase realized is up to 3 speed bins depending on the CPU type and the headroom available. Table 5 lists the number of additional speed bins possible based on the CPU type and number of active cores.

TABLE 5. TURBO BOOST TECHNOLOGY CAPABILITIES BY CPU TYPE AND NUMBER OF ACTIVE CORES

NUMBER OF ACTIVE CORES	INTEL XEON PROCESSOR E7-4870 AND INTEL XEON PROCESSOR E7-4860	INTEL XEON PROCESSOR E7-4820
9 or 10	1 step (+133 MHz)	Not Applicable (8-core CPU)
7 or 8	1 step (+133 MHz)	1 step (+133 MHz)
5 or 6	2 steps (+266 MHz)	1 step (+133 MHz)
3 or 4	3 steps (+400 MHz)	1 step (+133 MHz)
1 or 2	3 steps (+400 MHz)	2 steps (+266 MHz)

Turbo Boost is enabled and disabled in BIOS and controlled by the operating system. The amount of time the system spends in Turbo Boost depends on workload, operating environment, and system design. More information on Turbo Boost is available at *http://www.intel.com/technology/turboboost*.

- **Intel Intelligent Power Technology.** When a processor workload decreases, unneeded components — cores, cache, and memory — are put into sleep mode to reduce power consumption.
- **Intel Hardware-Assisted Virtualization.** This technology improves virtualization efficiency, and enhances connectivity within a virtualized server. For example, Intel Hardware-Assisted Virtualization can help reduce virtual machine monitor (VMM) interventions by eliminating the need for the VMM to listen, trap and execute certain instructions on behalf of the guest OS as is required in software-only virtualization. The technology also provides hardware support for transferring platform control between the VMM and guest operating systems. As a result, handoffs are faster, more reliable, and more secure when VMM intervention is required.

Intel 7500 Scalable Memory Architecture

Each Intel Xeon Processor E7-4800 product family CPU provides two integrated memory controllers that each operate on a pair of interlocked memory channels. By default, memory is interleaved between the two memory controllers belonging to each processor[2]. A pair of SMI links connects each integrated memory controller to the memory subsystem. To increase reliability and bandwidth, the SMI links that originate from the same memory controller operate in lock-step fashion to access memory DIMMs. Intel 7500 Scalable Memory Buffers (MBs) control SMI link access to the memory DIMMs. Each MB connects to one SMI link and up to four DIMMs using two DDR3 channels. Intel 7500 Scalable

[2] Interleaving across memory controllers requires that equivalent quantities of memory be populated below each memory controller.

Memory Buffers are responsible for handling Intel SMI channel transactions with the processor and memory requests to and from the local memory DIMMs interfacing with the memory buffer on its DDR3 channels. Each DDR3 channel supports up to two x4 and x8 Single-Rank, Dual-Rank, or Quad-Rank DIMMs and offers error detection and correction features. Figure 10 provides an illustration of the memory architecture utilized by Sun Server X2-4 systems.

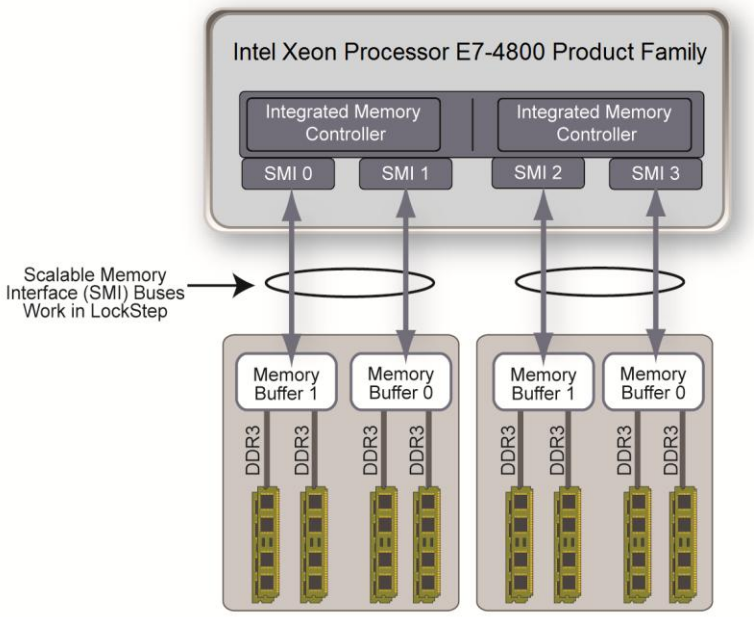

Figure 10. The Intel Xeon Processor E7-4800 product family CPU includes two integrated memory controllers and two pairs of scalable memory interface buses.

The integrated memory controller on the Intel Xeon Processor E7-4800 product family CPU supports several advanced RAS features, including the following:

- **x4 and x8 SDDC.** As an advanced form of ECC technology, x4 and x8 SDDC protects computer memory systems from any single memory chip failure. This technology is capable of detecting and correcting the complete failure of a single x8 DRAM device or the failure of two adjacent x4 DRAM devices. SDDC performs this function by scattering the bits of an ECC word across multiple memory chips, such that the failure of any one memory chip affects only one ECC bit.
- **Demand and patrol scrubbing.** This technology proactively searches system memory, repairing correctable errors. In the case of uncorrectable errors, the algorithm permanently marks the memory location as unusable.
- **Data Protection.** The memory controller performs replays on errors to recover from transient errors and supports lane failover and spare lanes to recover from single SMI channel lane failures. Along with each 64 bytes of cache line stored in memory, there are 16 bits for CRC.

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

Oracle's Sun Server X2-4 system Architecture: Virtualization Platform for Mission-Critical Applications
June 2011; Revosed July 2012

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

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

SOFTWARE. HARDWARE. COMPLETE.