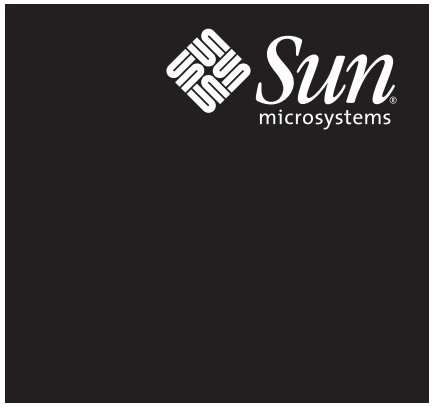
Sun Fire™ V20z Server Architecture



A Technical White Paper
April 2004
SunWIN Token# 400850

Table of Contents

Introduction 1
Architecture Overview 2

CPU Specifications 4
AMD Opteron Processor 4
CPU Packaging 5
Processor VRM 5
Memory Architecture 5
AMD64 Architecture 6
HyperTransport Technology 8
System Performance Enhancements 8

I/O Subsystem 10
AMD 8000 Series Chipset 10
AMD-8131 HyperTransport PCI-X Tunnel 10
AMD-8111 HyperTransport I/O Hub (Southbridge) 11
LSI 53C1020 SCSI Controller 11
SCSI Backplane 12
BCM5703 Gigabit Ethernet Controller 12
XGI Blade 3D Video 14
Super-IO Device Winbond 83627F 14
Floppy Disk Controller (FDC) 14
Serial Ports 14
Keyboard/Mouse controller 15
PCI-X Expansion Slots 15
CD-ROM Drive 16
Power and Reset Sequencer (PRS) 17

Service Processor 18
Service Processor Architecture 19
SP Hardware 19
SP Software 21
BMC 21
Lights Out Management (LOM) 22
Intelligent Platform Management Interface (IMPI) 22
IPMItool 22
SNMP 23

Motherboard Layout 24
Jumpers 24
Platform Power Lights 25
Service Processor Indicator Lights 26
Switches 27

Chapter 1
Introduction

The Sun Fire™ V20z server—the first in a new line of AMD Opteron™ processor-based servers and the first Sun product release resulting from the strategic alliance between Sun and AMD—demonstrates Sun's commitment to delivering the most compelling x86-based platforms in the market. The Sun Fire V20z server is a dual processor, 1U rackmount server that offers both 32-bit and 64-bit high-performance x86 computing at price points comparable to 32-bit only x86 servers.

The Sun Fire V20z server features 64-bit performance with 32-bit x86 compatibility; large memory support (up to 16 GB), I/O capability to support low-latency, high-bandwidth interconnects for compute clusters; Lights Out Management with in-band and out-of-band control; and a wide range of peripheral and ISV support. The Sun Fire V20z server is designed for compute-intensive applications in oil and gas, bio-sciences, EDA (electronic design automation), and MCAE (mechanical computer-aided engineering). The Sun Fire V20z server's wide range of features also helps address issues in areas such as Web services and grid-structured database management.

As part of an end-to-end architecture available from Sun, the Sun Fire V20z server—with its flexibility and performance—creates a myriad of opportunities to build technical and business solutions that match specific customer requirements.

Figure 1-1. Sun Fire V20z server

Architecture Overview

A symmetric multiprocessor, x86-based, rack-optimized system, the Sun Fire V20z server is one rack-unit (RU) high (or 1.69 inches/43 mm), 16.94 inches (430 mm) wide and 28 inches (711 mm) deep. The air-flow direction is from front to back and internal fans are included. I/O ports are located on the rear panel. Informational LEDs are located on the front panel. Access to the power connection is at the rear of the chassis.

The Sun Fire V20z server also has the following system architectural features:

- Embedded single channel DDR memory controllers on each CPU provide maximum memory capacity and bandwidth scaling—delivering up to 16 GB of capacity and 10.7 GB/sec. bandwidth total (3–4x faster than typical x86 servers that use the Northbridge architecture).
- AMD HyperTransport™ Technology CPU-to-CPU and CPU-to-I/O links deliver 6.4 GB/sec. aggregate bandwidth per link.
- Two PCI-X slots deliver high-performance I/O, over 1.5 GB/sec. of I/O plug-in bandwidth.
- Embedded dual Gigabit Ethernet and an Ultra320 SCSI controller leave the two PCI slots available for expansion needs.
- Embedded management and legacy I/O offer maximum operational flexibility without compromising PCI-X slots for optional features.

Figure 1-2 shows the front view of the Sun Fire V20z server while figure 1-3 shows the back view. Figure 1-4 shows a block diagram of the Sun Fire V20z server's architecture.

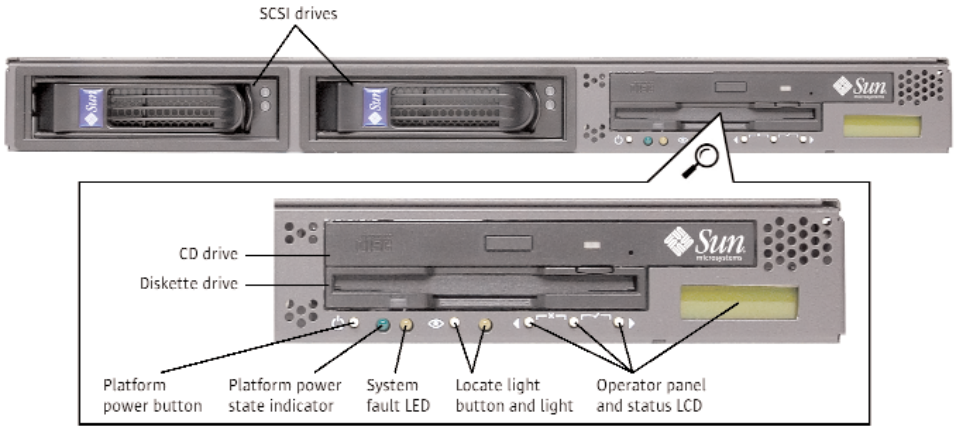


Figure 1-2. Sun Fire V20z server front view

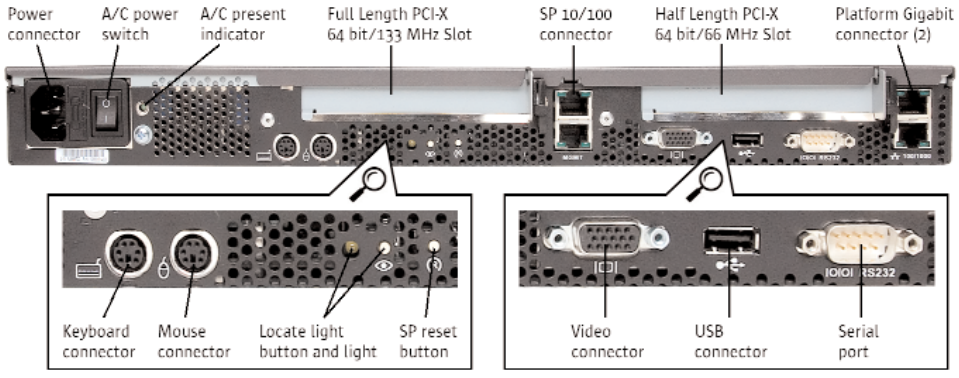


Figure 1-3. Sun Fire V20z server rear view

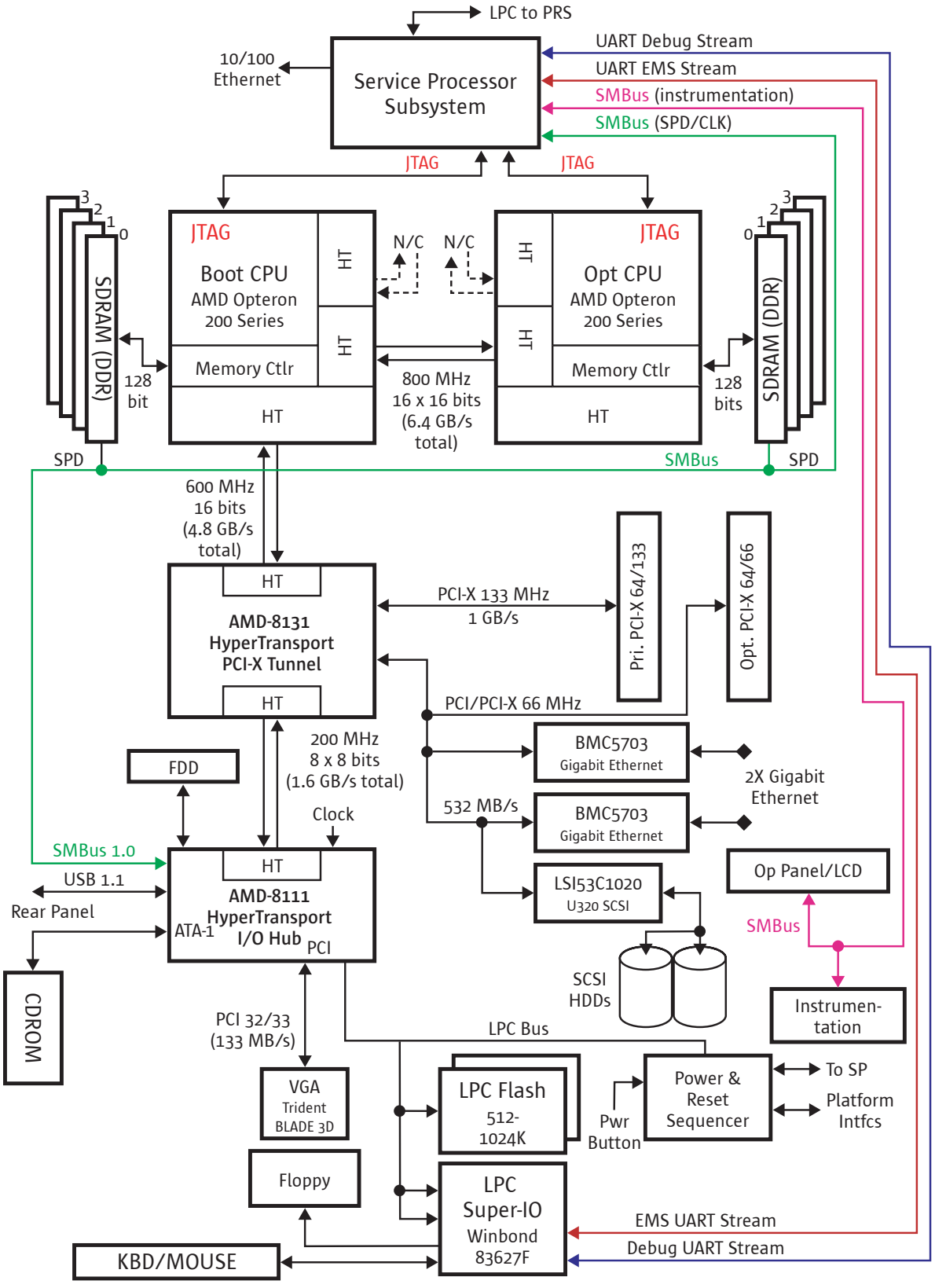


Figure 1-4. Sun Fire V20z server architecture block diagram

Chapter 2
CPU Specifications

The Sun Fire V20z server supports one boot and one optional AMD Opteron processor. The boot processor is inserted in physical position CPU0 (the rightmost socket when viewed from the front of the system). Each processor contains a memory controller supporting a 128-bit wide path to memory and three 16x16-bit HyperTransport links. Each HyperTransport link runs at up to 800 MHz and is clocked on both edges of the clock pulse, allowing for a maximum of 3.2 GB/sec. (1.6 GB/sec., 2 bytes wide) throughput in each direction (6.4 GB/sec. aggregate data rate bi-directionally). Memory support is for PC2700 (DDR333) 184-pin SDRAM DIMMs, four DIMM slots per CPU, accessed in pairs. Since memory attached to an unpopulated processor slot is unavailable, a single processor machine can support a maximum of four DIMMs. A dual CPU server supports a maximum of eight DIMMS or 16 GB (8 x 2 GB) of memory.

AMD Opteron Processor

Features of the AMD Opteron processor on the Sun Fire V20z server include:

- Dual AMD Opteron 200 Series Processors
- AMD64 architecture (64-bit extensions), integrated memory controller, and HyperTransport Technology
- Native support for the 32-bit x86 ISA, MMX, and 3DNow!
- ECC protection for L1 data cache, L2 unified cache, and DRAM
- HyperTransport links supporting up to 6.4 GB/sec. of direct inter-processor bandwidth and up to 6.4 GB/sec. of bandwidth to PCI-X
- CPU L1 Instruction cache: 64KB 2-way associative, parity protected
- CPU L1 Data cache: 64KB 2-way associative, ECC protected
- CPU L2 cache: 1MB 16-way associative, ECC protected

Figure 2-1 shows a block diagram of the AMD Opteron processor.

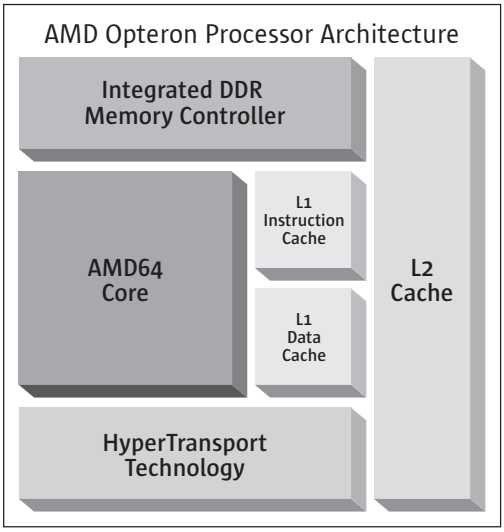


Figure 2-1. AMD Opteron processor block diagram

Note: In a single processor system, the CPU must be in socket 0. In a dual processor configuration, the CPUs must match.

CPU Packaging

AMD Opteron processor packaging specifications include:

- 940-pin ceramic microPGA package
- 1.27 mm pin pitch
- 31 x 31 row pin array
- 40 mm x 40 mm ceramic substrate, C4 die attachment
- 16.5 mm x 11.3 mm die size

Processor VRM

Each CPU receives core voltage from an individual voltage regulator module (VRM). The VRMs regulate the bulk 12 Volt supply to the value required by the CPU core, as specified by the VID code output of the CPU package. Figure 2-2 shows the VRMs on the Sun Fire V20z server.

Memory Architecture

Each CPU device includes an integrated memory controller. Up to four memory modules are supported per CPU. Access by one CPU to the memory residing on the other CPU incurs some latency as data must be forwarded through the interconnecting HyperTransport link. The AMD Opteron memory controller operates as a single channel controller and can be configured to be 8 or 16 bytes wide. The controller supports 1 bit per byte ECC, and supports PC3200 (DDR400) registered DDR SDRAM modules.

IMPORTANT: In single CPU configurations, only the four module positions attached to the boot CPU position (CPU0) can be accessed when populated. In dual CPU configurations, the modules installed for each processor are mapped into global memory space. Memory modules must be installed in pairs, starting with the slots attached to the boot CPU position. The manufacturer and capacity of both modules in the same memory bank (Banks 1–4) must be identical.

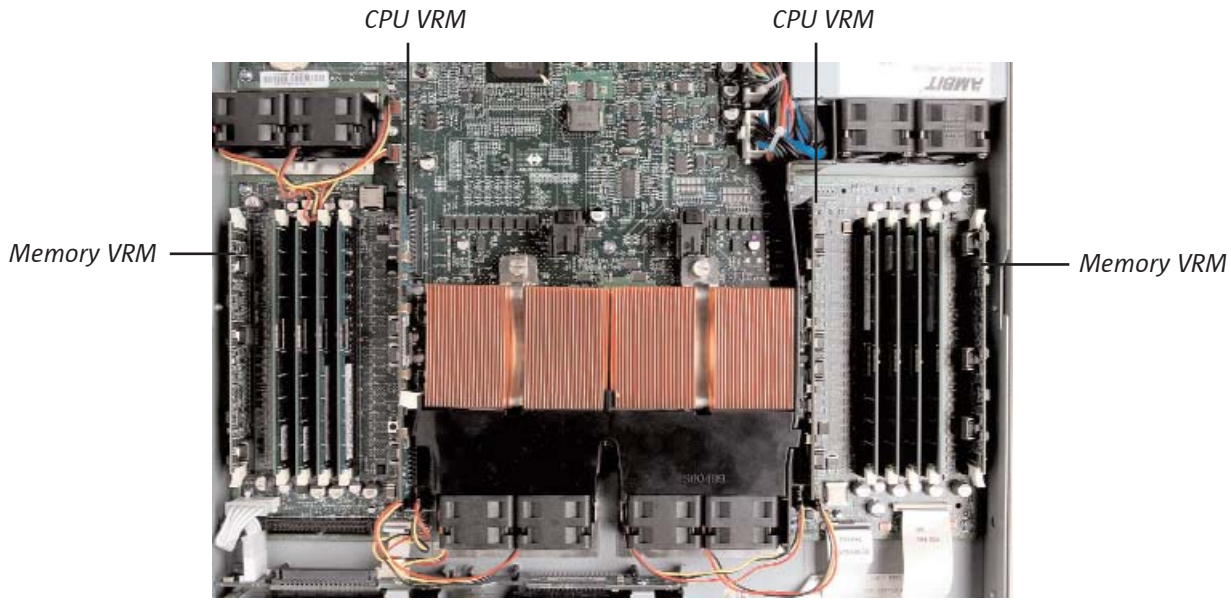


Figure 2-2. Memory and CPU voltage regulator modules on the Sun Fire V20z server

Note: Although the memory architecture supports PC3200 (DDR400) memory, the Sun Fire V20z server currently only supports PC2700 (DDR333) memory. Future system enhancements will include support for faster memory and processor speeds.

Other features of the Sun Fire V20z server's memory architecture include:

- Dedicated on-die 128-bit wide DDR memory controller
- Memory bandwidth of up to 5.3 GB/sec. per processor (with PC2700 modules)
- Under 80 ns latency
- Registered ECC DDR1/333 SDRAM support
- Up to 16 GB (8 GB per CPU) in 8 DIMM slots (4 slots per CPU)
- 256 MB to 2 GB DIMMs support
- Single supply (2.50 VDC)
- Standard SPD (VCC-SPD = 3.3V)

IMPORTANT: The DIMM height must be 1.2 inches (3.05 cm) or less. All four VRMs must be present in a system populated with two CPUs.

AMD64 Architecture

The AMD64 architecture allows end users to run existing 32-bit applications and operating systems at peak performance, while providing a migration path that is 64-bit capable. It is designed to enable 64-bit computing while remaining compatible with the vast x86 software infrastructure, and allows businesses, consumers, and hobbyists the opportunity to run applications that were previously available only on 64-bit workstations, including 3D modeling, rendering, animation, simulation, and software development.

The AMD64 Instruction Set Architecture (ISA) extends the existing x86 ISA and natively executes 32-bit code with no "emulation mode" to degrade performance. For 32-bit software that does not require immediate 64-bit

implementations, AMD64 processor-based systems are designed to provide full application performance while continuing to improve with AMD64 platform performance enhancements.

Many applications encounter architectural barriers that prevent efficient performance scaling. The AMD64 ISA is designed to allow continued performance scaling for applications that demand multiprocessor scalability, larger addressable memory, better multimedia performance, or improvements in computational accuracy.

The AMD64 ISA processor has been designed for applications that:

- Need large memory addressing to handle datasets larger then 3GB per process (financial and scientific modeling applications)
- Must manage a large number of concurrent users or application threads, such as large-scale thin-client solutions, large databases, and data warehouse applications for solutions in customer relationship management (CRM), supply chain management (SCM), enterprise resource planning (ERP), and digital rights management (DRM) systems
- Require real-time encryption and decryption for enhanced security, including e-commerce and protection of private or classified data
- Require mathematical precision and floating-point performance, including modeling, simulation, statistics and financial analysis, imaging/video/signal processing, physics, medical research, telecommunications, encryption, and compression
- Require large, high-power database performance, including decision support, searching and indexing, document and content management, and voice recognition
- Require x86 compatibility or the economies of scale of x86 as well as the large memory addressing capabilities of 64-bit computing, including many high-performance computing (HPC) cluster applications
- Provide digital content creation capabilities such as computer-aided design (CAD), computer-aided manufacturing (CAM), and computer-aided engineering (CAE), digital music production and video editing, and real-time media streaming solutions
- Require maximum performance for realistic and cinematic consumer experiences, including computer games, digital video, and real-time collaboration

AMD64 processors are designed to maintain full compatibility with x86 while providing the architectural enhancements that provide world-class 64-bit performance. With the AMD64 ISA, relevant instructions and encodings have evolved to support 64-bits, increasing the resources available to hardware and software.

Major enhancements over legacy x86 include:

- Sixteen 64-bit general-purpose integer registers that quadruple the general purpose register space available to applications and device drivers as compared to x86 systems
- Sixteen 128-bit XMM registers for enhanced multimedia performance to double the register space of any current SSE/SSE2 implementation
- A full 64-bit virtual address space with 52 bits of physical memory addressing that can support systems with up to 4 petabytes of physical memory—65535 times the amount of RAM supported by 32-bit x86 systems
- 64-bit operating systems to provide full, transparent, and simultaneous 32-bit and 64-bit platform application multitasking

AMD64 processors include HyperTransport Technology and are designed for flexibility and scalability. HyperTransport Technology provides links-based multiprocessing, simplifying the design of multiprocessor workstations and servers. Compatibility with x86 makes the AMD64 computing platform the first 64-bit platform designed to be compatible with mainstream PC applications while offering world-class performance, making it suitable for solutions ranging from consumer client PCs to high-performance clusters. The combination of flexibility and scalability reconciles the broad range of capability and performance requirements IT professionals face today.

HyperTransport Technology

HyperTransport Technology is a high-speed, high-performance, point-to-point link for integrated circuits. At a peak throughput of up to 6.4 GB/sec. bi-directional per link, HyperTransport Technology provides an I/O solution for the most demanding system applications. The AMD Opteron processor with built-in HyperTransport Technology links provides a scalable bandwidth interconnect between processors, I/O subsystems, and other chipsets.

System Performance Enhancements

HyperTransport Technology is designed to increase overall system performance by helping to remove I/O bottlenecks, which improves bandwidth and reduces latency.

Memory Interface

In traditional Northbridge/Southbridge architectures, memory transactions must propagate through the Northbridge chip fabric, creating additional latencies that reduce performance potential. To help resolve this performance bottleneck, AMD incorporates the memory controller into its AMD64 processors. The direct interface to the memory significantly reduces the memory latency seen by the processor. This latency will continue to drop as the processor frequency scales.

Additionally, hardware and software memory pre-fetching mechanisms can further reduce the effective memory latency seen by the processor. This reduction in memory latency, coupled with the additional increase in memory bandwidth available directly to the processor (resulting from this platform architecture design optimization), is critical as it greatly enhances system performance across all application segments.

Chip-to-Chip Interconnect

Current interface schemes offer throughput performance on the order of 266 MB/sec. to 1 GB/sec. Although these rates may be sufficient for desktop platforms, workstation, server, and other future platforms require a more robust interface. The simultaneous integration of high-speed technologies (such as Gigabit Ethernet, PCI-X, and the InfiniBand Architecture) onto high-end platforms will quickly dwarf the bandwidth capabilities of existing interfaces.

HyperTransport Technology provides a high-speed, chip-to-chip interconnect that virtually eliminates the I/O performance bottleneck while providing ample performance headroom for future growth.

I/O Expansion Capability to High-Speed Industry Buses

The traditional Northbridge/Southbridge architecture is not intended to support more than two “core-logic” elements. Adding additional high-speed functionality (such as Gigabit Ethernet, PCI-X, the InfiniBand Architecture, or any combination thereof) would have to occur in one of three ways:

- The functionality would have to be attached to an existing bus interface such as the PCI bus. However, an existing bus may not have sufficient bandwidth to support high-speed technologies, especially in instances in which multiple buses or combinations of buses must be supported simultaneously.
- The functionality would have to be directly attached to the higher-speed, proprietary, chip-to-chip interconnect bus via a bridging device. However, the proprietary nature of this solution may limit the number of components available from vendors, thus impacting cost and availability.
- The functionality would have to be integrated into one of the core logic components. This solution is the least flexible, as a wide range of components would have to be created for each desired combination of feature-set buses.

HyperTransport Technology, an industry standard, provides system designers a high-speed, daisy-chained interconnect between system components. Specific components can be connected in a building-block fashion to achieve a platform with specific feature-set and performance objectives. Figure 2-3 shows a sample Hypertransport technology architecture block diagram.

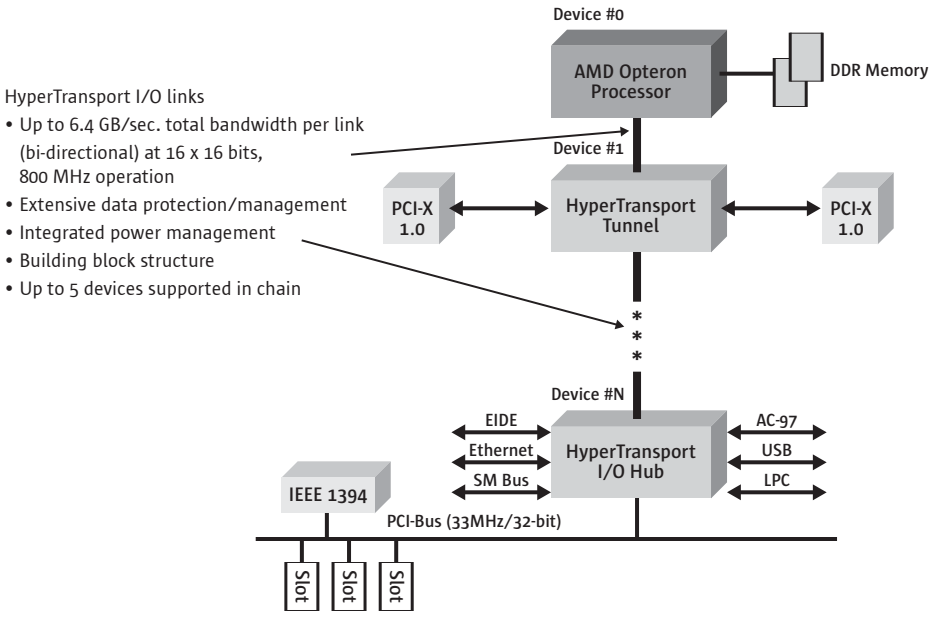


Figure 2-3. Sample HyperTransport Technology architecture block diagram

Chapter 3
I/O Subsystem

The I/O subsystem is connected to the CPU complex through a HyperTransport link from the boot processor. The I/O subsystem consists of the following components:

- A tunneling HyperTransport bridge (AMD-8131 HyperTransport PCI-X Tunnel)
- A HyperTransport-attached SouthBridge component (AMD-8111 HyperTransport I/O Hub)
- Flash memory for BIOS
- Additional peripheral I/O embedded on the motherboard consisting of:
  - A single channel SCSI controller (type LSI 53C1020) attached to the PCI-X Bridge AMD-8131
  - Two 10/100/1000 (Gigabit) Ethernet MAC/PHY devices (type Broadcom BCM5703)
  - A video controller device (type Trident Blade-3D) attached to the PCI bus of the Southbridge
  - Super-IO device (Winbond 83627F) which includes two 16550-compatible UARTs (one externally ported and capable of intercept by the Service Processor); a PS/2-style keyboard/mouse controller (KMC); and a floppy disk controller (FDC)

AMD 8000 Series Chipset

The AMD 8000 series consists of:

- HyperTransport interconnects (discussed in the previous chapter)
- AMD-8131 HyperTransport PCI-X Tunnel
- AMD-8111 HyperTransport I/O Hub

AMD-8131 HyperTransport PCI-X Tunnel

The AMD-8131 HyperTransport PCI-X Tunnel provides high-speed PCI-X capability and offers the following feature-set:

- 8x8 800 MHz downstream HyperTransport link (limited to 200 MHz by Southbridge)
- 16x16 800 MHz upstream HyperTransport link
- Dual PCI-X interface supporting 133 MHz, 100 MHz, 66 MHz, and legacy PCI speeds

The AMD-8131 is a tunneling HyperTransport device, which provides two PCI-X ports. It supports a 16-bit 800 MHz upstream bi-directional HyperTransport link and an 8-bit 800 MHz downstream bi-directional HyperTransport link. Each link supports independent transfer rates and bit width selection. The part supports the HyperTransport link disconnect protocol. The PCI-X bridges in AMD-8131 each support 64-bit addressing and a 64-bit data bus, with transfers at 66, 100, and 133 MHz in PCI-X mode and 33 and 66 MHz in conventional PCI 2.2 mode. Each bridge is independent and supports up to five PCI bus masters with clock, request, and grant signals. Each bridge also includes an IOAPIC that supports legacy interrupt modes.

AMD-8111 HyperTransport I/O Hub (Southbridge)

The AMD-8111 HyperTransport I/O Hub integrates the system I/O functions into a single component. The AMD-8111 HyperTransport I/O Hub feature set includes the following:

- 8-bit 200 MHz upstream HyperTransport link
- 10/100 Ethernet MAC (unused)
- Dual EIDE CD-ROM controller (one only is used)
- AC'97 audio (unused)
- Two USB 1.1 ports (one only is used)
- PC I/O functions (RTC, CMOS, PIT, DMAC, and port control)
- IOAPIC
- PCI 32/33 interface (8 arbiters) used for the XGI (Trident) Blade 3D video
- Low pin count (LPC) legacy bus
- SMBus 1.0 and 2.0 controllers
- ACPI register set and support logic

The AMD-8111 is a HyperTransport-attached Southbridge function that provides several I/O blocks supporting basic peripherals and support functions for the system. AMD-8111 provides an 8-bit 200 MHz incoming HyperTransport bus and a 33 MHz LPC bus for connection to ROM and legacy I/O functions.

LSI 53C1020 SCSI Controller

The Sun Fire V20z server includes an integrated LSI 53C1020SCSI Ultra320 SCSI controller which supports wide Ultra320 SCSI synchronous transfer rates of up to 320 MB/sec. on an LVD SCSI bus. Integrated LVDlink transceivers support both LVD and single-ended signals with no external transceivers required. Fast SCSI, Ultra SCSI, Ultra2 SCSI, Ultra160 SCSI, and Ultra320 SCSI are all supported by the LSI53C1020 controller. Ultra320 SCSI features include:

- Double transition (DT) clocking
- Packetized protocol
- Paced transfers
- Quick arbitrate and select (QAS)
- Skew compensation
- Intersymbol interference (ISI) compensation
- Cyclic redundancy check (CRC)
- Domain validation technology

SCSI Backplane

The SCSI bus from the integrated SCSI controller on the motherboard is routed to the backplane through standard 68-pin connectors and a single twisted-pair cable. The backplane provides support for low-voltage differential signaling (LVD) disk drives only. Full LVD termination for the bus is provided. Figure 3-1 shows a schematic of the SCSI bus routing while figure 3-2 (on the following page) illustrates a pictorial concept of the SCSI backplane.

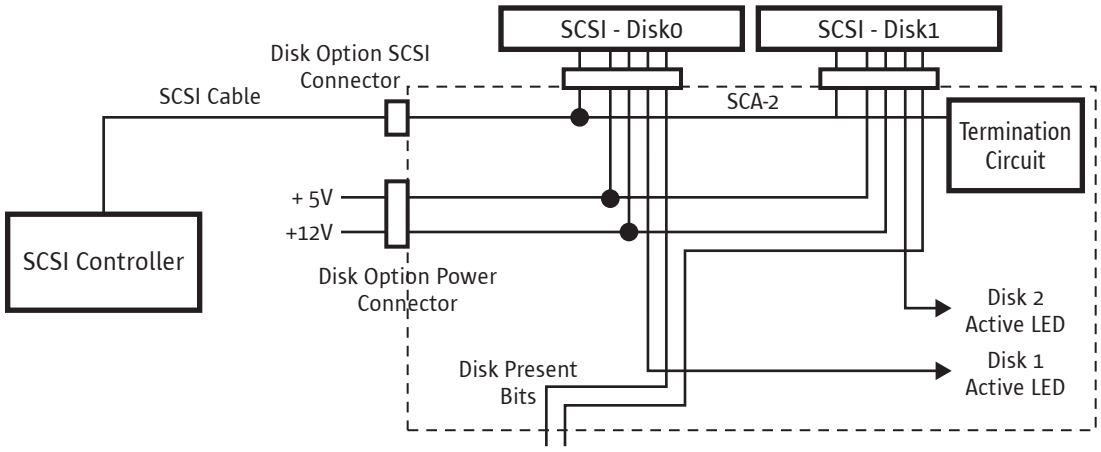


Figure 3-1. Schematic of SCSI bus routing

BCM5703 Gigabit Ethernet Controller

The BCM5703 is a fully integrated 64-bit 10/100/1000-Mbps Gigabit Ethernet media access control (MAC) and physical layer transceiver (PHY). The BCM5703 includes a 10/100/1000-Mbps Ethernet triple-speed MAC with full/half-duplex capability at all speeds, and 10/100/1000 copper PHY. Support for the following 802.3 functions is featured in the MAC:

- VLAN tagging
- Layer 2 priority encoding
- Link aggregation
- Full-duplex flow control

Figure 3-3 shows a block diagram of the BCM5703 Gigabit Ethernet controller's architecture.

Note: The BCM5703 provides Gigabit Ethernet connectivity to the platform side of the Sun Fire V20z server. The Service Processor side uses independent, dedicated 10/100 Ethernet ports for management traffic.

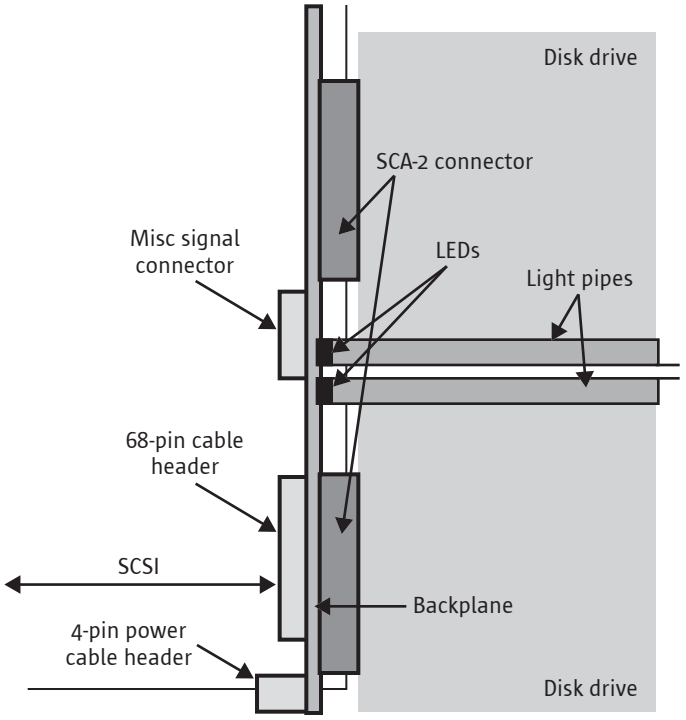


Figure 3-2. SCSI backplane illustration

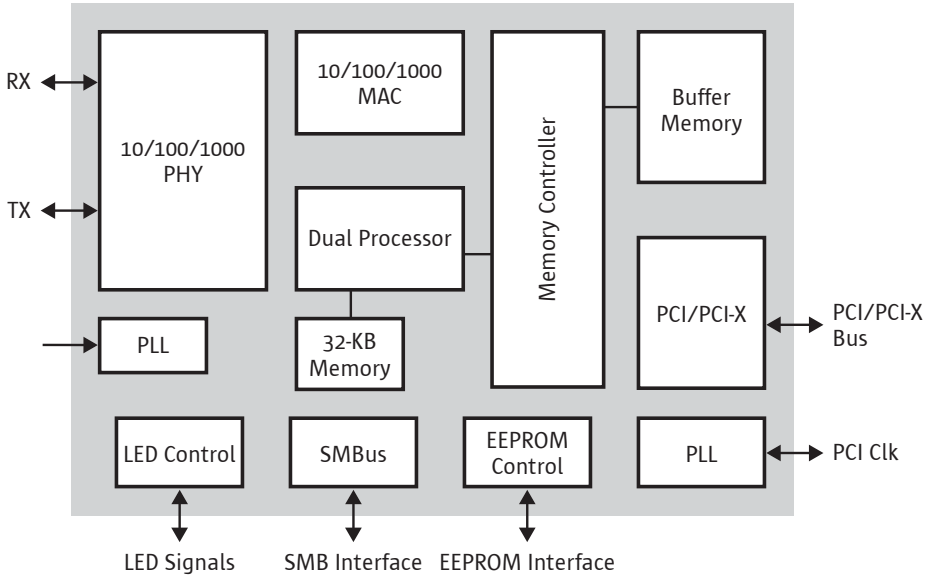


Figure 3-3. BCM5703 Gigabit Ethernet controller architecture block diagram

XGI Blade 3D Video

The Sun Fire V20z uses the XGI Technologies (Trident) Blade 3D as its graphics controller. Features include:

- 64-bit SDRAM/SGRAM memory bus
- 64-bit single cycle 2D/3D graphics engine
- DirectDraw acceleration
- 125 MHz system clock
- 24-bit 170 MHz true color RAMDAC
- Up to 1400 x 1024 pixels resolution
- 8/16/32-bit per pixel color rendering
- DirectX 6.0 and OpenGL support

Super-IO Device Winbond 83627F

The Super-IO device offers multiple I/O functions required to provide legacy I/O devices to the operating system. The Super-IO contains the following I/O devices used by the Sun Fire V20z server.

Floppy Disk Controller (FDC)

The Winbond 82627F supports floppy formats from 360K to 2.88M and transfer rates from 250K to 2Mbps. This device controls the diskette drive for the platform and is brought out from the motherboard via a 34-pin cable connection to a pin header. Table 3-1 shows the pin-out of the floppy disk drive connector.

Table 3-1. Floppy disk drive connector pin-out

Pin	Signal	Pin	Signal
1	5.0 VDC	2	INDEX
3	5.0 VDC	4	DRV_SEL
5	5.0 VDC	6	DISK_CH
7	NC	8	READY
9	HD_OUT	10	MOTOR_ON
11	NC	12	DIR_SEL
13	NC	14	STEP
15	Ground	16	W_DATA
17	Ground	18	W_GATE
19	Ground	20	TRACK0
21	NC	22	W_PROTECT
23	FDD_PRSNT#	24	R_DATA
25	Ground	26	SIDE1_SEL

Serial Ports

Serial ports provide a convenient way to connect the Sun Fire V20z server to terminal servers and other serial devices. The Winbond 82627F contains two high-speed 16550-compatible UARTs, each with 16-byte send/receive FIFOs, a programmable baud rate generator, complete modem control capability, and a processor interrupt system. These ports are routed in TTL form directly back to the Service Processor, with headers on the motherboard for attaching a serial terminal directly to the ports when the covers are open. Serial port 0 is routed to a back-panel panel connector and may be intercepted by the Service Processor. Table 3-2 shows the pin-out of the serial port.

Table 3-2. Serial port pin-out

Pin	Signal	Pin	Signal
1	DCD	2	RXD
3	TXD	4	DTR
5	Ground	6	DSR
7	RTS	8	CTS
9	RI		

Using jumper J19 (see Jumpers section in Chapter 5), the serial port can be switched via an internal multiplexer between connection to Super-IO port 0 and the Service Processor SP-SMC2.

Note: The serial port can also be redirected to the platform or the Service Processor using command line over SSH (see manual). Jumper J19 must be removed in order for the command to operate.

Keyboard/Mouse Controller

The keyboard/mouse controller interface is routed to a DIN connector accessible on the back panel of the server. Table 3-3 shows the pin-out of the keyboard connector while table 3-4 shows the pin-out of the mouse connector.

Table 3-3. Keyboard connector pin-out

Pin	Signal
1	Data
2	NC
3	Ground
4	5.0 VDC
5	CLK
6	NC

Table 3-4. Mouse connector pin-out

Pin	Signal
1	Data
2	NC
3	Ground
4	5.0 VDC
5	CLK
6	NC

PCI Expansion Slots

The leftmost slot (when viewed from the front) can support half- or quarter-length PCI or PCI-X expansion cards. The card can be 32-bit or 64-bit. The slot can operate up to 66 MHz in speed. Since a PCI-X 133 MHz card in the leftmost slot would operate at 66 MHz, high-performance PCI-X 133 MHz cards should be placed in the rightmost slot.

The leftmost slot shares the bus resources with the embedded SCSI hard disk drive and dual port Gigabit NIC. PCI-X architecture automatically degrades the bus to the lowest common denominator. Placing a PCI 33 MHz or a PCI 66 MHz card in the leftmost slot may affect the performance of the embedded disk and ethernet controllers.

The amount of degradation depends on the workload and whether or not the embedded controllers are active. The amount of degradation would be worse for PCI 33 MHz than it would be for PCI 66 MHz. For maximum performance, this slot should be left unpopulated, and, if populated, only 66 MHz-capable cards should be used.

The rightmost slot (when viewed from the front) can support full length PCI or PCI-X expansion cards. The card can be 32-bit or 64-bit. The slot can operate up to 133 MHz in speed. The rightmost slot is on an independent PCI-X bus that does not share PCI bus bandwidth with any other devices.

The Sun Fire V20z server PCI slots use a 3.3 Volt signaling bus and thus, only support 3.3V keyed cards. 5V keyed cards are not supported and may damage the motherboard.

CD-ROM Drive

The Sun Fire V20z server includes a standard CD-ROM drive. The IDE interface provided by the SouthBridge AMD-8111 drives the CD-ROM device. The CD-ROM drive connects to the motherboard via a 50-pin ribbon cable. Table 3-5 shows the pin-out of the CD-ROM connector.

Table 3-5. CD-ROM connector pin-out

Pin	Signal	Pin	Signal
1	NC (left audio)	2	NC (right audio)
3	Analog ground	4	NC
5	Reset	6	Host data 8
7	Host data 7	8	Host data 9
9	Host data 6	10	Host data 10
11	Host data 5	12	Host data 11
13	Host data 4	14	Host data 12
15	Host data 3	16	Host data 13
17	Host data 2	18	Host data 14
19	Host data 1	20	Host data 15
21	Host data 0	22	DMARQ#
23	Ground	24	IOREAD#
25	IOWRITE#	26	Ground
27	IORDY	28	DMACK#
29	IRQ	30	NC
31	DA1	32	PDIAG#
33	DA0	34	DA2
35	CS0#	36	CS1#
37	NC	38	5.0 VDC
39	5.0 VDC	40	5.0 VDC
41	5.0 VDC	42	5.0 VDC
43	Ground	44	CD_PRESENT#
45	Ground	46	Ground
47	CSEL	48	Ground
49	Reserved	50	Reserved

Power and Reset Sequencer (PRS)

The Service Processor (SP) interfaces to the platform primarily through the Power and Reset Sequencer (PRS). The PRS provides a number of functions to the Service Processor sub-system including:

- A path for communication with the platform via LPC
- A common point for interrupts and alerts
- A wide range of registers for monitoring the system power status
- Interfaces for voltage margining
- A heartbeat detector from the SP that is used to sense when the SP is functioning

Chapter 4
Service Processor

The Sun Fire V20z offers multiple management solutions, both in-band and out-of-band. 'In-band' refers to traffic that follows the same path as the normal data. In-band management on the Sun Fire V20z server goes through the 'platform side' (using the Gigabit Ethernet ports and the running operating system). 'Out-of-band' refers to the traffic that travels on a separate media from the data. Out-of-band management on the Sun Fire V20z uses the 'Service Processor side' (using the dedicated 10/100 MB/s management ports and the SP).

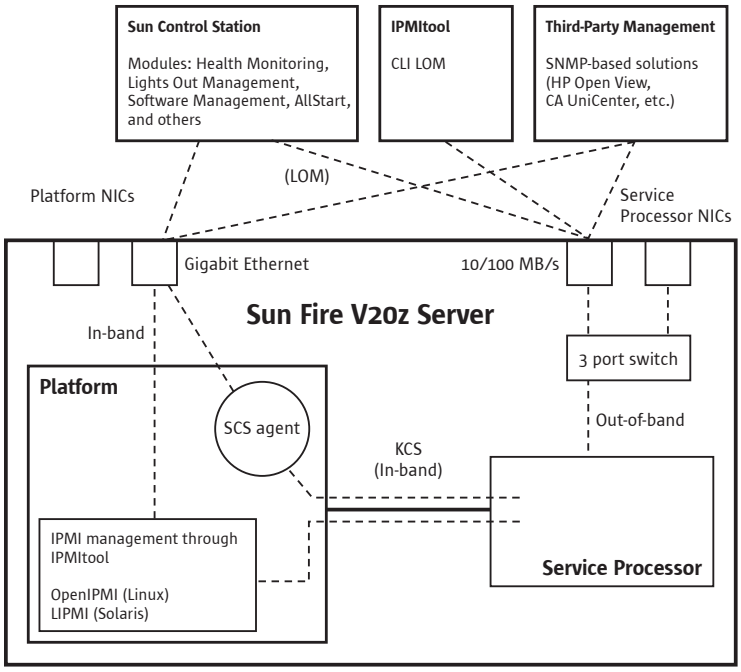


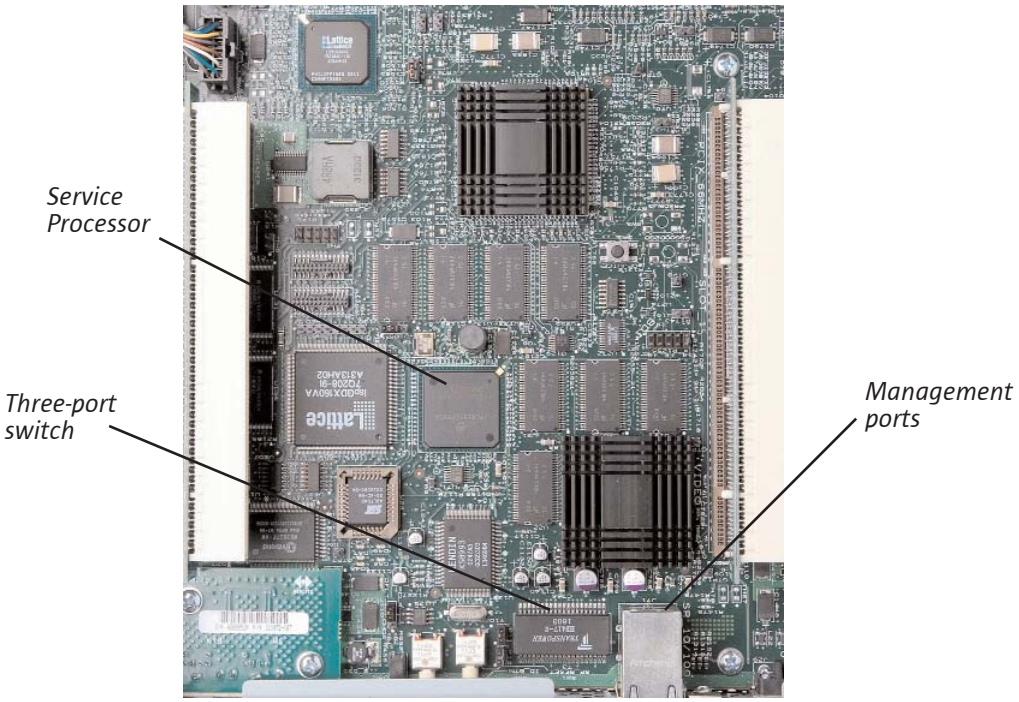
Figure 4-1. The Sun Fire V20z server includes a dedicated Service Processor

Service Processor Architecture

The Sun Fire V20z server includes a dedicated Service Processor for complete operating system independence and maximum availability of system management. The Service Processor enables:

- Environmental monitoring of the platform (such as temperature, voltage, fan speed, and panel switches)
- Alert messaging when problems occur
- Remote control of server operations (such as boot, shutdown, and reboot of the server's operating system, halting the server's boot process in BIOS, and upgrading BIOS)

Figure 4-2 shows the location of the Service Processor on the Sun Fire V20z server, as well as the three-port switch and management ports. The chipset, called Service Processor (SP), runs embedded Linux on an integrated Motorola MPC855T PowerPC processor.



Figures 4-2. The Service Processor, three-port switch, and management ports on the Sun Fire V20z server

SP Hardware

Service Processor specifications include:

- Dedicated on-board Motorola MPC855T integrated processor
- 64 MB RAM
- 16 MB NVRAM
- Kendin KS8993 integrated 3-Port 10/100 Base-T switch

The Service Processor is connected to the platform via multiple paths:

- A shared memory interface that provides the network channel for Ethernet connectivity between the Service Processor and the platform.
- A JTAG path that allows the Service Processor to diagnose the platform following a system crash.
- A serial path via the Super-IO Controller that permits the Service Processor to redirect the platform's serial port output to a user remotely logged in to the Service Processor.

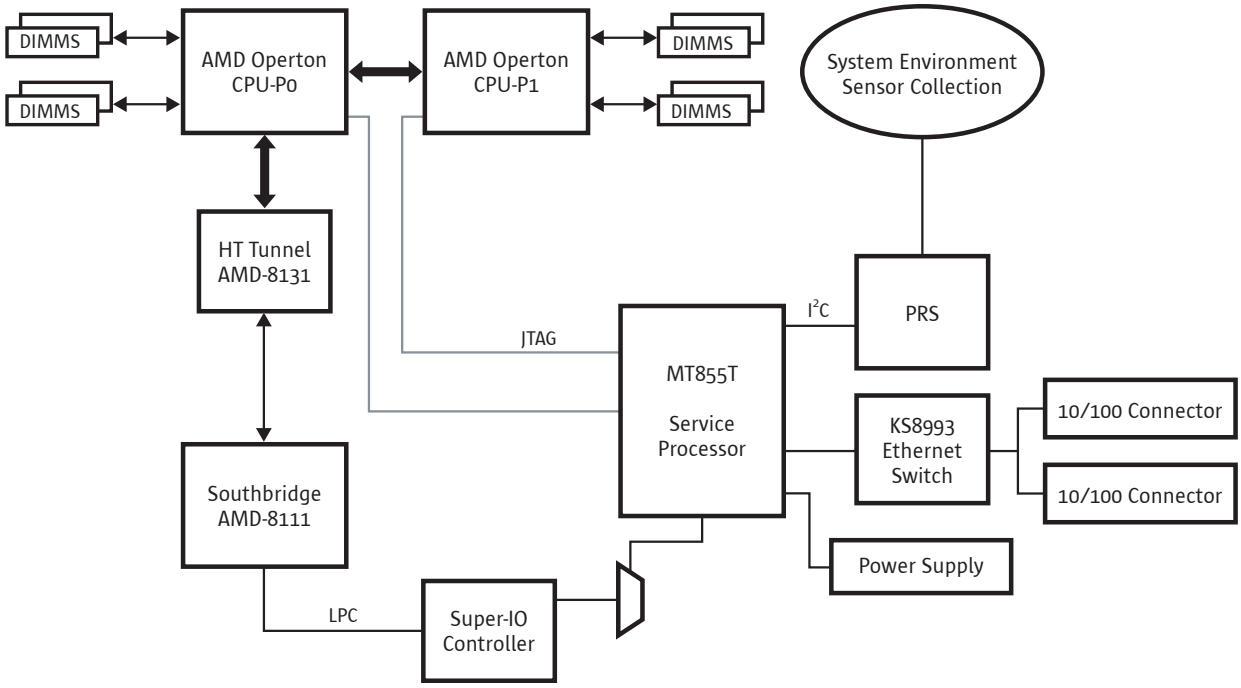


Figure 4-3. Service Processor connections

The Service Processor is connected to its Sensor Collection via I2C, and is powered via auxiliary power so that it is available even when the platform is powered-down. Access to the Service Processor is available via the serial port connector on the back of the box, or via its private LAN connection also on the back of the box.

The Service Processor's dedicated 10/100-Mbps Ethernet three-port layer-2 switch (Kendin KS8993) has one port internally connected to the MPC855T and two external ports to the rear panel of the Sun Fire V20z server. These two ports enable daisy-chaining using cross-over Ethernet cables.

Important things to consider when daisy-chaining Sun Fire V20z servers to a management network include:

- Performance degradation will occur if the cross-over cables are less than 1 meter long or if they are coiled.
- Redundant connections to the management network are strongly recommended (at both ends of the daisy-chain) since a hard power-off of a unit will break the chain.

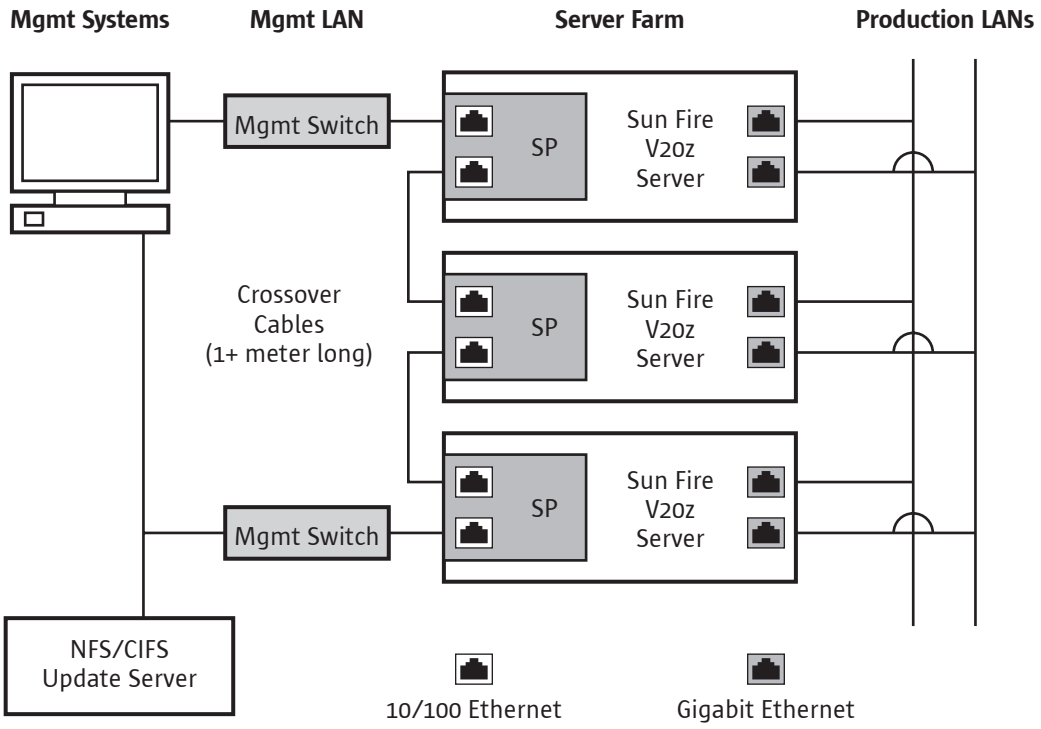


Figure 4-4. Sample daisy-chain configuration

SP Software

All system management functions are developed as standard Linux applications. The sole purpose of the Service Processor is to support system management. The flexibility offered by using Linux as the foundation of the Service Processor allows for the development and support of advanced system management features. The Service Processor —although technically a second, small computer—needs only a minimum of administration attention in order to fulfill its role.

Service Processor software includes:

- Embedded Linux
- IPMI BMC emulation
- Platform Control agent
- Diagnostics software
- LCD console control software

BMC

In order to perform autonomous platform management functions, the Service Processor runs embedded software or firmware. The Service Processor and its controlling firmware together are referred to as the Base Management Controller (BMC), which is the core of the IPMI structure. Tightly integrating an IPMI BMC and management software with platform firmware facilitates a total management solution.

The BMC is a service processor integrated into the motherboard design, providing a management solution independent of the main processor. The monitored server can communicate with the BMC via one of three defined interfaces which are based on a set of registers shared between the platform and the BMC.

The focus is on administering the platform under the Service Processor's control. As such, the full functionality of the Linux platform is not available in the Service Processor. Many familiar applications, such as ftp and telnet, are not provided as they are not required to support the system management feature set.

The Service Processor does include a command suite that enables system management and monitoring. Using these capabilities, administrators can individually execute commands or write shell scripts that remotely invoke these operations. For a multi-system environment in which configurations for all Service Processors must be synchronized, administrators can automate configuration changes using data-driven scripts. The system management command suite enables efficient management of each area of the system.

The Sun Fire V20z can be remotely managed using any of the following:

- Lights Out Management (LOM) through IPMItool
- Simple Network Management Protocol (SNMP)
- Sun Control Station (SCS)

Lights Out Management (LOM)

On the Sun Fire V20z server, Lights Out Management is performed through IPMItool, a command-line utility for controlling IPMI-enabled devices.

Intelligent Platform Management Interface (IPMI)

Platform management refers to the monitoring, logging, recovery, and inventory control features implemented in hardware and firmware. The key differentiator 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 Service Processor (or BMC) and System Management Software (SMS). The Service Processor is the brain behind platform management and its primary purpose is to provide autonomous sensor monitoring and event logging features. 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 SMS.

The Service Processor 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. This allows platform status to be obtained and recovery initiated under situations in which in-band delivery mechanisms are unavailable.

In modern systems, the Intelligent Platform Management Interface 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 v1.5 specification. IPMItool provides the ability to 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 but is also capable

of using the system interface as provided by a kernel device driver such as OpenIPMI. IPMItool is available under a BSD-compatible license.

System Management Software is generally complex and makes platform management only part of a much larger management picture. However, many system administrators and developers rely on command-line tools that can be scripted and systems that can be micro-managed.

IPMItool takes a different approach to SMS and provides a completely command-line oriented tool. Therefore, it 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. It 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

SNMP management provides remote access by SNMP-compliant entities 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.

Chapter 5

Motherboard Layout

Jumpers

Jumpers enable administrators to reconfigure the circuitry on a printed circuit board. When reconfiguring the system, administrators may need to change jumper settings on circuit boards or drives.

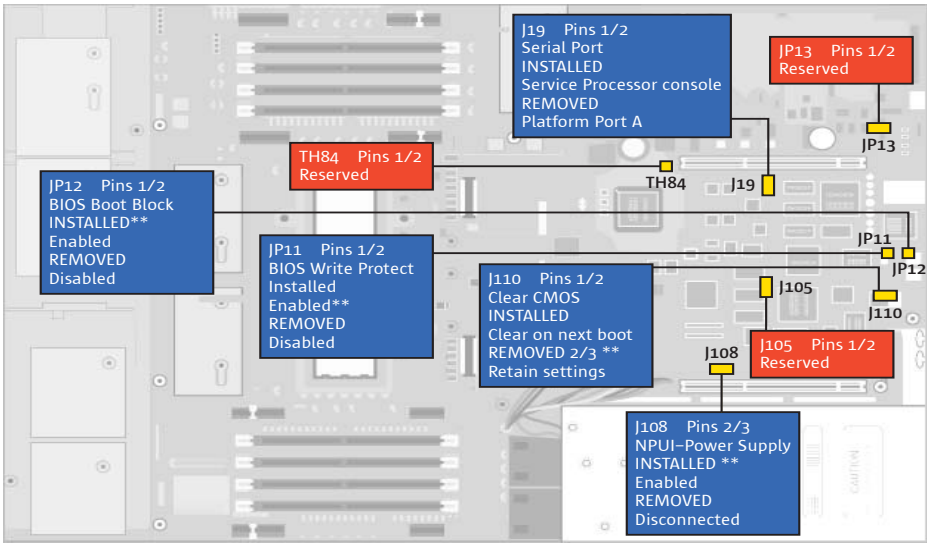


Figure 5-1. Sun Fire V20z jumper settings

Table 5-1. Sun Fire V20z jumper settings description

Jumper	Pins	Installed	Removed
J105	1/2	Internal use only	—
JP13	1/2	Internal use only	—
JP11	1/2	BIOS write protect disabled	BIOS write protect enabled
JP12	1/2	BIOS boot block protect enabled	BIOS boot block protect disabled
J19	1/2	Serial port is Service Processor console	Serial port is platform console
J108	2/3	Enable NPUI from AC power supply	Disconnect NPUI from AC power supply
J110	2/3	Clear CMOS on next boot	Retain CMOS settings
TH84	1/2	Internal use only	—

Platform Power Lights

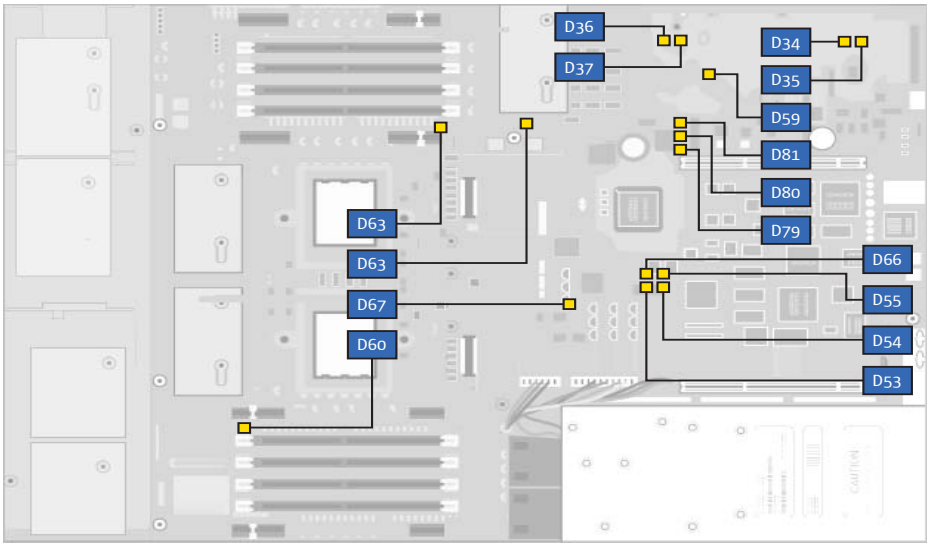


Figure 5-2. Sun Fire V20z platform power lights

Table 5-2. Sun Fire V20z platform power lights description

LED	DESCRIPTION	COLOR
D32	Fan in full on (error)	red
D34	GigaBit(0) 10/100 mode	yellow
D35	GigaBit(0) link activity	yellow
D36	GigaBit(1) 10/100 mode	yellow
D37	GigaBit(1) link activity	yellow
D53	Power Supply / Power OK	Green
D54	Power on indicator	Green
D55	Reset indicator	Yellow/off
D59	Thor rampower on	Green
D60	CPU 0 Power OK	Green
D63	CPU 1 Power OK	Green
D66	Thor Power Good	Green
D67	PRS Internal Error	Red
D79	SCSI Channel A Indicator	Green
D80	Not used	
D81	SCSI Controller Operational	Green

Service Processor Indicator Lights

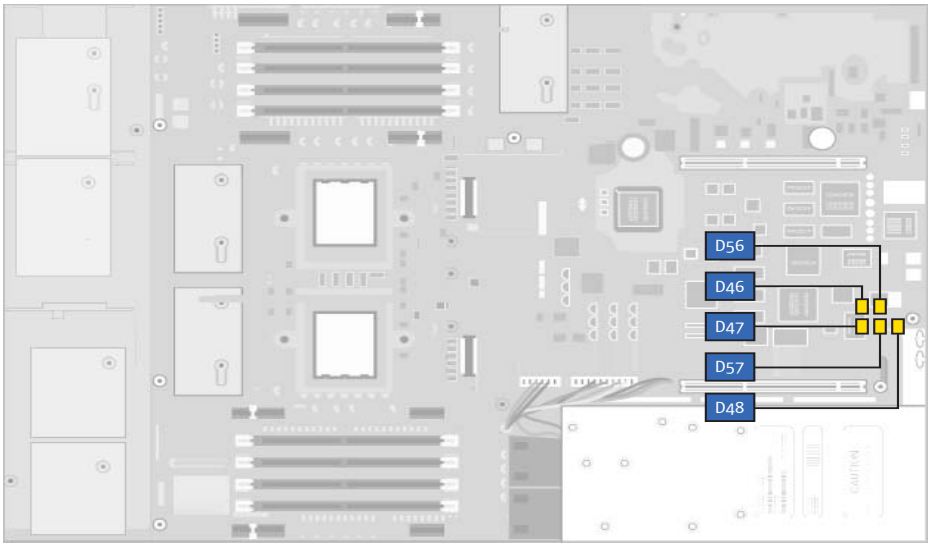


Figure 5-3. Sun Fire V20z service processor indicator lights

Table 5-3. Sun Fire V20z service processor indicator lights description

LED	Color		Comments
D48	Red	Off	Blinks 2X; Stays on
D47	Red	Off	Blinks 2X
D46	red	Off	Blinks 2X; Stays on
D56	Red	Off	Blinks 2X; Stays on if Enet cable installed in SP top
D57	Red	Off	Blinks 2X; Stays on if Enet cable installed in SP bottom

Activity is shown by D46 plus D56 and/or D57.

Switches

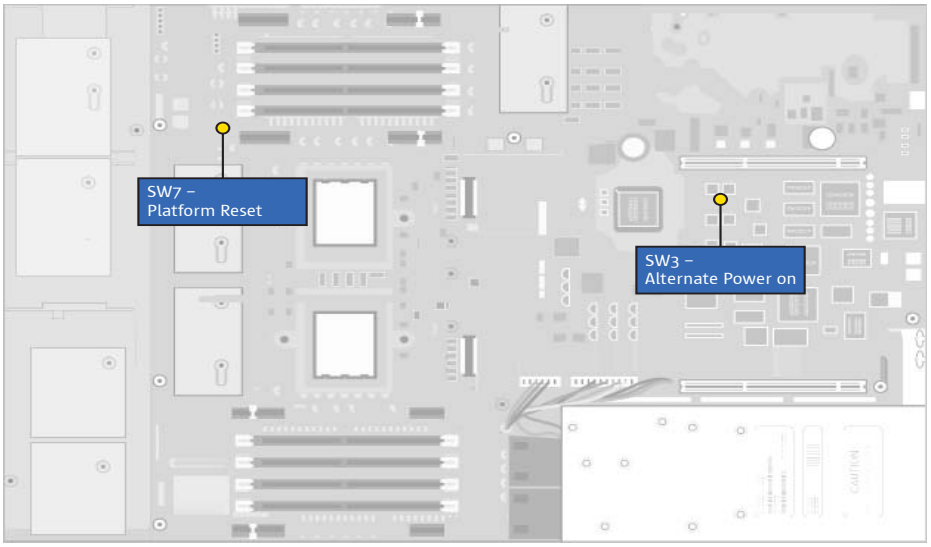


Figure 5-4. Sun Fire V20z switches

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN **Web** sun.com

We make the net work.

Sun Worldwide Sales Offices: Argentina +5411-4317-5600, Australia +61-2-9844-5000, Austria +43-1-60563-0, Belgium +32-2-704-8000, Brazil +55-11-5187-2100, Canada +905-477-6745, Chile +56-2-3724500, Colombia +571-629-2323 Commonwealth of Independent States +7-502-935-8411, Czech Republic +420-2-3300-9311, Denmark +45 4556 5000, Egypt +202-570-9442, Estonia +372-6-308-900, Finland +358-9-525-561, France +33-134-03-00-00, Germany +49-89-46008-0 Greece +30-1-618-8111, Hungary +36-1-489-8900, Iceland +354-563-3010, India–Bangalore +91-80-2298989/2295454; New Delhi +91-11-6106000; Mumbai +91-22-697-8111, Ireland +353-1-8055-666, Israel +972-9-9710500 Italy +39-02-641511, Japan +81-3-5717-5000, Kazakhstan +7-3272-466774, Korea +822-2193-5114, Latvia +371-750-3700, Lithuania +370-729-8468, Luxembourg +352-49 11 33 1, Malaysia +603-21161888, Mexico +52-5-258-6100 The Netherlands +00-31-33-45-15-000, New Zealand–Auckland +64-9-976-6800; Wellington +64-4-462-0780, Norway +47 23 36 96 00, People's Republic of China–Beijing +86-10-6803-5588; Chengdu +86-28-619-9333 Guangzhou +86-20-8755-5900; Shanghai +86-21-6466-1228; Hong Kong +852-2202-6688, Poland +48-22-8747800, Portugal +351-21-4134000, Russia +7-502-935-8411, Saudi Arabia +9661 273 4567, Singapore +65-6438-1888 Slovak Republic +421-2-4342-94-85, South Africa +27 11 256-6300, Spain +34-91-596-9900, Sweden +46-8-631-10-00, Switzerland–German 41-1-908-90-00; French 41-22-999-0444, Taiwan +886-2-8732-9933, Thailand +662-344-6888 Turkey +90-212-335-22-00, United Arab Emirates +9714-3366333, United Kingdom +44-1-276-20444, United States +1-800-555-9SUN or +1-650-960-1300, Venezuela +58-2-905-3800, or online at sun.com/store

Printed in USA 04/04