Sun™ Enterprise™ 10000 Server

Technical White Paper

Introduction 1

Sun Microsystems leads the world in providing the powerful symmetric multiprocessing servers that form the heart of today's enterprise networks. Building on this tradition of excellence, Sun invests its next generation of servers with even more performance, plus innovations in reliability, availability, and serviceability.

The Sun Enterprise™ 10000 server, known as Starfire™, is at the head of Sun Microsystem's Enterprise X000 family of servers, a comprehensive series of powerful, expandable, highly reliable servers for network computing. Like its companions in the series, Starfire is built around an uncompromised scalable symmetric multiprocessing (SMP) architecture. It offers the highest CPU, memory, and interconnect performance in its class yet is fully compatible with all applications and software tools written for the Enterprise series.

The Enterprise servers extend the performance of traditional network servers using key technologies such as UltraSPARC™ processors, a new interconnect architecture, hot-swap components (including processor boards) and a host of other features that improve reliability, availability, and serviceability.

With up to 64 processors, 64 Gbytes of memory, over 60 Tbytes of online disk, and the widest range of UNIX® application software, Starfire offers a straightforward path into client/server open systems. Importantly, it preserves existing investments in mainframe systems through a suite of connectivity and interoperability products.

Starfire provides scalable computing power for the vast range of existing applications for Solaris™ platforms, including the entire Enterprise X000 Series.

Starfire offers a cost-effective alternative to traditional mainframes for production data processing environments. It is an ideal application / data server for host-based or client / server applications like online transaction processing (OLTP), decision support systems (DSS), data warehousing, communications services, or multimedia services.

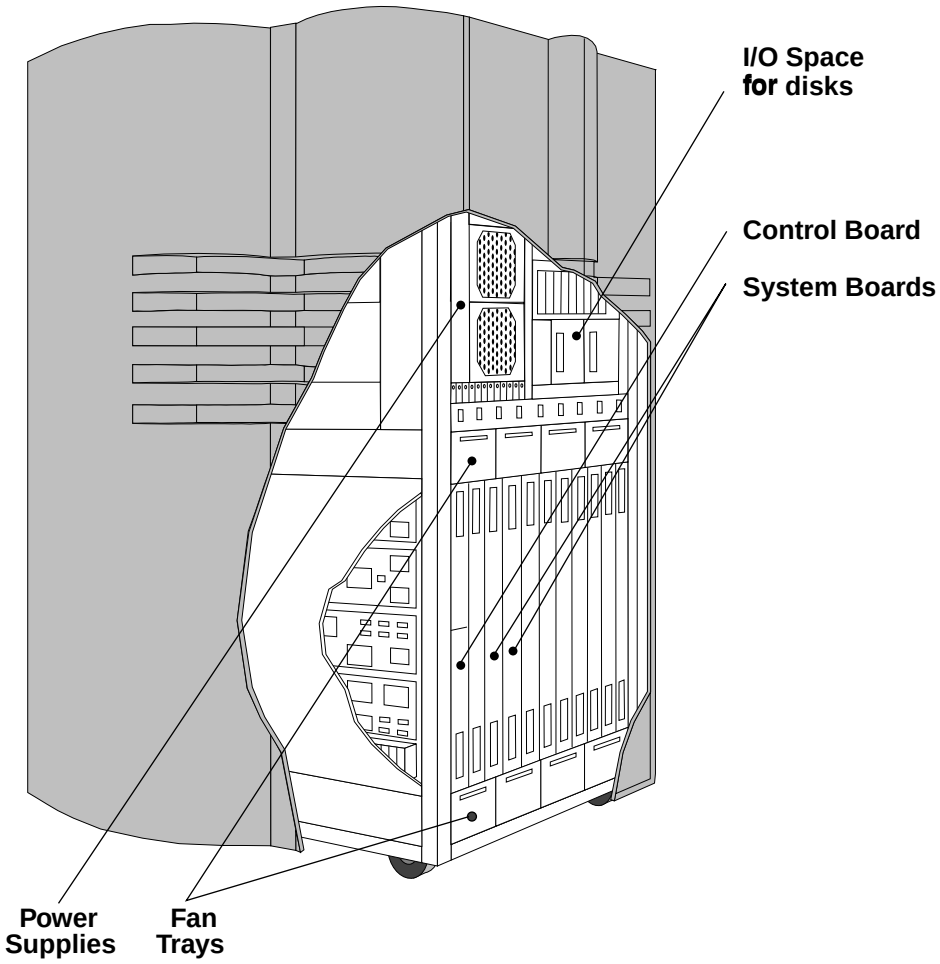

Figure 1-1 The Starfire Server

Starfire combines the power of system partitioning (known as Dynamic System Domains) and Gigaplane-XB™ interconnect technologies, plus the cost-effectiveness of industry-standard SPARC™-based hardware and software. It offers a wide range of high availability features and fast, responsive solutions for the most data-intensive applications.

The system provides a system bandwidth of up to 12.8 Gbytes/sec using Sun Microsystems' Gigaplane-XB coherent scalable crossbar interconnect at the core of the system. Starfire's UltraSPARC Port Architecture (UPA) works in conjunction with the crossbar interconnect to drastically reduce the number of cycles needed to transmit a standard data packet.

The computing power of the Starfire server is complemented by up to 64 I/O channels for high-speed communications. Its online storage capacity can exceed 60 Tbytes to support data warehousing, decision support systems, very large data bases, file serving, and other large scale data management applications.

In addition to a large system memory, each processor within the Starfire server utilizes an efficient fully coherent local cache to deliver optimal performance for a wide variety of applications and to support scalable multiprocessing in an SMP environment. Performance is further enhanced by a highly sophisticated virtual memory management system.

Scalability is a core feature of the Starfire server. The modular design of Starfire makes it easy to purchase a modest system and upgrade it as the demands for performance and capacity increase. There are no built-in barriers to expansion; the system can scale from minimum to maximum functionality in any axis without constraint. For example, the 4-processor entry-level configuration is easily expanded to the maximum capacity of 64 processors. Memory and I/O channels are equally flexible. All functional areas of Starfire are field upgradable, and most upgrades can be performed without halting the system.

The physical package of the Starfire server is a clean, modern single rack enclosure with access doors and detachable dress panels. The cabinet houses the system boards, the centerplane, the DC power supplies, and cooling fans. There is also room for more than 288 Gbytes of disk storage.

Each Starfire server has an Ethernet-based external System Service Processor (SSP), an external Sun™ workstation that is the system's administration console. It also monitors the Starfire server for problems and takes corrective actions.

1

Hardware Architecture 2

Starfire Server Features and Capabilities

- Internal frequency of up to 100 MHz with processors running at a clock frequency of 336 MHz. The system is designed to accept faster processors and interconnect components as future enhancements.
- High interconnect bandwidth: up to 12.8 Gbytes/sec.
- Fast floating point performance: up to 32 Gflops/sec peak.
- Cost-effective interconnect: a new generation of interconnect ASICs significantly reduces the cost of the system's interconnect architecture.
- High I/O bandwidth: up to 6.4 Gbytes/sec maximum aggregate SBus bandwidth. Starfire's individual SBuses can do 64-bit transfers, yielding a net data rate of 200 Mbytes/sec peak, 100 Mbytes/sec sustained per SBus. For PCI, the bandwidth of each bus is up to 528 Mbytes/sec.
- I/O flexibility: up to 32 independent SBuses or PCI busses in any mix.
- Error-correction interconnect: Data and address buses are protected by a combination of error correcting codes and parity.
- Starfire can be configured to automatically resume execution after most failures. This is achieved through a combination of redundancy and alternate pathing architecture.

- Dynamic System Domains (mainframe-like partitions): groups of system boards can be arranged in multi-processor Dynamic System Domains that can run independent copies of Solaris concurrently. Each domain is completely isolated from hardware or software errors that might occur in another domain.
- Dynamic reconfiguration: system components can be logically disabled or enabled during troubleshooting or prior to hot-swap operations.
- Hot-swapping: power supplies and most board-level system components —*including the processor boards*—can be exchanged while *hot*; that is, while the system is online.
- Scalable configurations: Memory and I/O slots can be added to Starfire without displacing processors.
- Service/maintenance process flexibility: the System Service Processor connects to Starfire via conventional Ethernet, permitting system administration from a remote location.

Independently Configured System Boards

The Starfire server houses a group of system boards interconnected by a centerplane. A single cabinet holds up to 16 of these system boards, each of which can be independently configured with processors, memory and I/O channels, as follows:

Up to four 336 MHz UltraSPARC microprocessor modules with supporting two level/4 Mbyte cache per module (64 per Starfire system).

Four memory banks with a capacity of up to 4 Gbytes per system board (64 Gbytes per Starfire server).

Two SBuses per board, each with slots for up to two adapters for networking and I/O (32 SBuses or 64 slots per system). Or two PCI busses per board - each accommodating one adapter. Starfire can have a mix of SBus and PCI adapters.

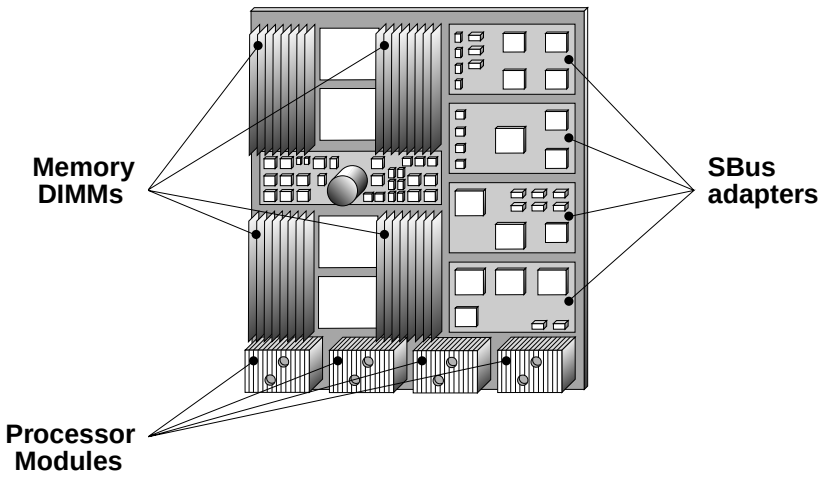

Figure 2-1 System Board

The Starfire server has a control board that contains the system-level logic central to all system boards. This includes the system clock generator, temperature and airflow monitoring, and an interface for the System Service Processor, which handles diagnostics, boot, shutdown and environmental monitoring. An optional, redundant control board is available should higher levels of system availability be required.

UltraSPARC™ Port Architecture and the Gigaplane-XB™ Interconnect

System architecture plays a dominant role in determining the performance, cost / performance, and scalability of today's high-end servers. Recognizing this, Sun Microsystems has incorporated an entirely new crossbar into the Starfire server: the Gigaplane-XB interconnect. This new feature, in combination with the UltraSPARC processors, enables the Starfire to provide the highest overall performance available in its class.

Physically, the Gigaplane-XB Interconnect is the centerplane of the server package. It is a circuit board with two symmetrical sides, each mounting up to eight system boards, a centerplane support board, and a control board.

The Starfire server interconnect design capitalizes on inherent UltraSPARC Port Architecture advantages. UPA is Sun Microsystems' standard I/O definition for UltraSPARC processors. The Gigaplane-XB interconnect uses a packet switched scheme with separate address and data paths. The segregation of address and data lines allows specific interconnect characteristics for each of the two entities. Data normally needs to be moved from point to point, (commonly from one system board to another), while addresses often must be distributed simultaneously throughout the system. Therefore the Gigaplane-XB interconnect handles data transactions with a fast crossbar interconnect, and address distribution with a broadcast router.

The Starfire server's UPA-mandated datapath width of 16 bytes halves the number of cycles needed to transmit a standard 64-bit data packet (when compared to previous architectures). The net outcome of this and all the innovations on the server centerplane is the highest interconnect bandwidth among open system servers.

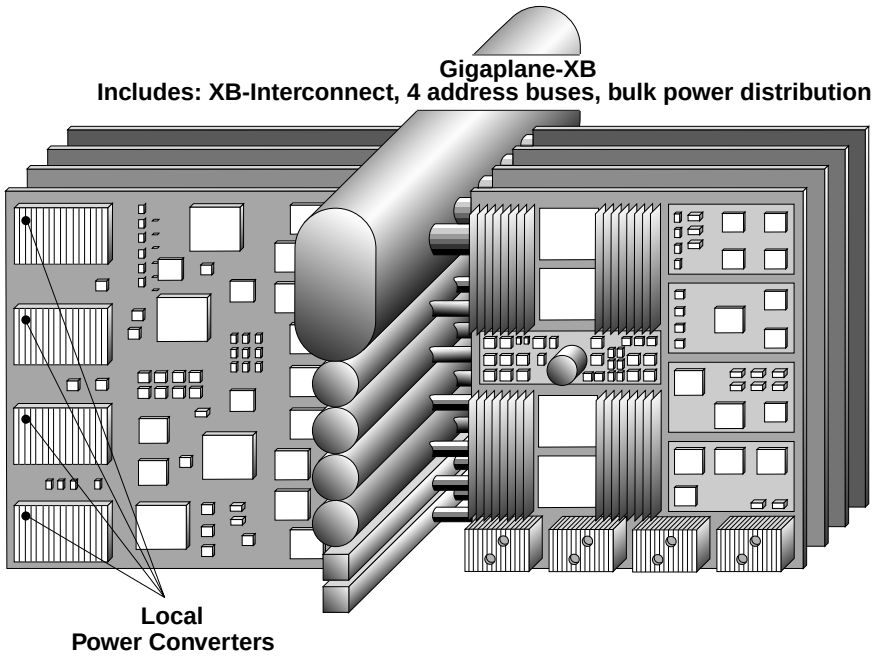

Figure 2-2 System Boards with the Gigaplane-XB Interconnect

Data and Address Routing

Data routing in the system is conducted at two levels: global and local.

The Global Data Router (GDR) is a 16 byte wide, 16 x 16 crossbar that steers data packets between the 16 system boards. From a hardware point of view, the GDR consists of 12 multiplexer ASICs situated on the centerplane board. Each ASIC is devoted to a 12-bit wide data slice. As a whole, the data is organized in 64-byte packets. The system requires four clock cycles to transmit each packet.

Data routing is carried out with a two-stage topology based on the Starfire's physical board partitioning. Local "many-to-one" routers on the system boards gather on-board requests and direct them to one port (per board). The global data crossbar connects the 16 system boards' ports together. With the 16 x 16 crossbar, any port can be connected to any other through the centerplane.

Starfire's address routing is implemented over a separate set of four global address buses, one for each of the four memory banks that can be configured on a system board. The buses are 48 bits wide including error correcting code bits. Each bus is independent, meaning that there can be four distinct address transfers simultaneously. An address transfer takes 2 clock cycles, equivalent to a snoop rate of 167 million addresses per second.

Scalable Symmetric Multiprocessing

Widespread use of more powerful workstations and networked personal computers is creating an increased burden on servers. An effective solution is symmetric multiprocessing that commits additional system resources as needed.

Starfire's system interconnect and cache coherency mechanism provide both the overall system performance and the throughput necessary to support a large number of processors, and to achieve near-linear scalability.

The Solaris™ operating system, too, scales proportionately with the other Starfire resources.

The combination of scalability, SMP, and SPARC/Solaris compatibility provides system performance, a large number of proven applications, and simplified capacity planning for new client/server applications or rehosted applications.

Mainframe-Style Partitioning via Dynamic System Domains

A Starfire can be subdivided into multiple computers, known as Dynamic System Domains, each consisting of one or more system boards. Each of these domains is a separate shared-memory SMP system that runs its own local copy of the Solaris operating system. And they do not have to be the same revisions of Solaris. The domains share the Gigaplane-XB interconnect but are isolated from each other. There can be up to eight domains.

The use of a single server with multiple domains is far more efficient than using many smaller, discrete servers. The multi-domain server is easier to manage and more flexible thanks to its integrated software hierarchy and its single-console operation. It also offers improved reliability, availability, and serviceability. And significantly, the multi-domain solution reduces the total cost of ownership: just one “infrastructure” (particularly the centerplane) supports many domains, equivalent to many discrete servers.

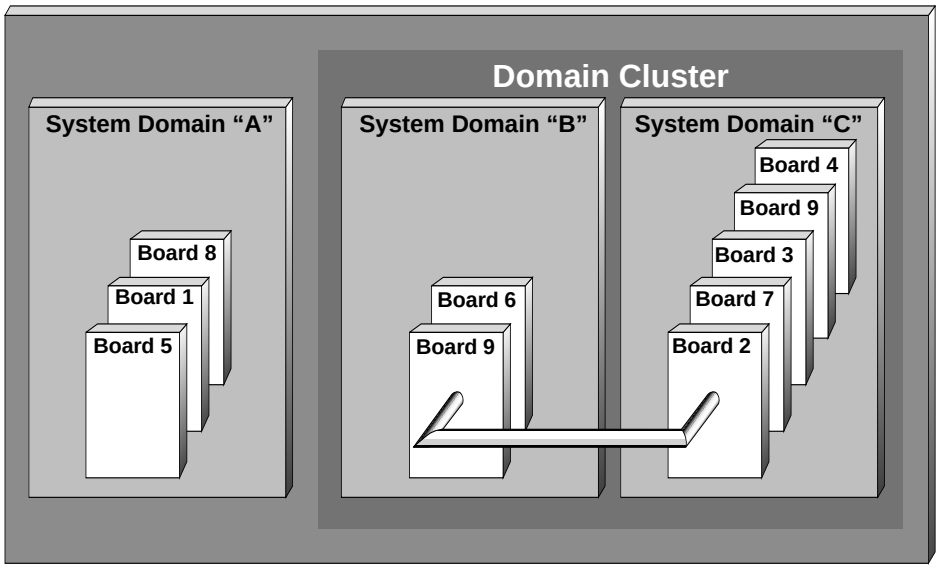

Figure 2-3 Dynamic System Domains

The Dynamic System Domain architecture supports a wealth of applications:

- Server consolidation: A single Starfire server can carry out the duties of several smaller servers. It preserves the independence and isolation of the discrete servers but is easier to administer and offers the flexibility to freely shift resources from one “server” to another. This is a benefit as applications grow, or when demand reaches peak levels requiring rapid deployment of additional computing resources.
- Simultaneous development, production, and test: These three functions can safely coexist within a single Starfire thereby eliminating the need for separate hardware for development and test purposes; instead, they are parcelled out to separate domains. The innate isolation of the domains enables development and test work to continue, while assuring that those efforts do not impact production.
- Software migration: Dynamic System Domains may be used as a means of migrating systems or applications software to updated versions. This applies to the Solaris operating system, database applications, new administrative environments, or any type of application.
- Special I/O or network functions: A domain may be established to deal with specific I/O devices or functions. For example, a high-end tape device could be attached to a dedicated domain, which is alternately merged into other domains which need to make use of the device for backup or other purposes.
- Departmental Systems: One partitioned Starfire server may be shared by multiple projects or departments, simplifying cost justification and cost accounting requirements.
- Configuring for resource requirements: Dynamic System Domains can be sized to match the needs of diverse applications. For example, projects with resource requirements that might overflow onto other applications can be isolated to their own domain. Conversely, for applications that can’t take advantage of all resources (i.e., those lacking scalability), multiple copies of the application can be run in separate domains.
- Data Warehousing: The normal configuration is to have a large data warehousing server, then a second tier of “data marts” on smaller servers, and finally the desktop tier. Using Dynamic System Domains, one Starfire with appropriately-allocated domains can be used as both the warehousing server *and* in place of all the discrete servers on the second tier. This adds flexibility in the administration of the data hierarchy.

2

UltraSPARC Processors/Processor Modules

Starfire can include up to 64 processors situated in groups of up to four per system board. The UltraSPARC processor itself is a superscalar microprocessor chip that can execute four instructions per clock cycle. Within the chip, Fetch and Prefetch units ensure that data for the chip's Floating Point and Integer units is always available without delay. A Data Cache holds results from floating point and integer calculations. Other on-chip elements include a Memory Management unit, a Load/Store unit, a Memory Interface unit, and a local Instruction Cache to store instructions prior to execution.

The processor mounts on a small daughter board, the UltraSPARC Module, which also houses the 4-MB Second-Level Cache and the UltraSPARC Data Buffer (UDB) circuitry. The Second Level Cache handles cache misses from the processor's on-chip cache memories. In total, the architectural elements on the processor chip and the module support Starfire's ability to execute two floating point instructions, add or subtract, and two integer instructions during a single clock cycle.

Because processor technology continues to advance more rapidly than other system technologies, Starfire is designed to be easily upgraded in the field by adding or replacing discrete processor modules.

Power Supply Redundancy and Hot Swap

Because power fluctuations are among the most common causes of unscheduled service interruptions, the Starfire server is designed for resiliency against such fluctuations. The system is configured to interface to standard uninterruptable power systems (UPS) that protect against line voltage variations and dropouts. Secondly, a non-maskable powerfail interrupt is issued when internal voltages exceed pre-defined limits. As a final measure, Starfire's own internal power supplies resist power variations.

Starfires are outfitted with multiple modular bulk power supplies that convert the incoming 220-volt AC line voltage to +48 volts DC. The aggregate capacity of these bulk supplies offers plenty of headroom over and above the maximum power demands from the system; for example, a maximum (64 processor) Starfire is equipped with eight supplies but can run on six. If a supply fails, then the remaining supplies pick up the load without affecting system operation. This redundancy is common to all Starfire configurations. The modular power supplies can be hot-swapped while the system is running.

Each side of the centerplane has its own 48V distribution bus, and each system board develops its own low-voltage supplies locally with on-board regulators. Should a regulator fail, the system adapts automatically by reconfiguring itself to exclude the offending board.

The system's AC wiring provides further safeguards against outright failure. The maximum Starfire server uses four line cords, each fed by a separate 220V single-phase 50/60 Hz AC circuit, to deliver the required input power to the bulk DC supplies. Should one cord fail, system operation is not affected. A fifth discrete line cord serves the system's I/O space. This level of redundancy ensures against a system-wide power loss and also reduces the current through any one circuit.

Memory System

Large-scale systems must provide sufficient memory capacity to sustain high performance from the processors and I/O channels. Additionally, memory must be quickly accessible in order to avoid interfering with other subsystem activities. Finally, the large concentration of data in today's data center systems and production environments necessitates a highly reliable design.

Starfire is designed to meet all of these requirements. By using 16 Mbit DRAM chips, a fully-configured system offers up to 64 Gbytes of system memory. The Solaris operating system provides scalability consistent with this memory capacity. Efficient memory interleaving provides the necessary access speeds. The entire memory data path is protected by error correcting code (ECC) mechanisms, and SIMM organization is specifically designed such that each DRAM chip contributes only one bit to a byte of data. Thus the failure of a DRAM chip causes only correctable single-bit errors.

Main memory is configured in multiple logical units. All units—memory, processors, and I/O buses, are equidistant in the architecture and all memory modules comprise a single global shared memory. Each memory unit consists of four banks of physical memory connected to the memory data buffer and the local data router to the crossbar switch (the global data router). Memory addresses are assigned through software at boot time.

A memory module can do a cache-line read or a write every four system clocks (48 ns), a bandwidth of 1.3 Gbytes/sec. This matches both the Global Data Router port bandwidth, and that of the processor's UPA port.

2

Input/Output Modules and SBus

The Starfire I/O subsystem employs the standard SBus (IEEE 1496-1993) which has been the mainstay of Sun Microsystems' architecture for years and is supported by several system manufacturers and many peripherals vendors. SBus provides high throughput and minimum latency using optimized block transfer modes, a non-blocking I/O cache, and a low-contention design. It provides a 64-bit data path and I/O drivers are adapted to support the Starfire dynamic reconfiguration capability.

The Starfire I/O module is a mezzanine card that plugs into the system board and connects the UPA to a pair of SBuses. Each of these SBuses can in turn be populated with one or two single-width SBus adapters, or one double-width adapter.

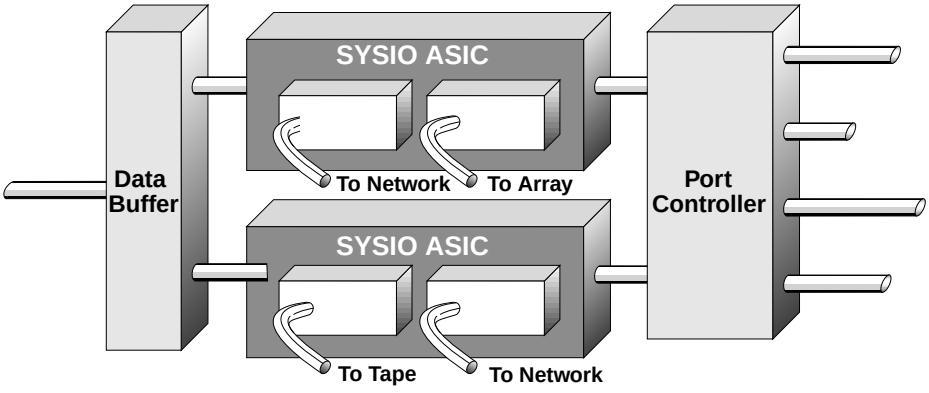

Figure 2-4 I/O Module

This allows Starfire to offer incremental expandable I/O by configuring up to 64 SBus slots on 32 independent buses. Access latency for every SBus is the same. The SBus operates at 25 MHz and can deliver 100 Mbytes/sec sustained throughput with capable peripherals. This is as much bandwidth as some alternate systems provide for multiprocessor CPU-to-memory traffic. The system's overall I/O capacity approaches system memory bandwidth, rather than just peripheral bus bandwidth as in most other comparable systems.

All SBuses are connected through buffers to their respective local data routers and ultimately to the Global Data Router. The heart of the I/O module is the SYSIO ASIC, which oversees transactions between the Starfire data routers and the SBus adapters.

By using virtual (rather than physical) addresses in its I/O scheme, the operating system gains a 64 Mbyte space for efficient management of I/O operations. Each SBus interface includes its own Memory Management Unit to translate between virtual and physical addresses.

The SBus provides three transfer modes: programmed I/O, consistent mode direct virtual memory access (DVMA), and stream mode DVMA. While programmed I/O requires direct CPU intervention, DVMA permits large amounts of data to be transferred between the SBus and memory without using the processor.

Network and Peripheral Connectivity With SBus

The following adapters and peripherals are supported by Starfire:

- SCSI adapter: A single-width SBus card with one fast-wide differential 16-bit SCSI-2 bus. This SCSI bus is used on Starfire for connections to discrete disks or tape devices. The raw data transfer rate on the bus is 20 Mbytes/sec.
- UltraSCSI: For connection to Sun's hardware RAID subsystem, the StorEdge A3500 disk array.
- Fiber Channel Arbitrated Loop (fc-al): for connection to the StorEdge A5000 disk array. The data transfer rate is 100 Mbytes/sec.
- Quad fast Ethernet adapter: This SBus adapter has four 10/100base-T ports. In conjunction with higher level protocols, it supports industry-standard twisted pair networking between Starfire, workstations, and other systems. Raw data rate is 10 or 100 Mbits/sec with auto data rate negotiation.
- Gigabit Ethernet adapter: to allow Starfire to participate in 1000 Mbits/sec backbone networks.
- FDDI: A networking medium used as an alternate to 100 Mbits/sec Ethernet. Raw data transfer rate is also 100 Mbits/sec but over fiber cabling. There are two FDDI options available with the E10000: single attach (SAS) and dual attach (DAS).
- ATM: This adapter allows the Starfire to interface to Asynchronous Transfer Mode local or wide area networks over fiber cable. Supported network speeds are 155 and 622 Mbits/sec.

- Token Ring: This adapter provides 32-bit data transfers at 4 or 16 Mbits / sec. For connection to mainframe environments.
- SunLink ISDN: For connection to digital telephone networks.
- High Speed Serial Interface: For WAN connections up to T1 speeds.

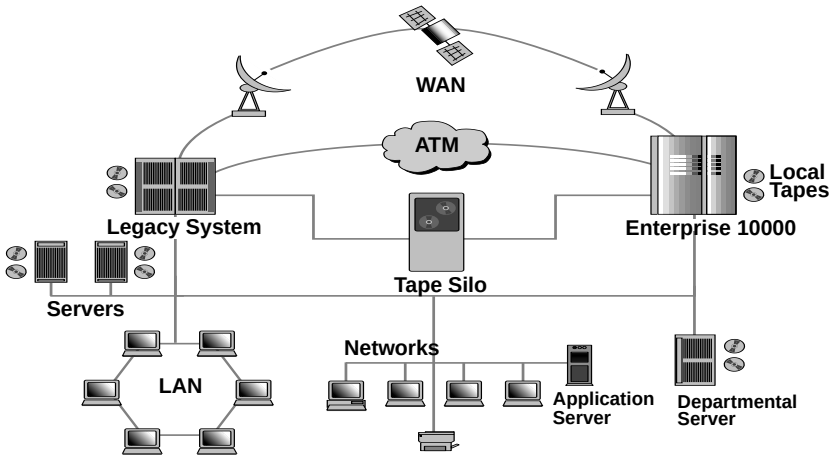

Figure 2-5 Networking

A wide range of peripheral options is available for Starfire. These currently include:

- Discrete disks (3.5" devices with 4.2, 9.1 or 18 Gbyte capacities). Drives are installed in hot-swap trays for easy replacement in the event of failure. Trays are outfitted with redundant power and cooling.
- The StorEdge A3500 array, a high availability disk subsystem with hardware RAID. Storage up to 1 Tbyte per cabinet using 18 Gbyte disks.
- StorEdge A5000 fc-al disk array. Storage capacity is 500 Gbytes per cabinet.
- StorEdge A7000 Intelligent Storage Server. Allows datasharing with mainframes and storage capacities up to 2 TBytes.
- Tape devices, including Exabyte 8mm and digital linear tape (DLT) with autoloaders and libraries.

In addition to these currently available storage options, Starfire will support all future, high-performance Sun peripheral solutions.

The software complement to all these storage options is Sun StorEdge Volume Manager, a tool that improves disk manageability, reliability, and performance. It supports disk striping (using multiple disk adapters to achieve higher I/O rates), disk mirroring (duplicating data on multiple disks and disk adapters to protect against disk or adapter failures), RAID-5 (for economical disk redundancy), and concatenation of files across logical disks to support large file systems.

PCI I/O on Starfire

PCI is an open I/O standard to which Sun products are moving. The main advantage of using PCI, in a server application, is PCI's higher speed compared to SBus. PCI adapters are available with 32 or 64 bit data paths and run at a clock frequency or 33 MHz or 66 MHz.

System Boards for Starfire are available with a PCI module in place of the standard SBus module. This PCI module has two 66 MHz busses and each can accommodate one PCI adapter.

Because it is only possible to package two PCI adapters per system board (in contrast to four SBus adapters), PCI is not cost effective where there is not a performance requirement. Therefore Starfire will remain basically as an SBus-based system with PCI available for selected uses. For instance, customers will prefer to use SBus for the interfaces detailed in the previous section (e.g. SCSI, Ethernet, fiber channel, FDDI, ATM, etc.). PCI can be used for fast data transfer situations such Gigabit Ethernet - also available as an SBus adapter. PCI, as a future capability, will be used for HIPPI (a high speed interface used by technical computing applications) and Sun Cluster interconnect.

System Service Processor (SSP)

The console functions for the Starfire server are carried out by the System Service Processor. The SSP is a Solaris workstation especially configured to support Starfire. It runs a suite of diagnostic and management programs, including POST and Hostview. These and other SSP-controlled programs are designed to minimize the need for active management by an operator while aiding that operator in supervising the system.

The SSP sequences the boot process, assists in configuration decisions at bootup, monitors environmental conditions, helps Starfire recover from interrupts, and optionally sends information to Sun Enterprise Services regarding interrupts. In addition, the SSP is the management station for dynamic reconfiguration activities (including hot swap) and Dynamic System Domains.

The SSP is the access port for administrators and service providers and provides intelligence separate from the host Starfire server. The SSP is connected via an Ethernet network. Using the Network Console feature, virtually any workstation in the network can open a host console session. This provides several important capabilities, including remote administration of Starfire and remote power control. Starfire and the SSP use Network Time Protocol (NTP) to maintain time synchronization.

For redundancy, Starfire can be provided with a second SSP that acts as a hot spare.

Power On Self Test (POST)

POST is the principal bring-up diagnostic program for Starfire It verifies that all components (memory, processors, I/O adapters, etc.) are functional before allowing them to be logically configured into the server. Using POST ensures that the system can be brought up with confidence, even if there are hardware problems, and that those problems will not impact the availability of the system because they are configured out of the system.

POST begins by testing each component individually, and then tests the system as components are configured in. Based on the results of its testing and parameters entered by the systems administrator, POST creates a list of viable configurations and selects which of these configurations is to be booted.

The amount of testing done by POST is controlled by the “level” of testing requested. At higher levels, more tests are run, and POST thoroughly tests additional components—even those with little chance of failure. Levels range from “1” (basic initialization and minimal testing) to “17” (the minimum diagnostic level) to “128” (the maximum diagnostic level, including many components with little chance of failure).

Hostview

Hostview is the graphical interface program that monitors system components and functions and interfaces with service and operator functions. Hostview simplifies monitoring and control of Starfire by providing a series of easy-to-follow menus.

Hostview also provides for control of dynamic reconfiguration functions, including:

- Create independent Dynamic System Domains
- Dynamically detach system components for diagnosis or as a Dynamic System Domain
- Hot swap/hot plug
- Dynamically attach system components
- Alternate pathing setup and control of disks or networks

Hostview is able to make reasonable decisions about environmental problems, making Starfire less susceptible to single-point failures. It enables the system administrator to monitor and respond to various system functions, such as:

- Hardware faults, e.g., over-voltage
- Environmental faults, e.g., excessive heating
- Loss of *heartbeat* from Starfire

Dynamic Reconfiguration and System Board Hot Swap

Dynamic reconfiguration provides maximum flexibility for system configuration, upgrades, and quick recovery from failures. It also is essential to Starfire's hot-swap processes, during which it is used to logically deactivate affected components prior to removal. This capability is available to the system administrator during normal system operation, and to the system boot process for automatic logical removal of an unavailable or failed component. Moreover, dynamic reconfiguration can be driven remotely, via network connection to the System Service Processor. This is an important resource that helps service providers efficiently carry out maintenance, repairs, and updates.

At boot time, a comprehensive diagnostic software system determines which units in the server are available and functioning. This software is collectively known as the Power-On Self Test (POST), although it is also run when recovering from errors as well as during each reboot. POST runs in several phases under the control of the SSP.

The POST software system enables automatic system recovery after a panic interrupt (UNIX system crash). To permit continued operation, the system immediately reboots and reconfigures itself. POST detects and isolates the failed component and logically disables it in anticipation of later removal.

There are many possible combinations to use in case of failures. A series of heuristics using site-specified parameters determines which configuration is to be used.

As part of the dynamic reconfiguration feature, Starfire allows a system board to be inserted or removed without powering off the system. This hot-swap procedure is equally applicable to servicing failed boards or to upgrading memory, I/O, or processor configurations. Only the immediate system board and its components are affected.

Following the logical detachment, the designated system board is powered off and removed for service. An upgraded or replacement board is then inserted and powered on. The SSP runs the POST diagnostics on the new board, after which the system board is logically reattached to the operating system. During this period, the system continues to operate and user applications are unaffected.

SNMP Support and the Management Information Base (MIB)

All events noted by Starfire's monitoring functions are broadcast to the network via SNMP messages. This means that, in addition to the hostview program, they may be trapped by any software package that processes SNMP messages. Thus, messages specific to Starfire servers can be integrated conveniently into a site's existing systems and network management toolsets.

A Management Information Base has been defined for the Starfire server family. This includes a wealth of information that can be used by third party system and network management products.

Network Console (NetCon)

Network Console is the function that enables remote logins to the Starfire System Service Processor. A system administrator can use Network Console to log in and operate Starfire remotely, from anywhere there is a network connection, via an X-Window interface.

The administrator or service providers may attach to an SSP session, perform diagnostic work, and conclude their own work, without disturbing the console session. Also, the system administrator can observe a service provider's commands, in order to understand what diagnostic and corrective steps are being taken.

Starfire RAS Capabilities

Reliability, availability, and serviceability features are essential to the Starfire server. These features encompass RAS characteristics in three categories:

- Reliability features prevent failures from happening
- Availability features keep the system operating in spite of failures

- Serviceability features help correct failures quickly after they have occurred

A host of strategies has been applied to ensure that data is protected under all circumstances, and that recovery from any fault can occur with little or no interruption of system availability. Redundant buses, a programmable configurable address space, full ECC or parity data checking on data storage and data paths, and extensive built-in monitoring equipment all permit Starfire to detect, avoid, and recover from most failures.

Following is an overview of some key RAS characteristics of the Starfire server:

- The system detects data integrity problems and corrects them when it can. For example, a failed DRAM in a memory module manifests itself as a single failed bit in each of four different data words—each of which is correctable by error-correcting code bits in those words.
- The system can reconfigure automatically around a hard failure, resuming operation without the failed component after an automatic reboot.
- Redundant components can be added to augment the system's already high reliability and availability. All components in the system may configured redundantly if the customer so desires.
- Most hardware maintenance can be performed via hot swap without powering down the system; only the components actually being worked on are taken out of service.
- Multiple operating system versions and/or diagnostics can be co-hosted by the hardware using several independent Dynamic System Domains.
- Starfire is fully compatible with commercially available Uninterruptable Power Systems (UPSs).
- Power and cooling is fault tolerant in that any single failure does not cause a system interruption.
- Connectors, cables, guides, etc. within the system are designed to withstand thousands of insertion/removal cycles.
- Power On Self Test software controlled from the System Service Processor tests logic blocks first in isolation, then in combination with others to locate any failing components, which are then removed from the active system configuration.

- The Hostview program on the SSP enables administrators/operators to monitor the state of the system and take action, if needed, beyond the normal automated measures. SNMP messaging enables the Starfire to be managed by commercially available system management tools.

Redundancy options are available for disk mirroring, redundant adapters, and dual ported devices. In addition, the system supports alternate path connectivity to peripherals and networks from two system boards; this allows continued access in the event of a system or adapter board failure. These options increase availability for mission-critical applications and meet high reliability requirements with no single point of hard failures.

Starfire Software Architecture 4

System Software

Starfire's software environment is SunSoft's Solaris 2.6 operating system. Designed for enterprise-level integration, Solaris delivers a competitive advantage to business through networked computing, scalability, and multi-architecture support. It offers higher throughput, higher quality, and more productivity in applications that demand mission-critical reliability.

Solaris 2.6 is Sun's symmetric multiprocessing implementation of UNIX System V Release 4, and conforms to all generally-accepted industry standards (such as the SVR4 interface definition, X/Open®, XPG4, and POSIX 1003.1). Solaris 2.x is of course backward-compatible with existing Solaris 1.x applications.

The Solaris operating system takes advantage of Starfire's scalability and ensures that all resources can be fully utilized. It is enhanced regularly to further exploit that scalability and to better support mission critical applications and data center operations.

The Solaris kernel is fully multithreaded for maximum efficiency in taking advantage of all system resources. Multithreading allows system and user applications to be structured as several independent computational threads rather than as a single thread of control. Multiple independent threads execute more efficiently because they can use multiple processes at the same time or move among available processors.

The system operating environment includes the Common Desktop Environment (CDE) graphical user interface. With Solaris CDE, mission-critical client-server applications and information are easily accessible across the enterprise. Solaris CDE provides interoperability and consistency across popular enterprise hardware platforms and operating systems. It is compatible with industry standards such as Motif, X11R6, MIME and IMAP4. The Solaris CDE desktop includes productivity tools such as MIME-compatible email, workgroup calendaring, an image viewer and file and print managers for individual users.

The cost of managing a geographically distributed networked facility is reduced by a group of powerful automation and administration facilities known as Solstice AdminSuite. These tools enable efficient systems management with full network security. Among other capabilities, AdminSuite expedites addition of user accounts, connection of clients to the network, and more.

The Solaris operating environment provides a sophisticated security system that controls the way users access files, protect system databases and use system resources. Solaris security is network-wide security, providing security over several different systems, not just one.

Solaris runs on SPARC-based systems and Intel X86 architectures. This provides a seamless path for expansion from, and seamless interoperability between these platforms.

Software Development

Starfire supports a range of standards-based computer aided software engineering (CASE) applications, fourth generation languages (4GL), and standard compilers, libraries and tools developed by Sun Microsystems and a large number of independent software vendors.

Other high-performance software tools are available for developers, including Sun™ WorkShop™, compilers for various languages (Fortran, C, C++, Ada, Cobol, and Pascal), and intuitive graphical tools for creating, optimizing, and debugging parallel and multithreaded programs.

Application Software

Starfire features the industry's broadest independent software vendor (ISV) applications offering. Thousands of commercial, off-the-shelf solutions currently are available for Starfire's Solaris environment. This extensive third-party market allows Starfire server users to benefit from the specialized expertise of hundreds of vendors of hardware and software products for commercial and technical applications and requirements.

Meeting the Requirements of Production Business Environments

Starfire supports a full suite of hardware and software solutions to address the requirements of commercial business computing. These include RDBMS packages for online transaction processing (OLTP), decision support systems (DSS); transaction processing monitors; visualization and multimedia; CASE and 4GL languages and tools; software for report generation; tools for system, tape and batch management; and software for connectivity to and interoperability with mainframe and desktop systems.

With the increasing availability of commercial system management software for UNIX environments, Starfire offers a compelling solution for a cost-effective and reliable enterprise-class server. And, as organizations prepare to move client/server systems from pilot tests to full production environments, Starfire provides the scalability to support the most demanding requirements of large databases and other server-based applications in DSS, OLTP, and data warehousing.

Summary 5

Sun Microsystems provides a unique and complete solution for today's data centers. The Enterprise 10000 (Starfire) system brings together a number of desired attributes to offer a complete solution for commercial data processing:

- Breadth of application solutions: The industry's largest base of UNIX applications and system and network management software is available on Starfire.
- Dependability: Starfire incorporates a wide range of reliability, availability, and serviceability features for a fault resilient and dependable solution.
- Affordability: Leveraging large-volume hardware and software components, maximizing availability, and minimizing service costs result in a total solution that is cost effective both in its components and as a whole. Sun Microsystems' aggressive pricing for processors, memory, peripherals, system software, and services set a new standard for cost effective enterprise-class servers.
- Scalable symmetric multiprocessing: Starfire offers scalability and flexibility unparalled by any other SMP architecture in its class.
- Performance: With up to 64 processors, 64 Gbytes of memory, 60 TBytes of online disk, and the Solaris operating system, Starfire provides scalable system performance, a large number of proven applications, and simplified capacity planning for new client/server applications or rehosted applications.
- Alliance: Sun Microsystems' unique understanding of network, data center and mission critical applications enables it to be a long term ally for your organization.

Sun Microsystems Computer Company
A Sun Microsystems, Inc. Business
2550 Garcia Avenue
Mountain View, CA 94043 USA
415 960-1300
FAX 415 969-9131

Sales Offices

Argentina: +54-1-317-5600
Australia: +61-2-9844-5000
Austria: +43-1-60563-0
Belgium: +32-2-716 79 11
Brazil: +55-11-5181-8988
Canada: +905-477-6745
Chile: +56-2-638-6364
Colombia: +571-629-2323
Commonwealth of Independent States:
+7-502-935-8411
Czech Republic: +420-2-33 00 93 11
Denmark: +45 4556 5000
Estonia: +372-6-308-900
Finland: +358-9-525-561
France: +33-01-30-67-50-00
Germany: +49-89-46008-0
Greece: +30-1-6188130
Hungary: +36-1-202-4415
Iceland: +354-563-3010
India: +91-80-5599595
Ireland: +353-1-8055-666
Israel: +972-9-9513465
Italy: +39-039-60551
Japan: +81-3-5717-5000
Kazakhstan: +7-3272-466774
Korea: +822-3469-0114
Latvia: +371-755-11-33
Lithuania: +370-729-8468
Luxembourg: +352-49 11 33 1
Malaysia: +603-264-9988
Mexico: +52-5-258-6100
The Netherlands: +31-33-450-1234
New Zealand: +64-4-499-2344
Norway: +47-2218-5800

People's Republic of China:
Beijing: +86-10-6803-5588
Chengdu: +86-28-678-0121
Guangzhou: +86-20-8777-9913
Shanghai: +86-21-6466-1228
Hong Kong: +852-2802-4188
Poland: +48-22-8747800
Portugal: +351-1-412-7710
Russia: +7-502-935-8411
Singapore: +65-438-1888
Slovak Republic: +421-7-522 94 85
South Africa: +2711-805-4305
Spain: +34-91-596-9900
Sweden: +46-8-623-90-00
Switzerland: +41-1-825-7111
Taiwan: +886-2-2514-0567
Thailand: +662-636-1555
Turkey: +90-212-236 3300
United Arab Emirates: +971-4-366-333
United Kingdom: +44-1-276-20444
United States: +1-800-821-4643
Venezuela: +58-2-905-3800
Worldwide Headquarters:
Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900 USA
Phone: 650-960-1300
Fax: 650-969-9131
Internet: www.sun.com

Printed in USA