ORACLE
SPARC SERVERS

Oracle's SPARC T7 and SPARC M7 Server Architecture

Software in Silicon: Enabling Secure Clouds for the Real-Time Enterprise

ORACLE WHITE PAPER | OCTOBER 2015

ORACLE

Table of Contents

Introduction

Modern technology initiatives are driving IT infrastructure in a new direction. Big data, social business, mobile applications, cloud, and real-time analytics all require forward-thinking solutions and enough compute power to deliver the performance required in a rapidly evolving digital marketplace. Customers increasingly drive the speed of business, and organizations need to engage with customers on their terms. The need to manage sensitive information with high levels of security as well as capture, analyze, and act upon massive volumes of data every hour of every day has become critical.

These challenges will dramatically change the way that IT systems are designed, funded, and run compared to the past few decades. Data security can no longer be treated as an afterthought, because billions of dollars are lost each year to computer intrusions. The massive explosion in data volume, variety, and velocity increases the need for secure and effective analytics so that organizations can make better and quicker decisions. Complex IT infrastructure poses an impediment by becoming more difficult and expensive to maintain at precisely the moment that organizations are under pressure to drive down costs, increase operating efficiencies, and deliver innovative technologies that can generate new revenue streams.

Oracle's new SPARC M7 processor–based servers take Oracle's server technology to new levels by offering the world's first implementation of Oracle's Software in Silicon technology to build clouds with the most secure platforms in the world. Offering both database and application security and acceleration, these servers offer silicon secured memory, in-memory query acceleration, data compression and decompression, and encryption at their core. The SPARC M7 processor also improves density by doubling the core count of previous-generation processors to support 32 cores per processor, with an architecture that supports up to 256 threads in as little as 2U of space. The SPARC M7 processor also sports improved per-thread performance, reliability, availability, and serviceability (RAS) capabilities, power-efficiency, and double the memory and I/O bandwidths of previous SPARC processors. A new cache and memory hierarchy, along with other improvements, helps to provide up to triple the processing speed of earlier systems.

Oracle's SPARC M7 processor–based servers power the real-time enterprise with the most efficient platforms in the world, allowing organizations to compete in today's digital marketplace, save money and time, and boost their bottom line. At the same time, Oracle's technology innovation creates value and drives lower costs and higher ROI for any organization. Based on SPARC M7 processor advances, the new SPARC server family from Oracle (Figure 1) provides new levels of performance and throughput. Scaling from one to 16 SPARC M7 processors, the servers constitute a flexible and extensible product family with very high levels of integration to help improve security, lower costs, and increase reliability. An optimized system design provides support for all enterprise services and application types. Uniformity of management interfaces and adoption of standards also help reduce administrative costs, while innovative chassis design provides density, efficiency, and economy for modern data centers.

SPARC T7-1 server SPARC T7-2 server SPARC T7-4 server SPARC M7-8 server SPARC M7-16 server

Figure 1. Oracle's SPARC M7 processor–based server product family.

Comparison of Features

Table 1 provides a feature comparison of the SPARC T7-1, T7-2, T7-4, M7-8 and M7-16 servers.

TABLE 1. SPARC M7 PROCESSOR–BASED SERVER FEATURES COMPARISON.

Feature	SPARC T7-1 Server	SPARC T7-2 Server	SPARC T7-4 Server	SPARC M7-8 Server	SPARC M7-16 Server
Form factor	2U, 737 mm / 29" deep	3U, 753 mm / 29.6" deep	5U, 835 mm / 32.9" deep	System Rack: 600 mm wide, 1200 mm deep, 2 m / 78.7" high Standalone: 10U, 813 mm / 32" deep	System Rack: 600 mm wide, 1200 mm deep, 2 m / 78.7" high
Physical domains	1			1 or 2 (Static)	1 ,2, 3, or 4 (Reconfigurable)
Processor	32-core 4.13 GHz SPARC M7 processor 64 MB Level 3 cache, fully shared and partitioned, 8 MB per each core cluster Up to 256 threads per processor Silicon Secured Memory 32 accelerator engines for In-Memory Query Acceleration and In-Line Decompression Encryption instruction accelerators in each core with direct support for 15 industry-standard cryptographic algorithms plus random number generation: AES, Camellia, CRC32c, DES, 3DES, DH, DSA, ECC, MD5, RSA, SHA-1, SHA-224, SHA-256, SHA-384, SHA-512				
Processor quantity	1	2	2 or 4	2–8	4–16
Maximum cores	32	64	128	256	512
Maximum threads	256	512	1,024	2,048	4,096
Memory	16 GB or 32 GB DDR4-2133 memory DIMMs, 8 or 16 DIMMs per processor DIMM sparing is a standard feature increasing system reliability and uptime.[1]				
Memory capacity[1]	Max 512 GB Min 128 GB	Max 1,024 GB Min 256 GB	Max 2,048 GB Min 256 GB	Max 4,096 GB Min 512 GB	Max 8,192 GB Min 1,024 GB
Internal 2.5-inch disk drive bays	8	6	8	NA	
SAS support for internal 2.5-inch disk drive bays	One integrated SAS3 HBA with RAID 0/1/10/1E supporting up to eight 2.5-inch SAS hard-disk drives (HDDs) or solid-state drives (SSDs)	Two integrated SAS3 HBAs with RAID 0/1/10/1E supporting up to six (2 + 4) SAS HDDs or SSDs	Two integrated SAS3 HBAs with RAID 0/1/10/1E supporting up to eight (4 + 4) SAS HDDs or SSDs	NA	
NVMe support for internal 2.5-inch disk drive bays	One optional factory-configured PCIe switch supporting up to four 2.5-inch NVMe SSDs	One or two optional factory-configured PCIe switches supporting up to four 2.5-inch NVMe SSDs	Two optional PCIe switches supporting up to eight (4 + 4) 2.5-inch NVMe SSDs	NA	

1. Raw memory capacities. DIMM sparing is enabled with fully populated memory and reserves one sixteenth of memory capacity. DIMM sparing enables automatic retirement of an entire DIMM without interrupting system operation, causing loss of memory capacity, or changing error protection capability.

TABLE 1. SPARC M7 PROCESSOR–BASED SERVER FEATURES COMPARISON (CONTINUED).

Feature	SPARC T7-1 Server	SPARC T7-2 Server	SPARC T7-4 Server	SPARC M7-8 Server	SPARC M7-16 Server
Maximum number of Oracle Flash Accelerator F160 PCIe Cards (NVMe)	6	6	8	16	32
Removable media	DVD+R/-W	DVD+R/-W	No DVD (accessed via USB and rKVMS)		
Management ports	One Ethernet 100BASE-T port One serial RJ45 port			Two Ethernet 100BASE-T ports (active/standby) Two serial RJ45 ports (active/standby)	
Video ports	One HD-15 VGA video port	Two HD-15 VGA video ports		NA	
USB ports	Two USB 2.0 (front) and two USB 3.0 (rear) ports		Four USB 3.0 ports	NA	
Ethernet	Four integrated 10GBASE-T ports[2] Two integrated Ethernet controllers			Via PCIe adapters cards	
PCIe 3.0 low-profile slots	6 slots Six x8 slots, or two x16 and two x8 slots Supported by 4 PCIe root complexes	8 slots Four x8 and four x16 slots Supported by 8 PCIe root complexes	16 hot-pluggable slots Eight x8 and eight x16 slots Supported by 12 PCIe root complexes	Up to 24 hot-pluggable slots Three x16 slots per processor One PCIe root complex per slot	Up to 48 hot-pluggable slots Three x16 slots per processor One PCIe root complex per slot
Total PCIe root complexes	5	10	20	Up to 32	Up to 64
N+N redundant power supplies	2 redundant hot-swappable AC 1000 W power supplies	2 redundant hot-swappable AC 2000 W power supplies	4 hot-swappable AC 3000 W power supplies	6 hot-swappable AC 3000 W power supplies	16 hot-swappable AC 3000 W power supplies
N+1 redundant hot-swappable fans	4 dual-fan modules, top loading	6 fans, top loading	5 dual-fan modules, rear loading	8 dual-fan modules, front loading	52 dual-fan modules, front and rear loading
Operating system	Oracle recommends Oracle Solaris 11.3 or later for enhanced performance and functionality, including features enabled by Software in Silicon technology Control, root, and I/O domains: » *Oracle Solaris 11.3 or later*[3] The following versions are supported within guest domains: » *Oracle Solaris 11.3 or later*[3] » *Oracle Solaris 10 1/13*[4] » *Oracle Solaris 10 8/11*[4] » *Oracle Solaris 10 9/10*[4] Applications certified for Oracle Solaris 8 or 9 only may run in an Oracle Solaris 8 or 9 Branded Zone running within an Oracle Solaris 10 guest domain.				

2. 10GBASE-T autonegotiates to 100 Mb/sec, 1 Gb/sec, and 10 Gb/sec, full-duplex only.
3. Versions of Oracle Solaris 11 prior to 11.3 are not supported on SPARC M7 processor–based servers.
4. Plus required patches.

SPARC M7 Processor

With its new Software in Silicon capabilities coupled with an innovative cache and memory hierarchy, Oracle's SPARC M7 processor delivers dramatically higher processing speed and revolutionary protection against malware and software errors.

The Silicon Secured Memory feature of the SPARC M7 processor (Figure 2) provides real-time data integrity checking to guard against pointer-related software errors and malware. It replaces very costly software instrumentation with low-overhead hardware monitoring. Silicon Secured Memory enables applications to identify erroneous or unauthorized memory access, diagnose the cause, and take appropriate recovery actions. The SPARC M7 processor has Crypto Instruction Accelerators integrated directly into each processor core. These accelerators enable high-speed encryption for over a dozen industry-standard ciphers, eliminating the performance and cost barriers typically associated with secure computing.

The SPARC M7 processor also incorporates hardware units that accelerate specific software functions or primitives. The on-chip data analytics accelerators offload database query processing and perform real-time data decompression. The In-Memory Query Acceleration delivers performance that is up to ten times faster compared to other processors. The In-Line Decompression feature allows up to two times more data to be stored in the same memory footprint, without any performance penalty.

Figure 2. The SPARC M7 processor combines 32 SPARC S4 cores along with Software in Silicon features to accelerate application and database performance.

The per-thread performance is improved with the entirely new on-chip L2 and L3 cache design and increased processor frequency. The 64 MB L3 cache is partitioned and fully shared, and hot cache lines are migrated to the closest partition to minimize latency and maximize performance. The architecture of the core clusters and partitioned cache is ideal for server virtualization and pluggable databases. System administration and performance tuning are easier, because the design minimizes interaction between logical domains or between databases. The processor can dynamically trade per-thread performance for throughput by running up to 256 threads, or it can run fewer higher-performance threads by devoting more resources to each thread. This flexibility allows the system to balance overall throughput versus per-thread performance for optimal results.

The 32-core SPARC M7 processor is binary-compatible with earlier SPARC processors and provides 256 hardware threads—more than any multicore processor previously available. It is ideal for virtualized cloud computing environments, supporting a large number of virtual machines and delivering excellent multithreaded performance. This processor enables organizations to rapidly scale the delivery of new network services with maximum efficiency and predictability.

Table 2 provides a comparison between Oracle's SPARC M7, SPARC M6, and SPARC T5 processors.

TABLE 2. SPARC M7, SPARC M6, AND SPARC T5 PROCESSOR FEATURE COMPARISON.

Feature	SPARC M7 Processor	SPARC M6 Processor	SPARC T5 Processor
CPU frequency	4.13 GHz	3.6 GHz	3.6 GHz
Out-of-order execution Dual-instruction issue Data/instruction prefetch	Yes Yes Yes	Yes Yes Yes	Yes Yes Yes
SPARC Core type	S4	S3	S3
Cores per processor Threads per core Threads per processor	32 8 256	12 8 96	16 8 128
Sockets in systems	Up to 16	Up to 32	Up to 8
Memory per processor	Up to 16 DDR4 DIMMs	Up to 32 DDR3 DIMMs	Up to 16 DDR3 DIMMs
Caches	16 KB L1 four-way instruction cache 16 KB L1 four-way data cache Shared 256 KB L2 four-way instruction cache (per quad cores) Shared 256 KB L2 eight-way data cache (per core pair) Shared 64 MB (L3) cache	16 KB L1 four-way instruction cache 16 KB L1 four-way data cache 128 KB L2 eight-way cache Shared 48 MB L3 twelve-way cache	16 KB L1 four-way instruction cache 16 KB L1 four-way data cache 128 KB L2 eight-way cache Shared 8 MB L3 sixteen-way cache
Large page support[1]	16 GB	2 GB	2 GB
Power management granularity	¼ of chip	Entire chip	Entire chip
Technology	20 nm technology	28 nm technology	28 nm technology

1. Large page support with Oracle Solaris 11.3

SPARC M7 Processor Architecture

In order to deliver commercial workloads with appropriate levels of throughput, the SPARC M7 processor implements a new cache and memory hierarchy coupled with other improvements that can provide up to triple the processing speed of previous-generation processors. Power management improvements are also key to the increased in-system performance, and dynamic voltage frequency scaling (DVFS) is provided.

Figure 3 illustrates the architecture of the SPARC M7 processor. The processor contains 32 SPARC S4 cores that are grouped into eight core clusters. Four memory controller units (MCUs) are provided, with each connecting to the buffer-on-board (BoB) ASICs via high-speed links. The BoB has two DDR4 channels with each connecting to a single memory DIMM. A total of up to 16 DDR4 DIMMs is supported per SPARC M7 processor. Two coprocessors are associated with each MCU and provide Software in Silicon functionality.

There are two Coherency Link Clusters (CLCs) that provide eight Coherency and Scalability Links (CL/SL) for connectivity and coherency with other SPARC M7 processors. Two I/O Links (IL) connect to the I/O controller ASICs in the SPARC M7 processor–based servers. Up to eight SPARC M7 processors can be connected in a single glueless symmetrical multiprocessing (SMP) system without additional logic. Larger systems are built using Oracle-designed switch ASICs that provide coherency switching.

A high-bandwidth, low-latency on-chip network (OCN) connects the eight L3 cache partitions to each other, to the four MCUs, and to the I/O and coherence gateways, each of which handles distinct sets of addresses. The OCN maintains coherency both on-chip and off-chip.

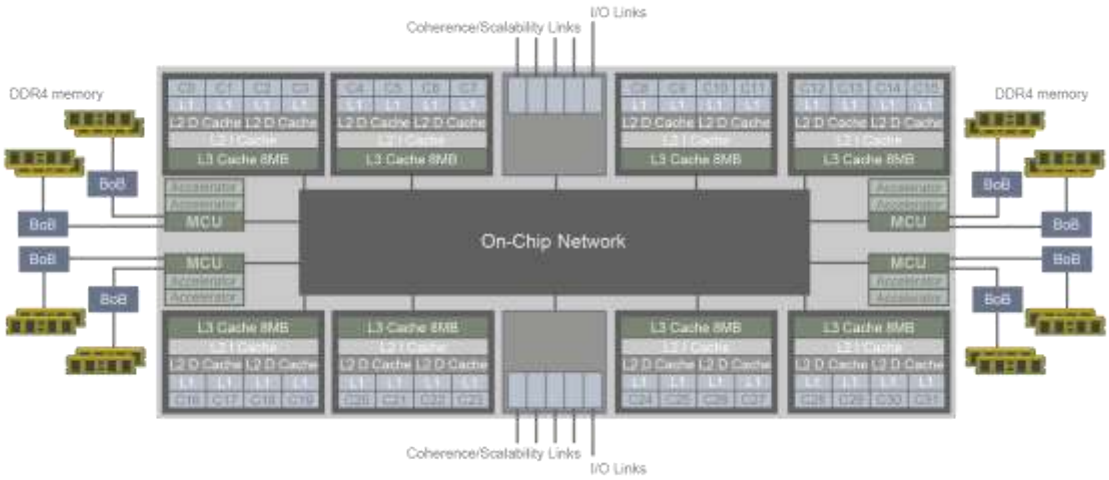

Figure 3. The SPARC M7 processor features 32 SPARC S4 cores, eight SPARC core clusters, four memory controller units (MCUs), and eight data analytics accelerators (DAXs).

SPARC S4 Core and Cache Architecture

The SPARC S4 core is dual-issue, out-of-order, and supports up to eight hardware threads. The core provides Dynamic Threading to optimize for the highest possible per-thread performance. Software can activate up to eight hardware threads (strands) on each core via Critical Threads Optimization. The processor hardware then dynamically and seamlessly allocates core resources among the active strands.

In the SPARC M7 processor, four SPARC S4 cores are combined into a SPARC core cluster with eight SPARC core clusters per SPARC M7 processor. Within the SPARC core cluster, each core has its own 16 KB L1 instruction and data cache. Two cores then share a 256 KB L2 data cache with the four cores sharing a 256 KB L2 instruction cache. The L3 cache is fully shared and partitioned. The L3 partition is eight-way set-associative with a 64-byte line size, and is composed of two address-interleaved banks. Any L3 partition may serve a request from any of the 32 cores of the SPARC M7 processor. Hot cache lines are migrated to the closest L3 cache partition to optimize for performance.

Software in Silicon Technology

Most processor chip development focuses on better and faster general-purpose processing. Several years ago Oracle initiated a revolutionary project to move in-memory database functions directly onto the chip, with hard-wired protection for data in memory. By innovating at the processor, system, and application levels, Oracle is in a unique position to optimize application performance through this approach. The SPARC M7 processor capitalizes on this ability by providing Software in Silicon functionality integrated into the SPARC M7 processor itself.

The SPARC M7 processor incorporates on-chip accelerators to off-load in-memory database query processing and perform real-time data decompression, while crypto instruction accelerators are integrated directly into each processor core. Together, the Software in Silicon features deliver significant advantages for security, performance and efficiency, including the following:

- Silicon Secured Memory provides real-time data integrity checking to guard against pointer-related software errors and malware, replacing very costly software instrumentation with low-overhead hardware monitoring. Silicon Secured Memory enables applications to identify erroneous or unauthorized memory access, diagnose the cause, and take appropriate recovery actions.
- Accelerated cryptography helps eliminate the performance and cost barriers typically associated with secure computing—which is increasingly essential for modern business operation.
- In-Memory Query Acceleration provided by the accelerators delivers performance that is up to ten times faster compared to other processors.
- The In-Line Data Decompression feature enables storing up to two times more data in the same memory footprint, without a performance penalty.

In addition to the Crypto Instruction Accelerators that are included in every core the SPARC M7 processor contains eight data analytics accelerators (DAX), each with four pipelines, or engines. These engines can process 32 independent data streams, offloading the processor cores to do other work. The data analytics accelerator engines can process query functions such as decompress, scan, filter and join.

The data analytics accelerators use very low-overhead interprocess communication and extremely fast atomic operations. For example, DAXs located on different processors can exchange messages and access remote memory locations, exchanging locks without CPU involvement. Utilizing this functionality requires Oracle Database 12*c* with the In-Memory option and Oracle Solaris 11.3 or later. The sections that follow describe Software in Silicon features enabled by the on-chip accelerators.

Silicon Secured Memory

Silicon Secured Memory in the SPARC M7 processor provides the first-ever hardware-based memory protection by placing dynamic pointer checking in hardware. Silicon Secured Memory detects and reports memory reference errors and stops unintentional or malicious accesses to the data in memory.

Some programming languages such as C and C++ remain vulnerable to memory corruption caused by software errors. These kinds of memory reference bugs are extremely hard to find, and victims usually notice corrupted data only long after the corruption has taken place. Complicating matters, databases and applications can have tens of millions of lines of code and thousands of developers. Importantly, errors such as buffer overflows are a major source of security exploits that can put an organization at risk.

Modern applications use many threads working on large shared-memory segments. Bugs or pointer problems in these applications can cause highly unpredictable behavior and consume excessive amounts of an application developer's time to troubleshoot and diagnose. Silent data corruption and buffer overruns are two of these difficult-to-diagnose problems. For both problems, Silicon Secured Memory in the SPARC M7 processor dramatically reduces the time it takes for application developers to troubleshoot memory reference bugs. For silent data corruption, Silicon Secured Memory can facilitate immediate action to be taken by the application, preventing costly recovery efforts.

Figure 4 illustrates the problem of silent data corruption where two application threads (A and B) accidentally access the same memory location. The color coding demarcates the memory areas that each thread should be accessing, respectively. However, a software programming error can mean that Thread A erroneously writes into the red-outlined area of Thread B. This error is typically not caught immediately, and is potentially detected only when that memory is read by Thread B. Thread B now has data that has been silently corrupted by Thread A, and the

source of the corruption is often extremely difficult to trace. This problem is extremely hard to diagnose, and typically is manifest as a software bug with potentially serious consequences.

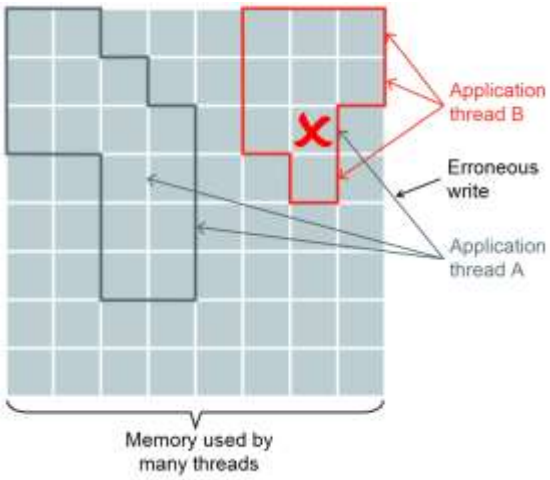

Figure 4. Silent data corruption occurs when two threads mistakenly write to the same memory location.

Buffer overruns are an additional problem that can occur in application development. Simply stated, buffer overruns imply that an application has erroneously started writing data beyond its allocated area (Figure 5). Through this error, sensitive data could be leaked into other memory locations and the application would not be aware of it. An application with malicious intent could then legitimately read all this sensitive information. Buffer overruns can present a catastrophic security nightmare, often seen in today's world in the form of malicious virus attacks.

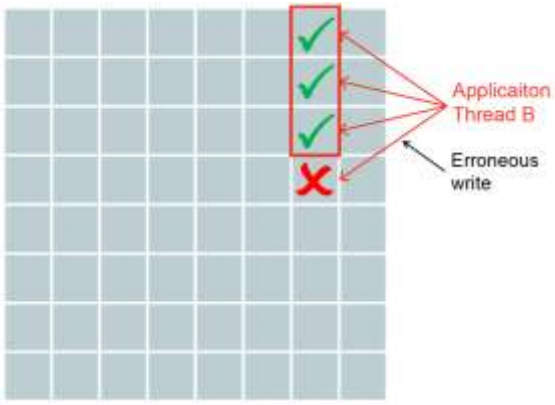

Figure 5. Buffer overruns can represent a significant security risk.

Silicon Secured Memory in the SPARC M7 processor combats these problems by using a key for each (memory) pointer to serve as the memory version. During the process of memory allocation, a corresponding code is written into the memory as its version. When this memory is accessed by any pointer, the key and code are compared by the hardware. If they match, the access is legal. If they do not match, there is a memory reference error, which is caught immediately.

Encryption Acceleration

Enhanced security has never been more important, and SPARC processors and systems have a long history of providing processor–based cryptographic acceleration. Each of the 32 cores in the SPARC M7 processor includes a Crypto Instruction Accelerator with direct support for 15 industry-standard cryptographic algorithms plus random number generation. Accelerated cryptography is supported through the Cryptographic Framework in Oracle Solaris.

The SPARC M7 processor permits access to cryptographic cypher hardware implementations with supported algorithms that include AES, Camellia, CRC32c, DES, 3DES, DH, DSA, ECC, MD5, RSA, SHA-1, SHA-224, SHA-256, SHA-384, and SHA-512. The cyphers are implemented within the appropriate pipeline itself rather than as a coprocessor. This approach yields a more efficient implementation of the hardware-based cyphers as well as no privilege-level changes, resulting in a large increase in efficiency in cryptographic algorithm calculations. In addition, database operations can make much more efficient use of the various cryptographic cyphers that are implemented within the instruction pipeline itself. Using the built-in encryption on the SPARC M7 processor across all layers of the Oracle stack provides greater data security with almost no loss in performance.

In-Memory Query Acceleration

In-Memory Query Acceleration was designed to work with Oracle Database In-Memory, which was architected with fast analytics response as its primary design tenet. The traditional way of storing and accessing data in a database employs a row format. This approach works well for transactional workloads that are subject to frequent inserts and updates as well as for reporting style queries. However, analytics run best on a columnar format. With Oracle Database In-Memory, it is possible to have a dual-format architecture that provides both row format for OLTP operations and column format for analytic operations.[1]

Oracle Database In-Memory populates the data in an in-memory column store. A set of compression algorithms is automatically run on the data being stored in the in-memory column, providing storage economies. Moreover, when a query is run, it scans and filters data in its compressed format, eliminating the need to decompress the data. The in-memory column store creates In-Memory Compression Units (IMCUs), as shown in Figure 6. The in-memory columnar data is fragmented into these smaller IMCUs so that parallelization is possible when running a query on the overall data.

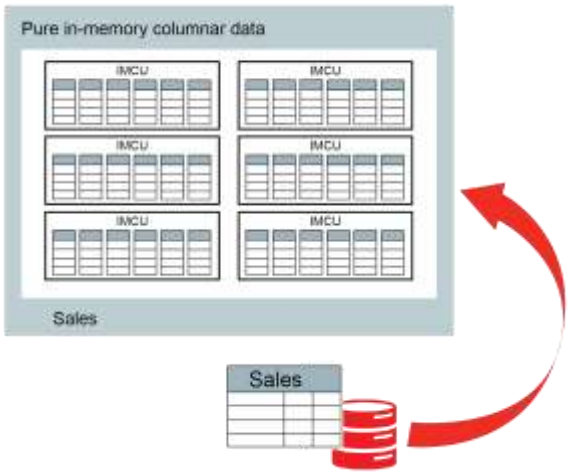

Figure 6. In-memory columnar data is fragmented into smaller IMCUs to enable parallelization.

1 In-Memory Query Acceleration is supported on SPARC T7 and M7 servers with these prerequisites: Oracle Solaris 11.3 or later and Oracle Database 12*c* 12.1.0.2 Bundle Patch and In-Memory option.

When a core in the SPARC M7 processor receives a database query, it can be offloaded to the on-chip accelerator. Accelerated database operations include the following:

- Select: Filter to reduce a column
- Scan: Search ("where" clause)
- Extract: Decompression
- Translate: Lookup to accelerate big-to-small joins

After the query is offloaded, the core is free to resume other jobs such as higher-level SQL functions. Meanwhile, the accelerator runs the query and places the result in the L3 cache for fast access by the core. Once the relevant core is informed of the completion of the query, it picks up the result.

Beyond accelerating operations, the other advantage of this query offload mechanism is the massive parallelization that is facilitated by the 32 accelerator engines within each SPARC M7 processor. Each of the 32 cores in the processor has access to all of these accelerator engines and can use them simultaneously to run a single query in a completely parallel fashion. The mechanics of this parallelism are achieved by the processor, and do not require the application code or the database to perform any extra operations. The accelerator can take data streams directly from the memory subsystem through the SPARC M7 processor's extremely high-bandwidth interfaces. As a result, queries can be performed on in-memory data at speeds determined by the memory interface, rather than being controlled by the cache architecture that connects to the processor cores.

In-Line Data Decompression

Compression is absolutely key to placing more data in memory and in storage. The speed of decompression is most important for database applications, where reading typically outweighs writing. Unfortunately, although the performance of decompression on today's processors is adequate for disk access, it is slow for flash memory, and presents an enormous bottleneck for in-memory database applications.

To address this challenge, a feature called In-Line Data Decompression is implemented in the SPARC M7 processor's DAX as an integral step of the query process. The accelerator decompresses data and runs the query function in a single step, thus eliminating multiple reads and writes. The result is no-penalty, in-line decompression that can run at memory speeds—greater than 120 GB/second. The decompression sequence involves the following steps:

- The processor core offloads query work to the accelerator, which reads the full compressed data (OZIP compression is used).
- The accelerator decompresses data on the fly and evaluates the query in a single step without any additional read or write operations.
- The processor core then writes out the final result as uncompressed data.

SPARC M7 Processor–Based Server Family Overview

The SPARC M7 processor-based servers are designed for cloud infrastructures that required high levels of security, performance, and efficiency. These SPARC servers are ideal for database, Java, middleware, and enterprise applications, offering exceptional throughput performance and memory bandwidth. The server product family provides support for one to 16 SPARC M7 processors, supporting a very broad range of applications, capabilities, and capacities.

Common new hardware features of the servers include the following:

- SPARC M7 32-core 4.13 GHz processors with Software in Silicon features
- 16 GB and 32 GB DDR4-2133 memory DIMMs
- PCIe 3.0 x16–capable expansion slots
- Support for NVM Express (NVMe) flash devices
- On-board 12 Gb/sec SAS3 I/O controllers (SPARC T7-1, T7-2 and T7-4 servers)
- Embedded USB (eUSB) storage device supporting booting over InfiniBand network

Memory Subsystem

Each SPARC M7 processor supports up to16 DDR4 memory DIMMs via eight buffer-on-board (BoB) ASICs. Up to 512 GB of memory is supported per processor with sixteen 32 GB DIMMs. Memory bandwidth for the 4.13 GHz SPARC M7 processor is 333 GB/sec. Half-populated and fully populated memory configurations are supported. Detailed configuration policies are discussed later in the model-specific sections.

The physical address space provided by the memory DIMMs and controlled by an individual SPARC M7 processor is interleaved to maximize performance. Half-populated memory configurations are 8-way interleaved, and fully populated configurations with16 DIMMs per processor are 16-way interleaved. The SPARC M7 processor also supports a 15-way interleaved configuration. The switch from 16-way to 15-way can be done dynamically. This capability is the basis of a new availability feature called DIMM sparing, first introduced with the SPARC M7 processor–based servers. DIMM sparing increases system uptime by reducing the need to perform DIMM replacement service work.

DIMM sparing is the function of removing a faulty DIMM from the configuration, thus preventing it from causing an unplanned system interruption. By leaving 1/16 of the capacity of each DIMM unused, the bad DIMM can be retired and its content remapped into the remaining 15 DIMMs. DIMM sparing is done automatically when a DIMM is determined to be faulty, with no interruption to application services. Through this process, the system memory capacity is unchanged and error protection remains intact after DIMM sparing is performed. The system simply continues to run with no loss of capacity and no increased exposure to failures. Consequently, there is no need to take the system down to service the hardware. The actual DIMM replacement process can wait until a second DIMM in the same memory bank needs to be replaced.

DIMM sparing is enabled in SPARC M7 processor–based servers with fully populated memory configurations (16 DIMMs per processor). DIMM sparing is not supported in half-populated memory configurations. While not recommended, it is possible to disable DIMM sparing in fully populated memory configurations.

I/O Subsystem

The SPARC T7 and M7 servers share the same basic design for the I/O subsystem. Each SPARC M7 processor is connected to one or two I/O controller ASICs via I/O Links (IL). There are two ILs on the SPARC M7 processor as well as in the I/O controller ASIC.

There are two different implementations of the connection between the SPARC M7 processor and the I/O controller ASIC being used. In the SPARC T7-1, M7-8 and M7-16 servers each SPARC M7 processor is connected to a single I/O controller ASIC using both ILs. The SPARC T7-2 and T7-4 servers utilize a cross-over connection scheme in order to provide connectivity to two I/O controller ASICs (and PCIe devices) even in case one of the processors or ILs is not available. In the cross-over connection, one IL in the processor connects to one I/O controller, and the other IL connects to another I/O controller. See more details in the model-specific sections that follow.

I/O Controller ASIC

The PCIe infrastructure is provided by the I/O controller ASICs. Each ASIC provides five PCIe 3.0 root complexes with an aggregated bandwidth of 72 GB/sec. Because the I/O controller ASIC hosts the entire PCIe fabric, it remains intact when processors are added or removed. As a result, the PCIe device paths do not change, because the connection to the device is fixed within the root complex provided by the I/O controller ASIC. The I/O controller ASIC utilized in SPARC M7 processor–based servers provides significant innovation, including the following:

- Two x16 I/O Links connect to the SPARC M7 processor, and each link comprises two x8 connections.
- A single lane failure is supported on each x8 I/O Link connection.
- Each I/O Link participates in hardware cache coherency.
- Dual-host processor failover is used in the SPARC T7-2 and T7-4 servers.
- SR-IOV compliance is provided.
- Address translation per DMA stream is provided.
- Relaxed packet ordering per DMA stream is provided.
- There are four PCIe 3.0 x16 ports, which are quadfurcatable and can be implemented as a single x16, four x4, or two x8 PCIe ports.
- There is one PCIe 3.0 x8 port, which is bifurcatable and can be implemented as a single x8 or two x4 PCIe ports.
- Each of the five PCIE 3.0 ports is an independent root complex.

NVM Express Technology

SPARC M7 processor–based servers provide support for an emerging flash storage technology known as NVM Express, or NVMe. The NVMe specification defines an optimized PCIe-based interface for solid-state drives (SSDs). Utilizing non-volatile memory, NVMe-based SSDs provide both lower latency and better throughput performance relative to SAS or SATA-based SSDs. NVMe utilizes PCIe signaling and provides an 8 GT/second, x4 interface per drive, yielding approximately 4 GB/second full-duplex to the drive.

All SPARC M7 processor–based servers support the Oracle Flash Accelerator F160 PCIe Card—an NVMe-based SSD device on a low-profile PCIe card. The SPARC T7-1, T7-2, and T7-4 servers support internal 2.5-inch small form factor (SFF) NVMe SSDs in select drive bays that can also support SAS-based HDDs and SSDs. A factory-configured NVMe PCIe switch card and cables are required when the SFF NVMe drives are used. The switch card uses an x8 PCIe 3.0 interface and provides fan-out and electrical retiming functions for up to four x4 downstream links (one per NVMe drive).

NVMe devices are hot-plug capable, but OS-specific hot-plug procedures must be followed. An `nvmeadm` command is provided that lets administrators list drive health and firmware level, check temperatures, get error logs, and access SMART data, as well as conduct a security erase and perform low-level formatting.

Embedded USB Storage and Oracle Solaris Boot Pool

SPARC M7 processor–based servers support a new boot process that allows booting from a greater variety of devices. A conventional boot process requires that the boot devices be accessible by the system firmware. For example, a network boot over InfiniBand has previously not been supported for that reason.

The new Oracle Solaris boot process includes a new concept called a *boot pool*. A boot pool is a boot device that is used to store boot archives. In SPARC M7 processor–based servers, one or more embedded USB (eUSB) storage devices are grouped together to form a boot pool, making the boot pool accessible to the OpenBoot PROM firmware. The eUSB storage is an internal USB flash memory device that is installed into the system at the factory. SPARC T7-1, T7-2, and T7-4 servers include one eUSB device. SPARC M7-8 and M7-16 servers include one eUSB device for every CPU, memory, and I/O unit (CMIOU) chassis board, allowing the boot pool to consist of multiple, striped eUSB storage devices.

The existence of the local boot pool allows the OpenBoot PROM firmware to load the boot archive and subsequently mount the root file system in the root pool using iSCSI over IP over InfiniBand (IPoIB). SPARC M7 processor–based servers also provide a fall-back mechanism for the new boot process, which can be used when the eUSB-based boot archives are not available. A boot archive exists in the flash memory of the system service processor (SP). This boot archive is loaded onto the SP at the factory and is intended to be used only when other means of booting are not available.

PCIe Adapter Cards

SPARC M7 processor–based servers feature both PCIe 3.0 x8 and x16 expansion card slots. Supported options and requirements vary by server models. At release, available adapter cards from Oracle include the following:

- Dual 16 Gb Fibre Channel HBA
- Dual 8 Gb Fibre Channel HBA
- Sun Dual Port GbE PCIe 2.0 Low Profile Adapter, MMF
- Sun Quad Port GbE PCIe 2.0 Low Profile Adapter, UTP
- Sun Dual Port 10GBase-T Adapter
- Sun Dual Port 10 GbE SFP+ PCIe 2.0 Low Profile Adapter
- SAS 3.0 – 8 Port (2 x4) 12 Gb External LP
- Oracle Dual Port QDR InfiniBand Adapter M3

SPARC T7-1, T7-2, and T7-4 Servers

The SPARC T7-1, T7-2, and T7-4 servers are designed to provide breakthrough security and performance while maximizing reliability and minimizing power consumption and complexity. These systems are ideal for scale-out applications where high availability is provided by replicating systems with applications that support failover between systems. With common, shared components and subsystems, a separate design is used in each of the SPARC T7-1, SPARC T7-2, and SPARC T7-4 servers to optimize the system for its particular design point and capabilities. The servers feature a robust chassis, component, and subassembly design, enhanced system and component serviceability, and minimized cabling for maximized airflow.

- The SPARC T7-1 server is the entry model with a single processor. However, with 256 hardware threads and 512 GB of memory it outperforms dual-processor alternatives. Like all the models in this server family, the SPARC T7-1 server features the full suite of Software in Silicon features, including secure operation with Silicon Secured Memory and high performance with In-Memory Database Query Acceleration and In-Line Data Decompression.
- The SPARC T7-2 server offers twice the capacity and more resources compared to the SPARC T7-1 server. The dual-processor SPARC T7-2 server also includes additional availability features, such as innovative cross-connections between the processor and the I/O controller ASICs, dual integrated SAS controllers, and optional dual NVMe PCIe switches for internal NVMe SSDs.
- The SPARC T7-4 server supports up to four processors, which are mounted on the front-accessible processor modules. The entry configuration of two processors can easily be expanded to four processors, if needed. The SPARC T7-4 server also offers individually hot-pluggable PCIe cards supported with a large number of dedicated PCIe root complexes, making the server ideal for workload consolidation into a virtualized private cloud.

SPARC T7-1 Server

The SPARC T7-1 server is the entry model and features a single SPARC M7 processor in a 2U enclosure. Standard features include eight on-board memory DIMM slots that are expandable to a total of 16 with optional dual mezzanine cards (memory risers)—yielding up to 512 GB of system memory. Figure 7 illustrates the front and rear perspectives of the SPARC T7-1 server. The server includes six low-profile PCIe 3.0 expansion slots, accessible from the rear of the system.

The eight 2.5-inch small form factor (SFF) front-loading drive bays are all supported by an onboard 12 Gb/sec SAS HBA, which provides RAID 0, 1, 10, and 1E protection. Oracle Solaris ZFS provides higher levels of RAID support. A factory-configured option is available to support up to four 2.5-inch SFF NVMe SSDs. The option includes an NVMe PCIe switch card installed in PCIe slot #3 and internal cabling to the drive cage. Mixing SAS and NVMe drives is supported. Other standard features of the SPARC T7-1 server include the following:

- A front-load DVD drive is supported by onboard USB-SATA.
- Four 10GBASE-T Ethernet ports (RJ45) are provided on the back of the system, supported via dual onboard NICs.
- Two hot-swappable N+N power supply units (PSUs), 1000 W at 200VAC–240VAC, are inserted from the rear of the system.
- Four top-loading hot-swappable fan modules are provided, each with dual counter-rotating fans.
- Dual USB 2.0 ports in the front and dual USB 3.0 ports and a VGA video port (HD15) in the rear are provided.

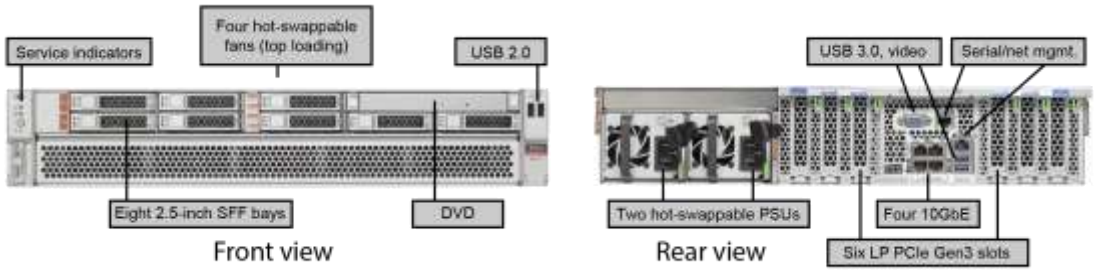

Figure 7. SPARC T7-1 server front and rear perspectives.

Figure 8 provides a block diagram of the SPARC T7-1 server. The SPARC M7 processor connects to onboard memory slots as well as optional memory risers to support up to 16 DDR4 DIMM slots. The SPARC M7 processor connects directly to the I/O controller ASIC, which furnishes the PCIe root complexes. The optional NVMe PCIe switch card enables up to four 2.5-inch NVMe devices to be placed into the center four SFF drive bays. Remote keyboard, video, and mouse (rKVM) functionality is provided by the Oracle Integrated Lights Out Manager (Oracle ILOM) remote console that runs on the service processor (SP). Separate serial (RJ45) and Ethernet (100BASE-T, RJ45) management ports are provided to interface with the SP. The onboard 10GBASE-T network ports can also be used (with the sideband feature enabled) to connect at speeds of up to 10 Gb/sec to the SP. Two of the six PCIe slots (#3 and #4) can support x16 connectivity if their adjacent slots (#2 and #5, respectively) are empty.

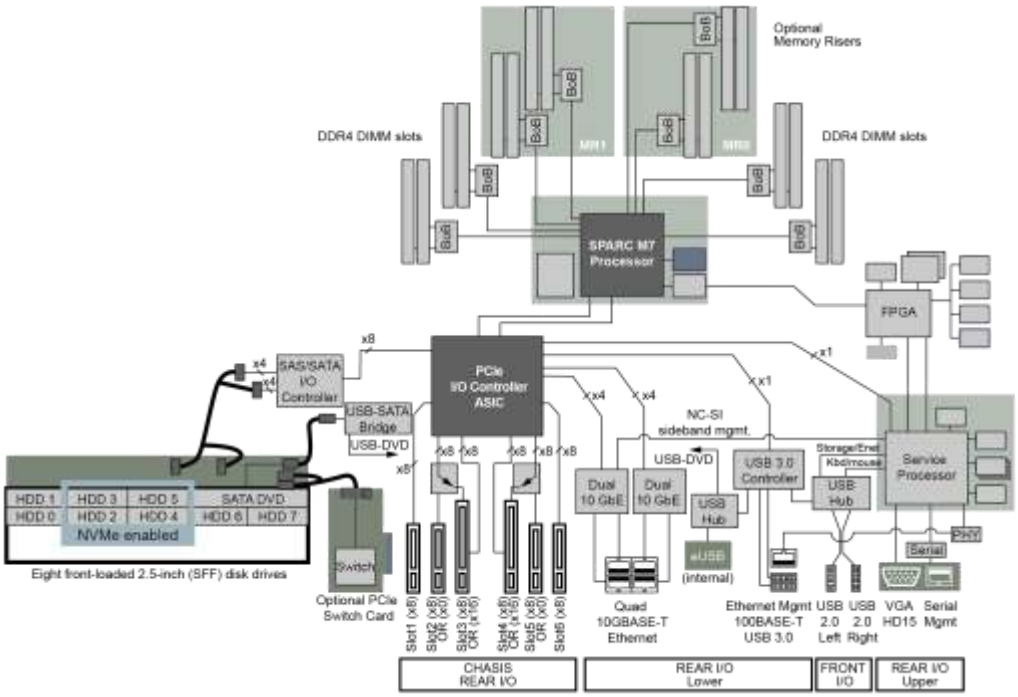

Figure 8. The SPARC T7-1 server features a single 32-core SPARC M7 processor and I/O controller ASIC.

Table 3 describes how PCIe devices share the five root complexes on the SPARC T7-1 server's I/O controller ASIC.

TABLE 3. SPARC T7-1 SERVER ROOT COMPLEX MAPPING.

Root Complex	Target	Speed
Root Complex 0	PCIe slot 6 SAS/SATA I/O controller	x8 x8
Root Complex 1	PCIe slot 4 PCIe slot 5	x8 or x16 x8 or x0
Root Complex 2	Dual 10 GbE Dual 10 GbE USB 3.0 controller Service processor (onboard graphics)	x4 x4 x1 of x4 port x1 of x4 port
Root Complex 3	PCIe slot 2 PCIe slot 3	x8 or x0 x8 or x16
Root Complex 4	PCIe slot 1	x8

SPARC T7-2 Server

The SPARC T7-2 server features dual SPARC M7 processors housed in a 3U rackmount enclosure. Up to 1 TB of server memory is mounted on eight riser cards, each with two or four memory DIMMs. Figure 9 illustrates front and rear perspectives of the SPARC T7-2 server. The server includes eight low-profile PCIe 3.0 expansion slots, accessible from the rear of the system.

The six 2.5-inch SFF front-loading drive bays are supported by two onboard 12 Gb/sec SAS HBAs (split two and four drives across the HBAs), which provide RAID 0 and 1 protection. The SAS HBA with four bays can also support RAID 10 and 1E. Oracle Solaris ZFS provides higher levels of RAID support. Factory-configured options are available to support up to four 2.5-inch SFF NVMe drives. The options include an NVMe PCIe switch card internally cabled to the drive cage. Mixing SAS and NVMe drives is supported with the following available options:

- **Single NVMe PCIe switch card**. The switch card is installed in PCIe slot #1 and supports all four NVMe–capable drive bays (the upper four drive bays).
- **Dual NVMe PCIe switch cards**. Switch cards are installed in PCIe slots #1 and #2 (separate root complexes) and each switch supports two NVMe–capable drive bays.

Other standard features of the SPARC T7-2 server include the following:

- A front-load DVD drive is supported by an onboard USB-SATA bridge.
- Four 10GBASE-T Ethernet ports (RJ45) are provided on the back of the system, supported via dual onboard NICs.
- Two hot-swappable N+N PSUs, 2000 W at 200VAC–240VAC, insert from the rear of the system.
- Six hot-swappable fan modules load from the top of the chassis.
- Dual USB 2.0 ports in the front and dual USB 3.0 ports in the rear are provided.
- One VGA video port (HD15) is provided in the rear of the chassis.

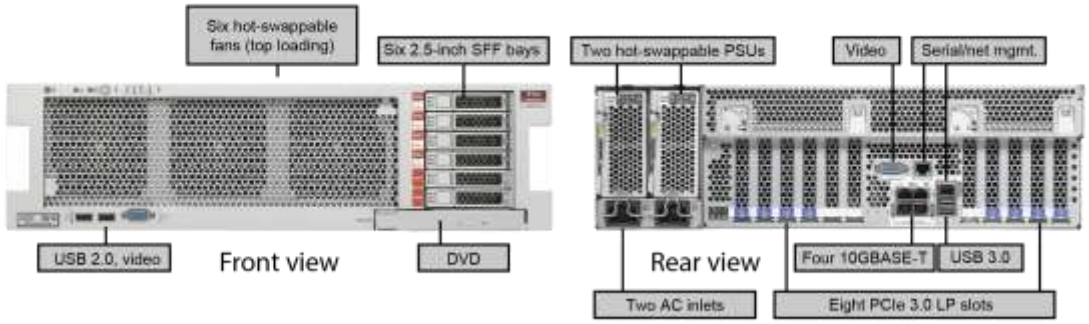

Figure 9. SPARC T7-2 server front and rear perspectives.

Figure 10 provides a block diagram of the SPARC T7-2 server. Two sockets for SPARC M7 processors connect to the memory risers to support up to 32 DDR4 DIMM slots (16 per processor socket). Redundant coherency links connect the processors together. Two I/O controllers are connected to both SPARC M7 processors. The cross-connected design maintains access to all I/O devices even in the event of a processor failure. If the system boots with only one processor, the PCIe device paths and all 10 PCIe root complexes remain intact. Four of the six PCIe slots are x16 capable and four are wired for x8. Essential for high-throughput devices, the x16 capable PCIe slots have dedicated root complexes, with no other devices sharing their root complex.

Remote keyboard, video, and mouse (rKVM) functionality is provided by the Oracle ILOM remote console that runs on the service processor (SP). Separate serial (RJ45) and Ethernet (100BASE-T, RJ45) management ports are provided to interface with the SP. The onboard 10GBASE-T network ports can also be used (with the sideband feature enabled) to connect at speeds of up to 10 Gb/sec to the SP.

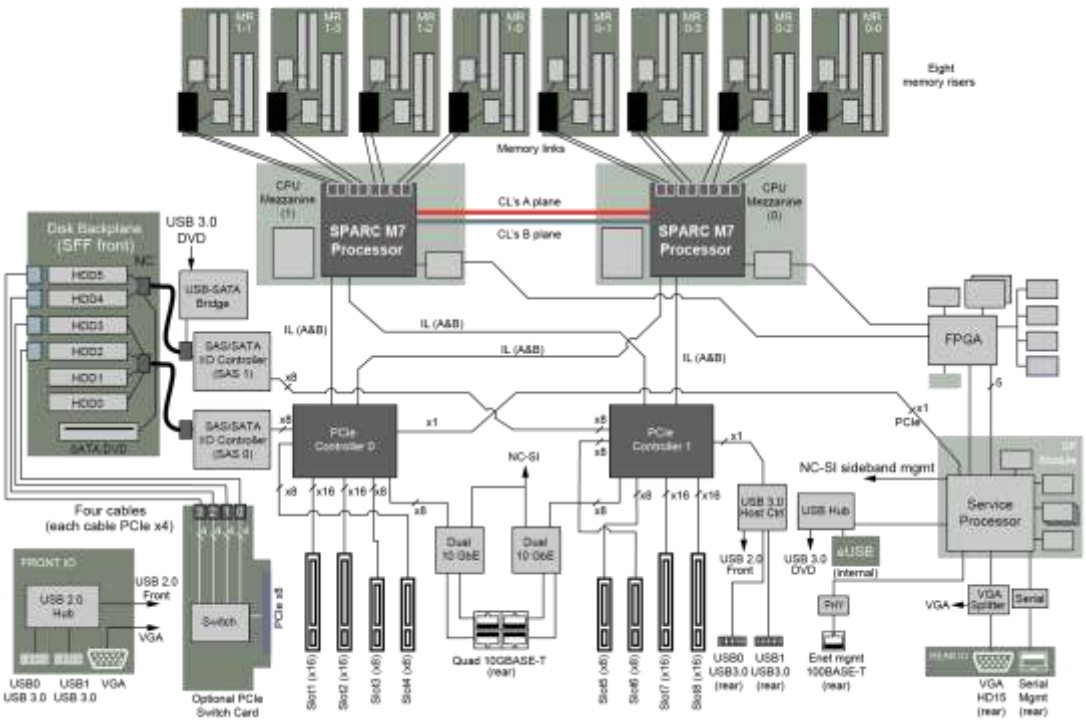

Figure 10. SPARC T7-2 server has dual SPARC M7 processors cross-connected to two I/O controller ASICs for higher availability.

Table 4 describes how the PCIe devices share the root complexes on each of the I/O controllers.

TABLE 4. SPARC T7-2 SERVER I/O CONTROLLER AND ROOT COMPLEX MAPPING.

I/O Controller	Root Complex	Target	PCIe Speed
I/O Controller 0	Root Complex 0	SAS/SATA I/O controller PCIe slot 4	x8 x8
	Root Complex 1	PCIe slot 3 Dual 10 GbE	x8 x8
	Root Complex 2	PCIe slot 2	x16
	Root Complex 3	PCIe slot 1	x16
	Root Complex 4	Service processor (onboard graphics)	x1 of x4 port
I/O Controller 1	Root Complex 0	SAS/SATA I/O controller PCIe slot 6	x8 x8
	Root Complex 1	PCIe slot 8	x16
	Root Complex 2	PCIe slot 7	x16
	Root Complex 3	Dual 10 GbE PCIe slot 5	x8 x8
	Root Complex 4	USB controller	x1 of x4 port

SPARC T7-4 Server

The SPARC T7-4 server features four SPARC M7 processors and up to 2 TB of memory housed in a 5U rackmount enclosure. Figure 11 illustrates the front and rear perspectives of the SPARC T7-4 server. The server includes eight x8 and eight x16 PCIe 3.0 expansion slots fitted with hot-pluggable PCIe carriers—accessible from the rear of the system.

The eight 2.5-inch SFF front-loading drive bays are split evenly across two onboard 12 Gb/sec SAS HBAs, which provide RAID 0, 1, 10, and 1E. Oracle Solaris ZFS provides higher levels of RAID support. In addition to SAS HDDs and SAS SSDs, factory-configured options are available to support up to eight 2.5-inch SFF NVMe SDD drives. The options include up to two NVMe PCIe switch cards internally cabled to the drive cage. Each NVMe PCIe switch card enables four drive bays to support NVMe drives. Mixing SAS and NVMe drives is supported.

Other standard features of the SPARC T7-4 server include the following:

- Four 10GBASE-T Ethernet ports (RJ45) are provided on the back of the system, supported via dual onboard NICs.
- Sixteen PCIe low-profile hot-pluggable carrier I/O slots are supported by four I/O controllers ASICs, including
  - Eight PCIe 3.0 x8 slots
  - Eight PCIe 3.0 x16 slots
- Four hot-swappable N+N redundant PSUs, 3000W at 200VAC–240VAC, insert from the front of the system.
- Five hot-swappable fan modules load from the top of the chassis.
- Local (front and rear USB and video) and remote keyboard, video, and mouse (KVM) functionality is provided.
- Separate serial and network management ports interface with the onboard service processor.

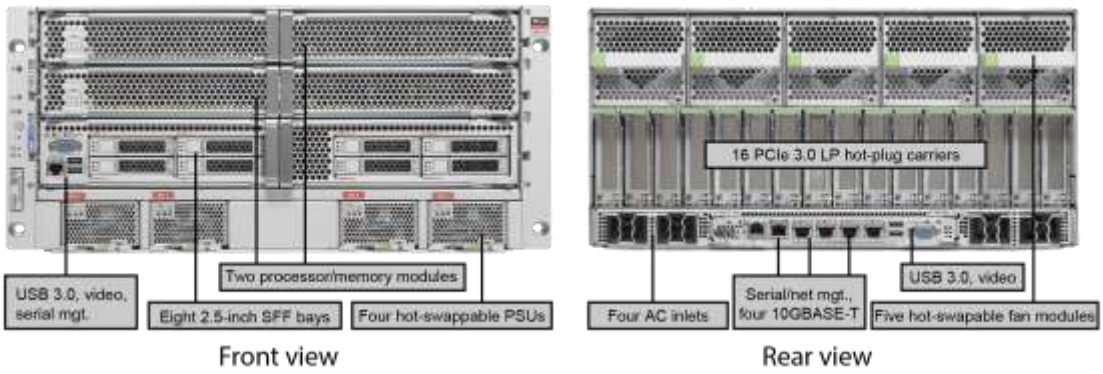

Figure 11. SPARC T7-4 server front and rear perspectives.

Figure 12 provides a block-level diagram of the SPARC T7-4 server. Up to two specialized processor/memory modules each provide two SPARC M7 sockets and 32 DDR4 DIMM slots (16 DIMM slots per processor socket, 64 DIMM slots per server). A midplane redundantly connects the coherency links of each processor socket to a processor socket on the other processor/memory module. Redundant ILs connect each processor socket to two of the four I/O controllers. Four I/O controllers furnish PCIe root complexes for all PCIe slots in the system, providing substantial I/O capabilities. SAS HDDs, SAS SSDs, or NVMe SSDs can be installed in any of the drive bays. Importantly, optional NVMe PCIe switch cards do not consume any of the system's hot-pluggable PCIe carrier slots at the rear of the chassis, leaving more capacity for additional I/O.

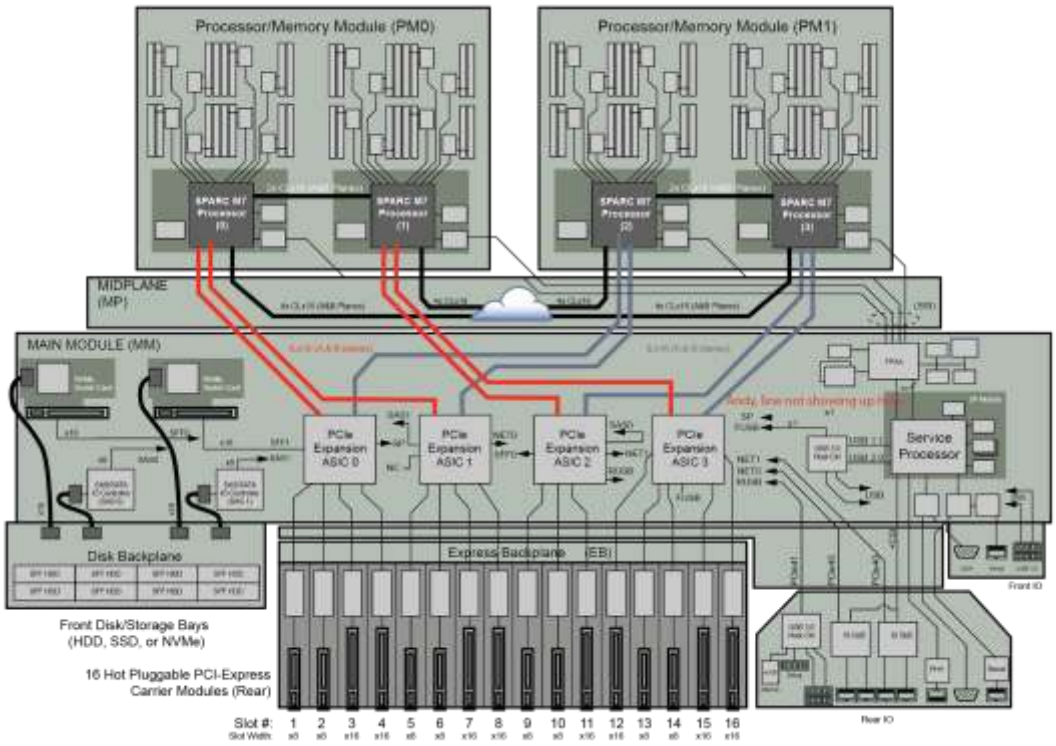

Figure 12. SPARC T7-4 servers feature two dual processor/memory modules and four I/O controllers ASICs.

Table 5 describes how the PCIe devices share the root complexes of the four I/O Controllers.

TABLE 5. SPARC T7-4 SERVER I/O CONTROLLER AND ROOT COMPLEX MAPPING.

I/O Controller	Root Complex	Target	PCIe Speed
I/O Controller 0	Root Complex 0	PCIe slot 1 PCIe slot 2	x8 x8
	Root Complex 1	NVMe PCIe switch card slot	x16
	Root Complex 2	PCIe slot 4	x16
	Root Complex 3	PCIe slot 3	x16
	Root Complex 4	Service processor Front USB 3.0 controller	x1 of x4 port x1 of x4 port
I/O Controller 1	Root Complex 0	PCIe slot 5 PCIe slot 6	x8 x8
	Root Complex 1	SAS/SATA I/O controller	x16
	Root Complex 2	PCIe slot 8	x16
	Root Complex 3	PCIe slot 7	x16
	Root Complex 4	Dual 10 GbE	x8

I/O Controller 2	Root Complex 0	PCIe slot 9 PCIe slot 10	x8 x8
	Root Complex 1	Dual 10 GbE	x8
	Root Complex 2	NVMe PCIe switch card slot	x16
	Root Complex 3	PCIe slot 11	x16
	Root Complex 4	Rear USB 3.0 controller	x1 of x4 port
I/O Controller 3	Root Complex 0	PCIe slot 13 PCIe slot 14	x8 x8
	Root Complex 1	PCIe slot 12	x16
	Root Complex 2	PCIe slot 16	x16
	Root Complex 3	PCIe slot 15	x16
	Root Complex 4	Not connected	NA

SPARC M7-8 and M7-16 Servers

The SPARC M7-8 and M7-16 servers are designed for modern cloud infrastructures. They are ideal for database and commercial business applications requiring operational efficiency, reliability, and scalability for large mission-critical computing environments. The large memory footprint and outstanding memory bandwidth of the SPARC M7-16 server makes it ideal for deploying database and enterprise applications completely in memory for extreme performance and capacity.

Based on an innovative new CPU, memory, and I/O unit (CMIOU) chassis, SPARC M7-8 and M7-16 servers provide additional scalability, data center features, and functionality making them ideally suited for demanding business and data center applications. These servers scale effectively to support large processor and memory capacities, while offering additional features to extend reliability and availability. SPARC M7 processor–based servers are specifically designed with multiple options that allow their considerable resources to be partitioned and deployed as required, for example:

- The SPARC M7-8 server is offered in two factory-configured variants. It can be ordered with either one physical domain (PDom) or two separate and independent PDoms.
- The SPARC M7-16 server combines up to 16 SPARC M7 processors in one to four PDoms.

Server Components

SPARC M7-8 and M7-16 servers utilize many common components. The sections that follow provide an overview of the key components in these servers.

CPU, Memory, and I/O Unit Chassis

The SPARC M7-8 server is contained within one CMIOU chassis—the enclosure that houses the processor/memory boards, service processors (SPs), and connectors for the interconnect assemblies—and the SPARC M7-16 server includes two CMIOU chassis. The enclosure also includes power supplies and cooling fans. Figure 13 shows the front and rear view of the CMIOU chassis and it key components.

The front of the chassis features hot-swappable fan modules and power supply units (PSUs) as well as interconnect assemblies that connect the CMIOU boards together into a system. Front-facing features include the following:

- Eight hot-swappable fan modules
- Six hot-swappable N+N redundant PSUs, 3000 W at 200VAC–240VAC

The rear of the CMIOU chassis includes the following:

- Up to eight CMIOU boards
- Two redundant SPs

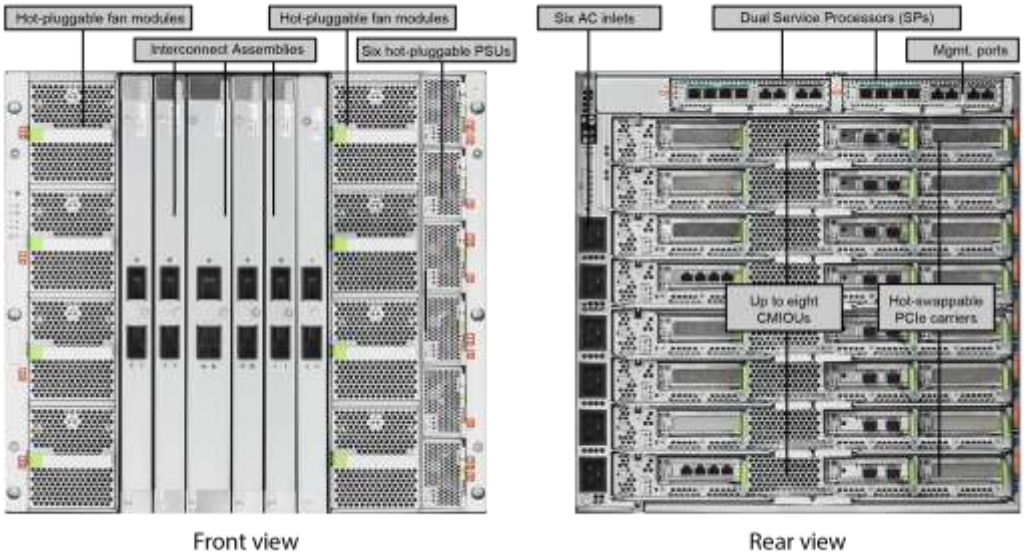

Figure 13. Front and rear perspectives of the CMIOU chassis.

CPU, Memory, and I/O Unit Board

Each CMIOU board assembly contains one SPARC M7 processor on a mezzanine board plus associated memory and I/O. All 16 memory DIMM slots are on the board. An I/O controller ASIC provides dedicated root complexes for three PCIe 3.0 (x16) slots. PCIe hot-pluggable carriers are included with the board. Figure 14 provides a block diagram of the CMIOU board, while Figure 15 provides a top-down photographic perspective. When inserted into the CMIOU chassis, the CMIOU board connects with the interconnect assemblies that provide the connectivity between the CMIOU boards, SPs, and the switch units (in SPARC M7-16 servers).

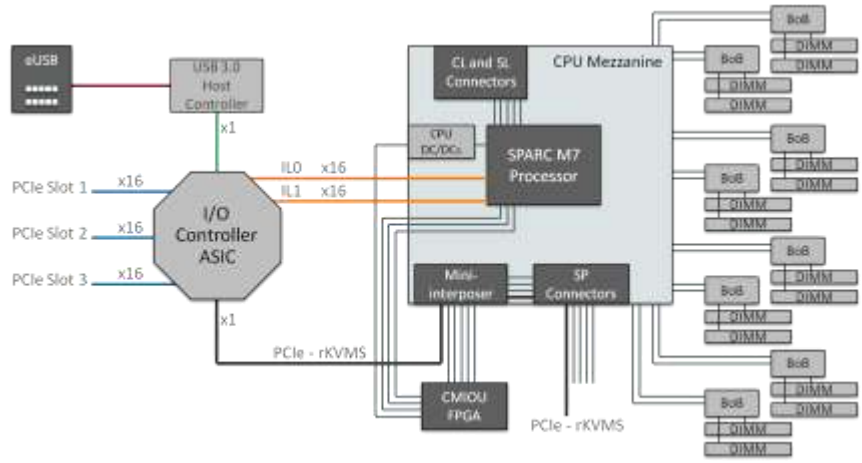

Figure 14. Architectural perspective of a single CMIOU board.

Figure 15. Each CMIOU board contains a SPARC M7 processor, 16 DDR4 DIMM slots, an I/O controller ASIC, and three PCIe x16 hot-pluggable carriers.

Interconnect Assembly

The systems interconnect for SPARC M7-8 and M7-16 servers is implemented with the interconnect assemblies. The SPARC M7-8 server uses interconnect assemblies that are contained within the 10U space of the CMIOU chassis. Five interconnect assemblies provide the coherency link (CL) connections (glueless systems interconnect) between the CMIOU boards. One interconnect assembly is used for communication between the CMIOU boards and the dual SPs at the top of the CMIOU chassis. The wiring in the interconnect assemblies is different for the two variants of the SPARC M7-8 server in order to support either one or two electrically isolated PDoms.

In the SPARC M7-16 server, each CMIOU chassis includes one interconnect assembly that provides the coherency links for the glueless CPU-to-CPU connections, and an SP interconnect assembly for communication between the CMIOU boards and the dual SPPs. Each of the two CMIOU chassis are also connected with four interconnect assemblies to the switch chassis. These eight interconnect assemblies provide the scalability links (SL) that form the glued systems interconnect in the SPARC M7-16 server.

Switch Chassis and Switch Unit

The switch chassis is used only in the SPARC M7-16 server. It houses six switch units (SWUs) that provide the glue for the systems interconnect in the server. Each SWU includes dual switch ASICs that are connected to every SPARC M7 processor in the server. A total of eight interconnect assemblies are used to provide the connectivity to the CMIOU chassis. Four interconnect assemblies connect to the bottom CMIOU chassis and the other four connect to the top CMIOU chassis in the SPARC M7-16 server. One SP interconnect assembly is used for communication between the SWUs and the SPs in the switch chassis. Figure 16 shows the front and rear view of the switch chassis.

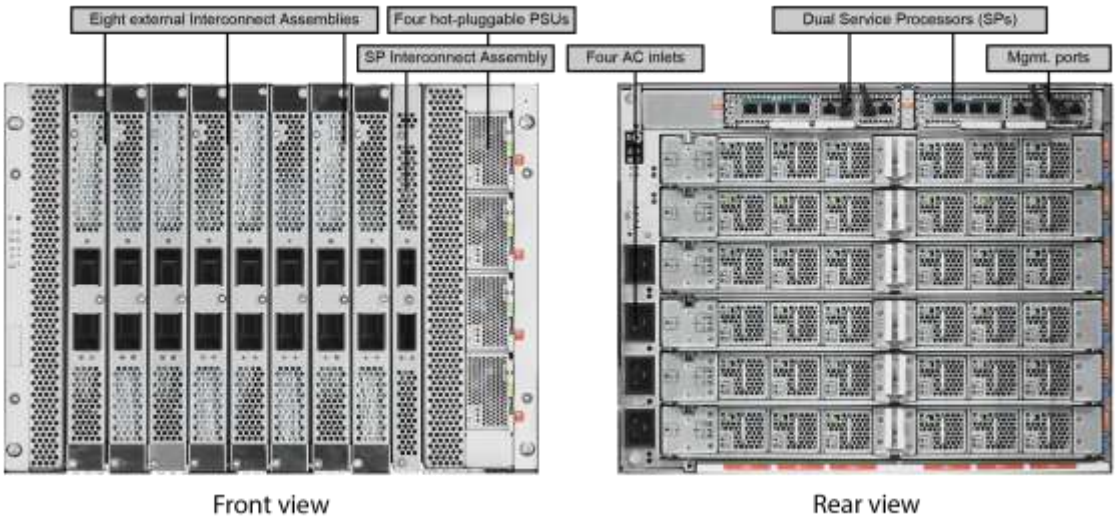

Figure 16. SPARC M7-16 switch chassis front and rear perspectives.

Service Processor, Service Processor Proxy, and Service Processor Module

The SPARC M7-8 and M7-16 servers feature redundant hot-pluggable service processors (SPs). There are two SPs located in the CMIOU chassis and two in the switch chassis. For the SPARC M7-16 server, the SP hardware in the CMIOU chassis is used as Service Processor Proxies (SPPs), and the SPs in the switch chassis become system SPs. Each SP has one serial (RJ45) and one 1000BASE-T management port (RJ45, autonegotiates to 10/100/1000 Mb/sec). The SP or SPP communicates with the CMOIU boards or switch boards via the SP interconnect assembly. The SPPs in the SPARC M7-16 CMIOU chassis are connected to the SPs in the switch chassis via network cables in the rear of the chassis.

Each SP or SPP includes one or two service processor modules (SPMs). The SPM is the component that runs the Oracle ILOM software and provides the SP functionality for the server system. In order to always have redundant SP functionality and failover capability, some configurations include dual SPMs in each SP. The following describes the variation depending on the server model:

- In the SPARC M7-8 server with a single PDom, each of the two SPs has only one SPM. This is sufficient to provide redundancy because the server that has a single PDom. One of the SPs functions as the active SP to manage the platform, and the other acts as a standby SP that assumes the active SP role in the event of a failure.
- In the SPARC M7-8 server with dual PDoms, each SP has two SPMs, and the server uses two pairs of SPMs to manage the two PDoms in the system. Two SPMs provide the active SP functionality, while the remaining two are standby. The active and standby SPMs are located in separate SPs.
- In the SPARC M-16 server, each SP on the switch chassis has only one SPM, and each SPP has two SPMs. This configuration is necessary because the system is capable of redundant support for up to four PDoms.

System Rack and Power Distribution Units

The SPARC M7-8 and M7-16 servers are factory-configured into a system rack that has the same form factor as Oracle's Sun Rack II 1242. The rack includes two power distribution units (PDUs) for a three-phase dual-grid power source. The PDUs have single-phase outputs that feed power to the CMIOU chassis and switch chassis. PDUs include a power monitoring module with one serial (RJ45) and one 100BASE-T (RJ45) management port.

Oracle recommends that the SPARC M7-8 server be ordered factory-configured in the system rack. However, this system is optionally available as a standalone system to be rackmounted onsite. The standalone SPARC M7-8 server enclosure is 10 units (10U) high. In addition, 3U is reserved for the three-phase power cables, either at the bottom or at the top of the rack. The PDUs are mounted vertically in the cabinet and do not consume units. Twenty-nine units are available for other devices when the SPARC M7-8 server is rackmounted in the Sun Rack II 1242. There is no available space for additional devices in the SPARC M7-16 server system.

SPARC M7-8 Server

Both variants of the SPARC M7-8 server support up to total of eight processors. The SPARC M7-8 server with a single PDom is configured as a single symmetric multiprocessing (SMP) server. The SPARC M7-8 server with dual PDoms consists of two electrically isolated hardware partitions. All SPARC M7-8 server hardware is contained in a single CMIOU chassis, which Oracle recommends be factory-mounted in the rack along with dual PDUs. Most hardware components in the two variants of the SPARC M7-8 server are identical. However, the wiring in the interconnect assemblies is different to support either one up to 8-way or two isolated 4-way PDoms. This wiring is set at the factory and cannot be changed later. Also, the service processor hardware varies based on how many PDoms are supported with redundancy. Table 6 shows the main characteristics of the SPARC M7-8 server.

TABLE 6. SPARC M7-8 SERVER FEATURES.

Feature	Description
Max processors	8 or 2x 4
Max threads	2,048 or 2x 1,024
Max memory	4 TB or 2x 2 TB (based on 32 GB DIMMs)
Max PCIe 3.0 (x16) slots	24 or 2x 12
Service processors (SPs)	One PDom: Two hot-pluggable SPs, each with a single SP module (SPM) providing redundancy for one PDom Two PDoms: Two hot-pluggable SPs, each with dual SPMs providing redundancy for two isolated PDoms
Fans	Eight hot-swappable N + 1 fan modules with dual counter-rotating fans Front access
Power supplies	Six hot-swappable N+N redundant power supply units (3000 W each at 200VAC–240VAC); front access
PDUs	Two PDUs in the system rack. Three 3-phase power cords per PDU for dual-grid configuration. Six single-phase power cords from PDUs to the system chassis.

SPARC M7-8 Server with a Single Physical Domain

The SPARC M7-8 server with a single PDom is contained in a single CMIOU chassis. It is configured as a single PDom with up to eight SPARC M7 processors. Each SP in the SPARC M7-8 server includes a single service processor module (SPM), because only one pair of SPMs is required to provide redundancy for the SP function. By default, the first CMIOU board (in slot #0) is connected to SP0/SPM0 and is active. The CMIOU board in slot #1 provides the stand-by connection to SP1/SPM0. Figure 17 illustrates the high-level SPARC M7-8 server I/O device map, including a total of 24 PCIe 3.0 (x16) slots supported by the dedicated root complexes of the I/O controller ASIC.

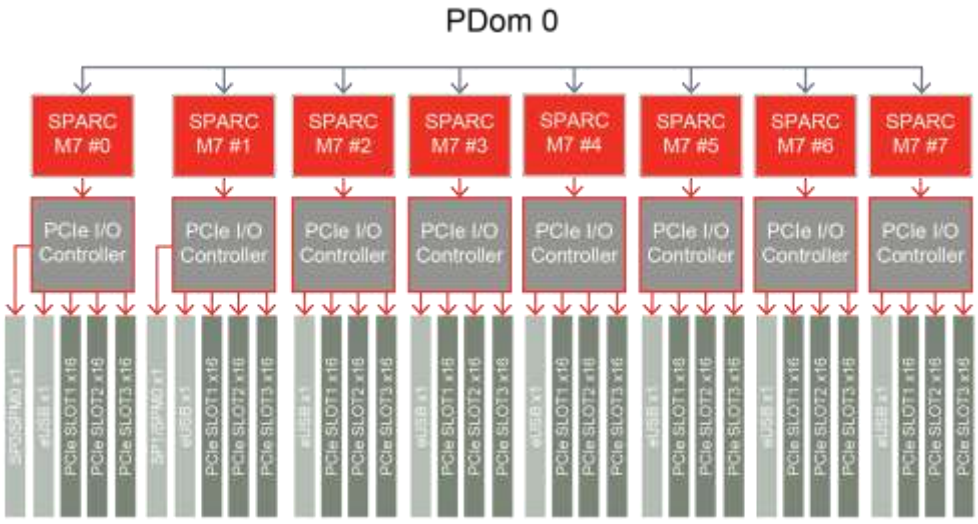

Figure 17. This SPARC M7-8 server includes eight SPARC M7 processors and is factory-configured with a single PDom.

The eight-way, glueless CPU-to-CPU fabric is composed of seven x16 coherency links (CLs) from each processor to the other seven processors (Figure 18).

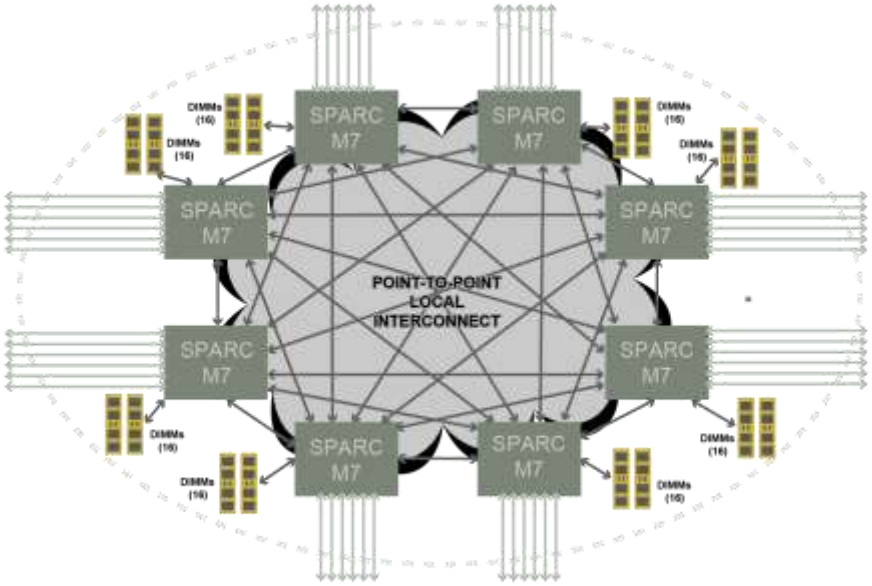

Figure 18. The eight-way all-to-all interconnect of the SPARC M7-8 server.

SPARC M7-8 Server with Two Physical Domains

The SPARC M7-8 server with two PDoms is also contained in a single CMIOU chassis, but it is electrically partitioned into two PDoms at the factory. The wiring of the interconnect assemblies is different from the wiring used in the SPARC M7-8 server variant that has only one PDom.

Each of the two SPs includes dual SPMs to provide redundant service processor functionality for two PDoms. A PDom is associated with one SPM in each of the service processors, which then act as an active/passive pair. The connections to the SPMs are from the first two CMIOU boards in each PDom. The SP active/passive functionality switches over to the standby connection if the active SPM or SP fails.

Figure 19 provides a high-level PCIe device map for the SPARC M7-8 server with two PDoms. The figure illustrates the connection between the PDoms and the SPs. By default, for PDom 0 the connection from the CMIOU board in slot #0 to SP0/SPM0 is active. The CMIOU board in slot #1 provides the standby connection to SP1/SPM0. Similarly, for PDom 1, the CMIOU boards in slots #4 and #5 have the active/standby connections to SP0/SPM1 and SP1/SPM1, respectively. The device map also shows the x16 connections to the PCIe 3.0 slots supported by the dedicated root complexes of the I/O controller ASIC.

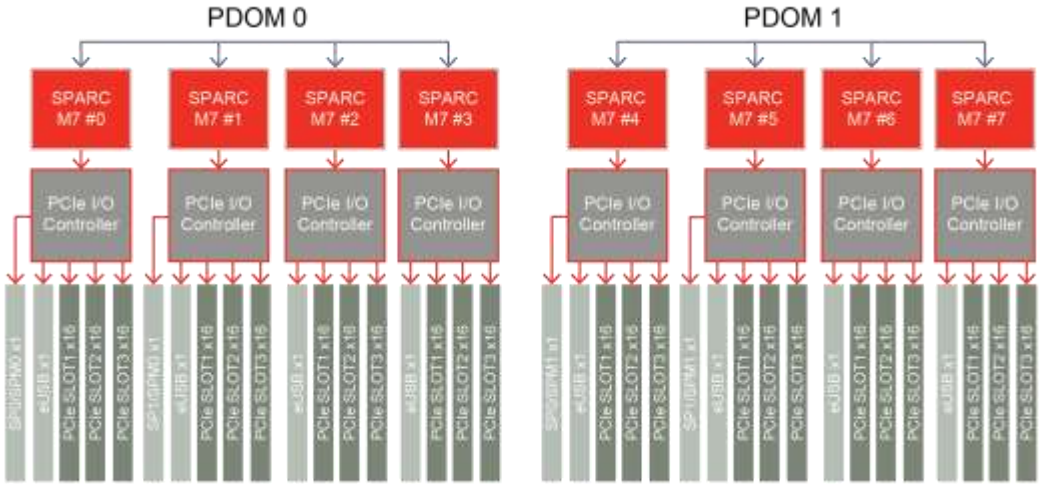

Figure 19. The SPARC M7-8 server can be ordered factory-configured with two electrically isolated PDoms.

Figure 20 illustrates the glueless systems interconnect in the SPARC M7-8 server with two PDoms. This system configuration has two separate four-way, all-to-all interconnects. The four-way CPU-to-CPU fabric consists of six x16 coherency links (CL) from each processor to all the other processors.

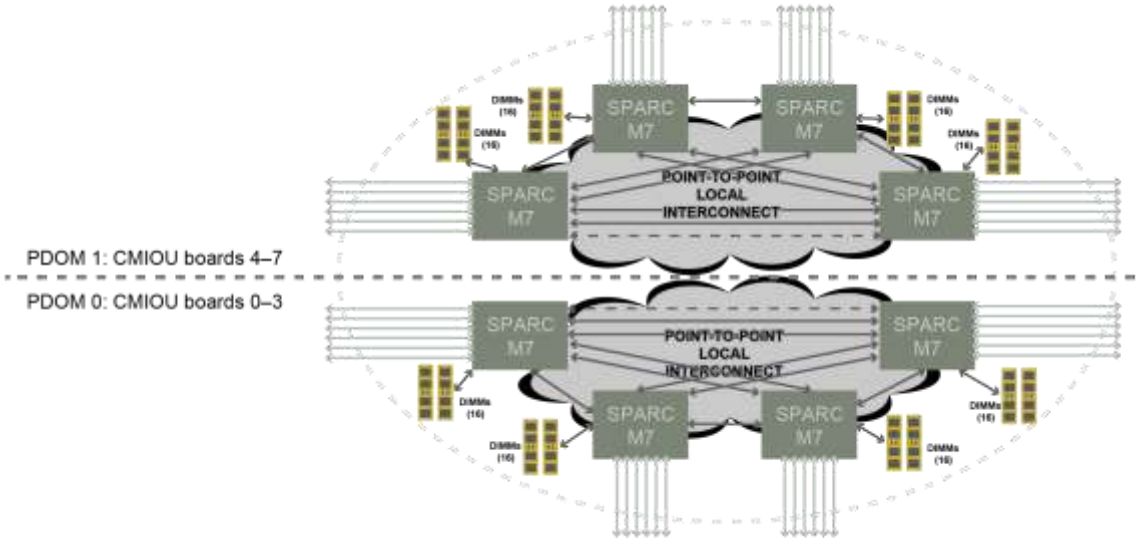

Figure 20. The SPARC M7-8 server with two PDoms is composed of two all-to-all interconnects.

SPARC M7-16 Server

The SPARC M7-16 server provides additional flexibility, scalability, and bandwidth with up to16 SPARC M7 processors, supporting up to 512 cores and 4,096 threads and up to 8 TB of memory in a single system. The SPARC M7-16 server is built at the factory into a system rack that includes two CMOIU chassis, one switch chassis, interconnect assemblies, and two power distribution units (PDUs). Figure 21 illustrates the front and rear perspectives of the SPARC M7-16 server.

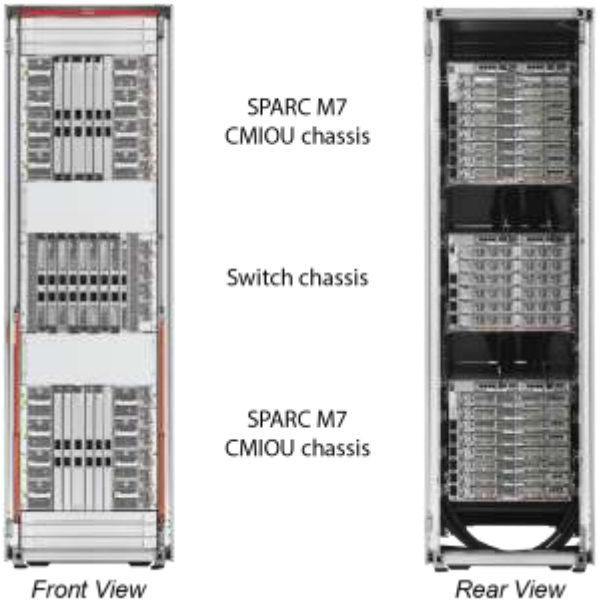

Figure 21. In the SPARC M7-16 server, two CMIOU chassis are connected via a switch chassis.

The SPARC M7-16 server can be partitioned into one, two, three, or four PDoms. The software-defined PDoms can be reconfigured, without making changes to the physical hardware, by accessing the service processor. The SPARC M7-16 server is divided into four domain configurable units (DCUs). Each DCU comprises four CMIOU slots (that is, half of a CMIOU chassis). The DCUs are the hardware building blocks of the PDoms. A PDom can consist of one, two, three, or all four DCUs.

The configuration of the connection from the PDoms (host) to the SPM depends on how the PDoms are defined. The first two CMIOU slots (lowest slot numbers) in each PDom establish the active/standby connections, respectively.

Figure 22 illustrates the high-level I/O device map for an example configuration with three PDoms. Each PDom has redundant SP connections to SPMs in separate SPPs. The three PCIe 3.0 (x16) slots per CMIOU board are also shown.

Figure 22. The SPARC M7-16 server is divided into four domain configurable units (DCUs) which can be flexibly combined to create one, two, three, or four PDoms (three PDoms shown).

The systems interconnect in the SPARC M7-16 server is based on a combination of coherence links (CL) and scalability links (SL). CLs are direct local (glueless) connections between the four processors in the DCUs. SLs connect processors via the switch ASIC located in the switch units, enabling scalability of up to 16 processors in a single PDom.

Figure 23 shows the coherency and data interconnect for the four-way DCU. The local four-way CPU-to-CPU fabric has three x12 CLs from each processor to the other three processors. The wiring for the CLs is physically located in the one interconnect assembly in each of the CMIOU chassis.

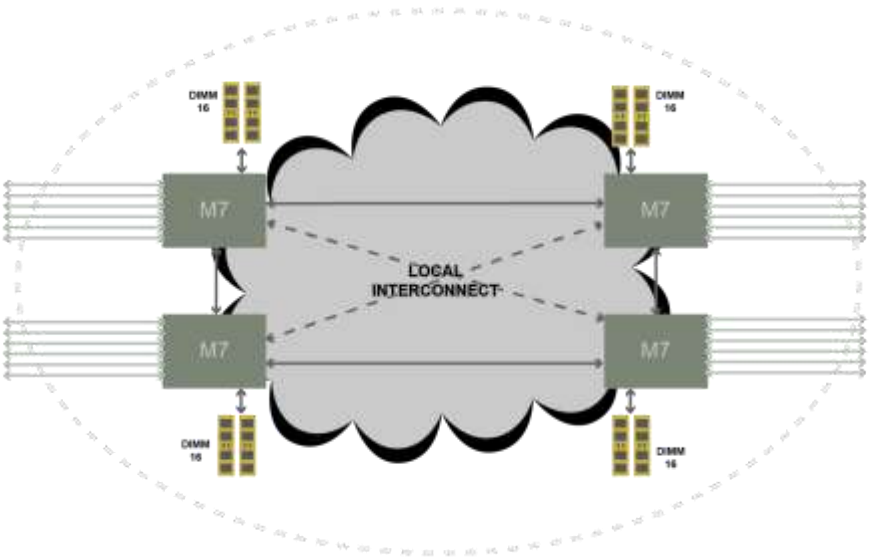

Figure 23. The four-way DCU in the SPARC M7-16 server has a local glueless CPU-to-CPU interconnect fabric using coherency links (CLs).

Figure 24 illustrates the global systems interconnect. In addition to the local CL connections described above, each processor is connected with two x4 SLs to each of the twelve switch ASICs in the switch chassis. This arrangement provides a fully connected all-to-all topology for the system, with two links connecting every processor to every switch ASIC, and a coherence directory distributed among switches. The wiring for the SLs is physically located in the eight interconnect assemblies that connect the CMIOU chassis to the switch chassis.

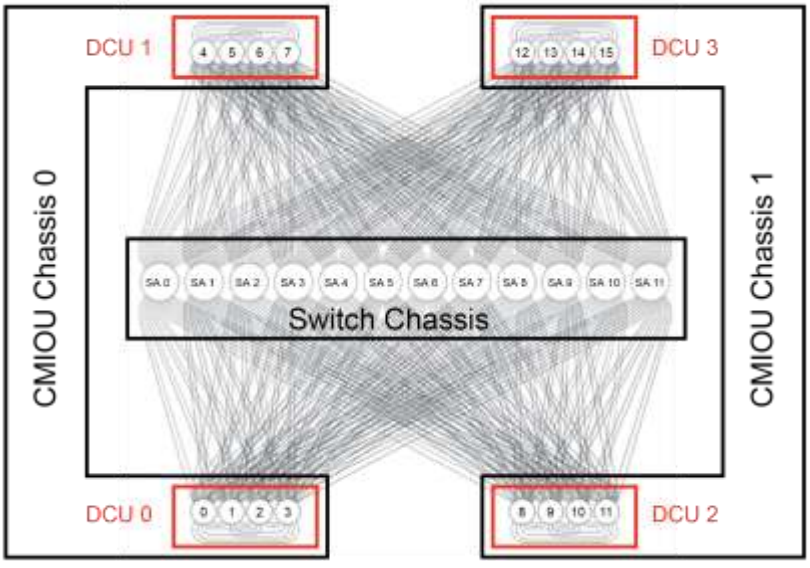

Figure 24. The 16-way SPARC M7-16 global systems interconnect uses scalability links and switch ASICs for the CPU-to-CPU fabric.

The SPs in the SPARC M7-16 server reside in the switch chassis, and each of them includes a single SPM. Only one pair of SPMs is required to provide redundancy for the SP function, because of the hierarchical SP infrastructure. Both CMIOU chassis include two SPPs, each with dual SPMs. This provides a redundant SP function that can recover from an SPM, SPP, or SP failure and continue to provide access to an SP separately for all PDoms.

Oracle Solaris

Oracle Solaris provides key functionality for virtualization, optimal use, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. Oracle Solaris runs on a broad range of SPARC- and x86-based systems, and compatibility with existing applications is guaranteed. The Oracle Solaris 11.3 release is specifically designed to take full advantage of the considerable resources of SPARC M7 processor–based servers, including the complete set of advanced Software in Silicon features. Oracle Solaris 10 support is available only in guest domains that exclusively use virtual disk, console, and network devices provided by service domains (for example, control, root, or I/O domains). A guest domain does not have physical I/O devices and relies on service domains. Table 7 provides the supported releases of Oracle Solaris 11 and 10 in various domains of SPARC M7 processor–based servers.

TABLE 7. SUPPORTED ORACLE SOLARIS RELEASES ON SPARC M7 PROCESSOR–BASED SERVERS.

Oracle Solaris Version	Control Domain	Root Domain	I/O Domain	Guest Domain
Oracle Solaris 11.3[1]	✓	✓	✓	✓
Oracle Solaris 10 1/13[2]				✓
Oracle Solaris 10 8/11[2]				✓
Oracle Solaris 10 9/10[2]				✓

1. Versions of Oracle Solaris 11 prior to 11.3 are not supported on SPARC M7 processor–based servers
2. Plus required patches

One of the most attractive features of systems based on the SPARC M7 processors is that they appear as a familiar SMP system to Oracle Solaris and the applications it supports. In addition, Oracle Solaris has incorporated many features to improve application performance on Oracle's multicore/multithreaded architectures.

» **OpenStack cloud management**. Oracle Solaris 11 includes a complete OpenStack distribution, allowing administrators to centrally share and manage data center resources—including infrastructure and virtualization offerings provided by other vendors—through a single management pane. Integrated into the core technology foundations such as Oracle Solaris Zones, the ZFS file system, Unified Archives, and comprehensive software defined networking, OpenStack on Oracle Solaris provides self-service computing, allowing IT organizations to deliver services in minutes rather than weeks, with enterprise-grade reliability, security, and performance.

- **Oracle Solaris software defined networking (SDN).** Oracle Solaris 11 enhances its existing, integrated software defined networking technologies to provide much greater application agility without the added overhead of expensive network hardware. It now enables application-driven, multitenant cloud virtual networking across a completely distributed set of systems; decoupling from physical network infrastructure; and application-level network service-level agreements (SLAs)—all built in as part of the platform. Enhancements and new features include the following:
  - Network virtualization with virtual NICs (VNICs) and virtual switching
  - Network resource management and integrated quality of service (QoS) to enforce bandwidth limits on VNICs and traffic flows
  - Cloud-readiness: a core feature of the OpenStack distribution included in Oracle Solaris 11
  - Application-driven, multitenant cloud virtual networking with Oracle Solaris Elastic Virtual Switch and Virtual Extensible LANs (VXLANs)
  - Application-level QoS with application-driven SDN
  - Tight integration with Oracle Solaris Zones and Oracle Solaris 10 Zones
- **Lifecycle management.** Oracle Solaris 11 includes a complete and integrated set of technologies for managing the software lifecycle of the platform. With support for secure end-to-end provisioning with the Automated Installer, failsafe software updates with the Image Packaging System, ZFS Boot Environments, rapid application deployment using Unified Archives, and a comprehensive compliance framework, Oracle Solaris 11 helps to increase productivity, reduce human error, and greatly reduce IT costs.
- **Accelerated cryptography**. Accelerated cryptography is supported through the cryptographic framework in Oracle Solaris as well as through the SPARC M7 processor. The SPARC M7 processor permits access to cryptographic cypher hardware implementations. For the first time, through user-level instructions, the cyphers are implemented within the appropriate pipeline itself rather than as a coprocessor. This means a more efficient implementation of the hardware-based cyphers as well as no privilege-level changes, resulting in a large increase in efficiency in cryptographic algorithm calculations. In addition, database operations can make much more efficient use of the various cryptographic cyphers that are implemented within the instruction pipeline itself.
- **Critical-thread optimization**. Oracle Solaris releases 11 and 10 permit either a user or a programmer to allow the Oracle Solaris Scheduler to recognize a critical thread by means of raising its priority to 60 or above through the use of either the command-line interface (CLI) or system calls to a function. The Oracle Solaris Scheduler will usually evenly distribute threads among all the cores on the system. However, when a process is running in FX60 mode, the scheduler will attempt to give that thread exclusive access to a whole core, and allocate the remaining runnable threads to the other available cores.
- **Multicore/multithreaded awareness**. Oracle Solaris releases 11 and 10 are aware of the SPARC M7 processor hierarchy, so the Oracle Solaris Scheduler can effectively balance the load across all available pipelines. Even though it exposes each of these processors as 256 logical processors, Oracle Solaris understands the correlation among core clusters, L2 and L3 cache hierarchies, and cores and the threads they support to provide a fast and efficient thread implementation.
- **Fine-granularity manageability**. For the SPARC M7 processor, Oracle Solaris releases 11 and 10 have the ability to enable or disable individual cores and threads (logical processors). In addition, standard Oracle Solaris features, such as processor sets, provide the ability to define a group of logical processors and schedule processes or threads on them.
- **Binding interfaces**. Oracle Solaris allows considerable flexibility in that processes and individual threads can be bound to either a processor or a processor set, as required or desired.
- **Support for virtualized networking and I/O**. Oracle Solaris contains technology to support and virtualize components and subsystems on the SPARC M7 processor. As part of a high-performance network architecture, Oracle multicore/multithreaded–aware device drivers are provided so that applications running within virtualization frameworks can effectively share I/O and network devices.

- **Non-uniform memory access (NUMA) optimization in Oracle Solaris**. With memory managed by each SPARC M7 processor, these implementations represent a NUMA architecture. In NUMA architectures, the time needed for a processor to access its own memory is slightly shorter than that required to access memory managed by another processor. Oracle Solaris provides the following technology, which can specifically help to decrease the impact of NUMA on applications and improve performance on NUMA architectures:
  - **Memory placement optimization (MPO)**. Oracle Solaris uses MPO to improve the placement of memory across the physical memory of a server, resulting in increased performance. Through MPO, Oracle Solaris helps ensure that memory is as close as possible to the processors that access it, while still maintaining enough balance within the system. As a result, many database applications are able to run considerably faster with MPO.
  - **Hierarchical Lgroup Support (HLS)**. HLS improves the MPO feature in Oracle Solaris by optimizing performance for systems with more-complex memory latency hierarchies. HLS lets Oracle Solaris distinguish between the degrees of memory remoteness, allocating resources with the lowest-possible latency for applications. If local resources are not available by default for a given application, HLS helps Oracle Solaris allocate the nearest remote resources.
- **Oracle Solaris ZFS**. Oracle Solaris ZFS offers a dramatic advance in data management, automating and consolidating complicated storage administration concepts and providing unlimited scalability with the world's only 128-bit file system. Oracle Solaris ZFS is based on a transactional object model that removes most of the traditional constraints on I/O issue order, resulting in dramatic performance gains. Oracle Solaris ZFS also provides data integrity, protecting all data with 64-bit checksums that detect and correct silent data corruption.
- **Multipathing software**. Mutipathing software in Oracle Solaris allows organizations to define and control redundant physical paths to I/O devices such as storage devices and network interfaces. If the active path to a device becomes unavailable, the software can automatically switch or fail over to an alternate path to maintain availability. To take advantage of multipathing capabilities, the server must be configured with redundant hardware, such as redundant network interfaces or two host bus adapters connected to the same dual-ported storage array.
- **A secure and robust enterprise-class environment**. Existing SPARC applications continue to run unchanged on SPARC T7 and M7 platforms, protecting software investments. Certified multilevel security protects Oracle Solaris environments from intrusion. The Fault Management Architecture in Oracle Solaris means that elements such as Oracle Solaris Predictive Self Healing can communicate directly with the hardware to help reduce both planned and unplanned downtime. Effective tools, such as Oracle Solaris DTrace, help organizations tune their applications to get the most out of the system's resources.

Virtualization

Virtualization is an essential technology as organizations strive to consolidate disparate workloads onto fewer, more-powerful systems, while increasing utilization. SPARC M7 processor–based servers contain built-in virtualization capabilities: multiple PDoms (in SPARC M7-8 and M7-16 servers), Oracle VM Server for SPARC, as well as support for a third layer in the form of OS-based virtualization with Oracle Solaris Zones.

- Physical domains (PDoms) are used to divide a single hardware system into multiple, security- and fault-isolated servers or hardware partitions.
- Logical domains (LDoms) are created using Oracle VM Server for SPARC and are used to virtualize a server or physical domain to host multiple virtual machines (VMs), each running its own instance of Oracle Solaris. Oracle VM Server for SPARC is a free-of-charge feature that is included in all of Oracle's SPARC servers.
- Oracle Solaris Zones enable OS virtualization so that a single instance of Oracle Solaris can securely isolate applications from each other and allocate system resources to each zone. This essentially allows the creation of multiple virtual machines within a single instance of the Oracle Solaris operating system.

These virtualization technologies are complimentary to each other. Indeed, best practices often include layered virtualization where two or all three technologies are deployed to achieve optimal security, availability, performance and manageability. Furthermore, Oracle's virtualization technologies are included at no incremental cost. The virtualization technologies available in SPARC M7 processor–based servers are shown in Figure 25 and described in more detail in the information that follows the figure.

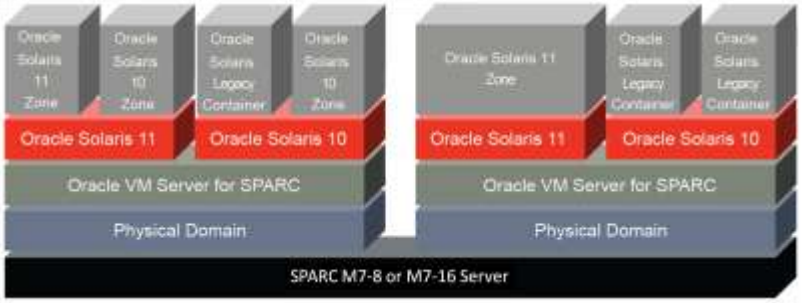

Figure 25. SPARC M7-8 and M7-16 servers can provide multiple PDoms while SPARC T7-1, T7-2, T7-4, and M7-8 servers offer a single PDom.

» **Physical domains (PDoms)**. PDoms on SPARC M7-8 and M7-16 servers provide IT organizations with the ability to divide a single hardware system into multiple, security- and fault-isolated servers. The SPARC M7-8 server can be ordered from the factory with two PDoms. SPARC M7-16 servers can be reconfigured via software to include one, two, three, or four PDoms. With proper configuration, hardware or software faults in one domain remain isolated and unable to impact the operation of other domains.

» **Oracle VM Server for SPARC**. Like prior generations of SPARC processors, the SPARC M7 processor supports a hypervisor that enables the creation of logical domains (LDoms) within a single PDom, or within a single server in the case of SPARC T7-1, T7-2, T7-4, and M7-8 servers. This hypervisor differs from traditional ones, because it is a small firmware layer that provides a stable virtual machine architecture that is tightly integrated with the processor. Multithreading is crucial, because the hypervisor interacts directly with the underlying multicore/multithreading processor. Supported in all servers that use Oracle's multicore/multithreaded technology, Oracle VM Server for SPARC provides full virtual machines that run an independent operating system instance. These full virtual machines can be configured as root domains—in which case, they are assigned root complexes for direct access to I/O—or as guest domains, in which case, they access virtualized I/O devices. Each operating system instance contains processor, memory, storage, console, and cryptographic devices.

» **Oracle Solaris Zones**. Oracle Solaris 11 provides a unique partitioning technology called Oracle Solaris Zones. This technology can be used to create an isolated and secure environment for running applications. An Oracle Solaris Zone is a virtualized operating system environment created within a single instance of Oracle Solaris. Oracle Solaris Zones can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability, because processes in one Oracle Solaris Zone are prevented from interfering with processes running in another Oracle Solaris Zone. Virtual CPUs in a multiprocessor system (or threads in the SPARC M7 processor) can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to an Oracle Solaris Zone. Resource pools provide the capability to separate workloads so that the consumption of CPU resources does not overlap. They also provide a persistent configuration mechanism for processor sets and scheduling class assignment. In addition, the dynamic features of resource pools enable administrators to adjust system resources in response to changing workload demands. Oracle Solaris 11 provides Immutable Zones, which preserve the zones' configuration by implementing a read-only root file system for non-global zones. Unless performed as specific maintenance operations, modifications to system binaries or system configurations are blocked.

Oracle's layered virtualization technologies can be used together to create resilient high-availability systems. For example, Oracle engineered systems and Oracle Maximum Availability Architectures employ these virtualization technologies and virtualization best practices in order to achieve very high reliability, availability, and serviceability (RAS). To read more about virtualization and RAS in SPARC M7 processor–based servers please refer to table 9 on page 41.

Systems Management

Providing hands-on, local system administration for server systems is no longer realistic for most organizations. Around the clock system operation, disaster recovery hot sites, and geographically dispersed organizations lead to requirements for remote management of systems. One of the many benefits of Oracle servers is the support for lights-out data centers, enabling expensive support staff to work in any location with network access. The design of the SPARC M7 processor–based servers combines with powerful service processors (SPs) running the Oracle Integrated Lights Out Manager (Oracle ILOM) software; this, along with Oracle Enterprise Manager Ops Center software, helps administrators remotely execute and control nearly any task that does not involve physical access to hardware. These management tools and remote functions lower administrative burden, saving organizations time and reducing operational expenses.

Oracle ILOM and Service Processor

The Oracle ILOM software on each SP provides the heart of remote monitoring and management capabilities for SPARC servers. The SP consists of a dedicated processor that is independent of the server system and runs the Oracle ILOM software package. The larger SPARC M7-8 and M7-16 servers feature redundant SPs capable of automatic failover and support hot-serviceability for continuous operation. While input power is supplied to the server, the SP constantly monitors the system even if all domains are inactive.

The SP regularly monitors the environmental sensors, provides advance warning of potential error conditions, and executes proactive system maintenance procedures, as necessary. For example, the SP can initiate a server shutdown in response to temperature conditions that might induce physical damage to the system. The Oracle ILOM software package running on the SP helps administrators to remotely control and monitor physical domains and virtual machines, as well as the hardware platform itself.

Using a network or serial connection to the SP, operators can effectively administer the server from anywhere on the network. Remote connections to the SP run separately from the operating system and provide the full control and authority of a system console. The Oracle ILOM SP acts as a system controller, facilitating remote management and administration. The SP is full-featured and is similar in implementation to that used in Oracle's other servers. As a result, the server integrates easily with existing management infrastructure. Critical to effective system management, the Oracle ILOM SP does the following:

- Implements an IPMI 2.0–compliant SP, providing IPMI management functions to the server's firmware, OS, and applications and to IPMI-based management tools accessing the SP via the Oracle ILOM 3.2 Ethernet management interface. The SP also provides visibility to the environmental sensors on the server module and elsewhere in the chassis.
- Manages inventory and environmental controls for the server, including processors, DIMMs, fans, and power supplies and provides HTTPS, CLI, and SNMP access to this data.
- Supplies remote textual console interfaces.
- Provides a means to download upgrades to all system firmware.

The Oracle ILOM and SP also allow administrators to remotely manage the server, independent of the operating system running on the platform and without interfering with any system activity. Oracle ILOM can send e-mail alerts about hardware failures, warnings, and other events related to the server. Its circuitry runs independently from the server, using the server's standby power. As a result, Oracle ILOM firmware and software continue to function when the server operating system goes offline or when the server is powered off. Oracle ILOM monitors the following server conditions:

- CPU temperature conditions
- Hard drive presence
- Enclosure thermal conditions

- Fan speed and status
- Power supply status
- Voltage conditions
- Oracle Solaris Predictive Self Healing, boot timeouts, and automatic server restart events

In the SPARC M7-16 server, the PDom configuration is managed on the SP using Oracle ILOM commands via the CLI or in a web browser.

Power Management

Power and cooling costs for servers are becoming significant, and lowering these costs is a top challenge in the corporate data center. Limitations in the availability of power and space to expand data centers force customers to look closely at the power efficiency of servers. Contracts with power providers, which specify penalties for exceeding the stated power consumption, require servers to be able to cap their power consumption under customer control. Power efficiency and carbon footprint have become factors when customers evaluate servers.

Beyond the inherent efficiencies of Oracle's multicore/multithreaded design, the SPARC M7 processors incorporate unique power management features at both the core and memory levels of the processor. These features include reduced instruction rates, parking of idle threads and cores, and the ability to turn off clocks in both cores and memory to reduce power consumption. In addition, while previous SPARC processors allowed power management at the chip level, the SPARC M7 processors allows for subchip power management. The SPARC M7 processor is divided into four quadrants, of eight cores each. Each quadrant can set different power consumption levels independently of other quadrants. This allows for the possibility of different performance levels per quadrant and, importantly, different SLAs when deploying LDoms in different quadrants.

In addition to the power management support in Oracle ILOM, Oracle Solaris 11.3 or later provides a power manager that supports SPARC M7 processor–based servers. Oracle Solaris can determine which power savings features to enable based on the `poweradm` settings, which are set by the platform based on the system (Oracle ILOM) policy but can be overridden by the Oracle Solaris administrator. Substantial innovation is present in the following areas:

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

In a virtualized environment using Oracle VM Server for SPARC, the power management manager performs the following tasks when managing LDom guests:

- Determining which power savings features to enable based on the power management policy
- Calling the Power Management engine to initiate power state changes on its resources to achieve a power adjustment or utilization level (for resources not owned by Oracle Solaris 11.3 guests) or telling the hypervisor to enable or disable hypervisor/hardware–managed power states. Only Oracle Solaris 11.3 guests have a power management peer.

Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center delivers a converged hardware management solution for SPARC M7 processor–based servers that integrates management across the infrastructure stack. With advanced virtualization management and reporting capabilities, application-to-disk management, intelligent configuration management, and more, Oracle Enterprise Manager Ops Center helps IT managers reduce complexity and streamline and simplify infrastructure management. The inclusion of Oracle Enterprise Manager Ops Center with every SPARC server enables data center administrators to monitor and manage the storage, network, servers, Oracle Solaris, and virtualized environments from a single interface. This improves operational efficiency and lowers operational costs.

Oracle Enterprise Manager Ops Center is the most comprehensive management solution for Oracle servers and Oracle engineered systems infrastructure. Offering a single console to manage multiple server architectures and myriad operating systems, Oracle Enterprise Manager Ops Center can manage the components in SPARC M7 processor–based servers using asset discovery, the provisioning of firmware and operating systems, automated patch management, patch and configuration management, virtualization management, and comprehensive compliance reporting (Figure 26).

Oracle Enterprise Manager Ops Center automates workflow and enforces compliance via policy-based management—all through a single, intuitive interface. With Oracle Enterprise Manager Ops Center, IT staff can implement and enforce data center standardization and best practices, regulatory compliance, and security policies while efficiently deploying infrastructure to meet business requirements.

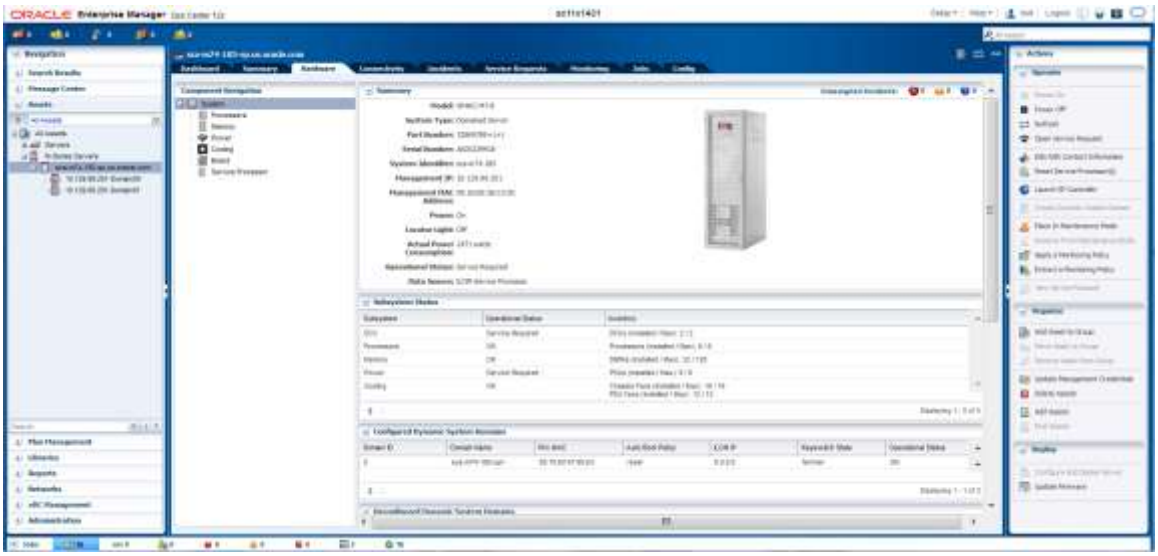

Figure 26. Oracle Enterprise Manager Ops Center provides detailed management capabilities for SPARC servers.

Reliability, Availability, and Serviceability

Reducing downtime—both planned and unplanned—is critical for IT services. System designs must include mechanisms that enable fault resilience, quick repair, and even rapid expansion without impacting the availability of key services. Specifically designed to support complex network computing solutions and stringent high availability (HA) requirements, SPARC M7 processor–based servers include redundant and hot-serviceable system components, diagnostic and error recovery features throughout the design, and built-in remote management features. The advanced architecture of these reliable servers fosters high levels of application availability and rapid recovery from many types of hardware faults, simplifying system operation and lowering costs for enterprises.

Advanced Reliability Features

Advanced reliability features included within the components of SPARC M7 processor–based servers increase the overall stability of the platform. Reduced component count and complexity within the server architecture contributes to reliability. In addition, advanced CPU integration and guaranteed data path integrity provide for autonomous error recovery by the SPARC M7 processors, reducing the time to initiate corrective action and subsequently increasing uptime. Redundancy in key components allows for failover. Hot-pluggable and hot-swappable components allow for increased serviceability. With SPARC M7 processor–based servers, advanced I/O controllers host the PCIe root

complexes, allowing them to remain unchanged if a processor is added or removed. Together these features present a very robust system design.

The Fault Management Architecture (FMA) that is implemented both in firmware and in the Oracle Solaris Predictive Self Healing software further enhances the reliability of SPARC servers. The FMA provides constant monitoring of processors, memory, and I/O devices. Depending upon the nature of the error, persistent CPU soft errors can be resolved by automatically off-lining a thread, a core, or an entire processor. In addition, the memory page retirement function supports the ability to take memory pages offline proactively in response to multiple corrections to data for a specific memory DIMM. As described earlier, SPARC M7 processor–based servers support DIMM sparing that allows an entire DIMM to be retired without interrupting system operation, and while memory capacity and error protection capability remain intact.

Error Detection, Diagnosis, and Recovery

SPARC M7 processor–based servers feature important technologies that correct failures early and keep marginal components from causing repeated downtime. Architectural advances that inherently increase reliability are augmented by error detection and recovery capabilities within the server hardware subsystems. Ultimately, the following features work together to raise application availability:

- End-to-end data protection detects and corrects errors throughout the system, ensuring complete data integrity.
- State-of-the-art fault isolation helps the server isolate errors within component boundaries and take offline only the relevant chips, or part of a chip, instead of the entire component. Isolating errors down to the smallest possible entity improves stability and provides continued availability of maximum compute power. This feature applies to processors, memory ASICs and DIMMs, switch ASICs, connectivity links, I/O controller ASICs, and SPs.
- Constant environmental monitoring provides a historical log of pertinent environmental and error conditions.
- The host watchdog feature periodically checks for the operation of software, including the domain operating system. This feature also uses the SP firmware to trigger error notification and recovery functions.
- The FMA capability of the system and dynamic CPU resource deconfiguration of the SPARC M7 processor enable powerful isolation and recovery. If necessary, the system can dynamically retire processor resources (for example, a core) without interrupting the applications that are running.
- Periodic component status checks are performed to determine the status of many system devices to detect signs of an impending fault. Recovery mechanisms are triggered to prevent system and application failure.
- Error logging, multistage alerts, electronic field replaceable unit (FRU) identification information, and system fault LED indicators contribute to rapid problem resolution.

Redundant and Hot-Serviceable Components

Today's IT organizations are challenged by the pace of nonstop business operations. In a networked global economy, revenue opportunities remain available around the clock, forcing planned downtime windows to shrink and, in some cases, disappear entirely. To meet these demands, the SPARC M7 processor–based servers employ built-in redundant, hot-pluggable, and hot-swappable hardware to help mitigate the disruptions caused by individual component failures or changes to system configurations. In fact, these systems are often able to recover from hardware failures—often with no impact to users or system functionality.

The SPARC M7 processor–based servers feature redundant, hot-swappable power supply and fan units, as well as the option to configure multiple processors, memory DIMMs, and I/O cards. Administrators can create redundant internal storage by combining hot-pluggable disk drives with disk mirroring software. SPARC M7-8 and M7-16 servers also support redundant, hot-pluggable SPs and SPPs. If a fault occurs, these duplicated components support continued operation. Depending upon the component and type of error, the system might continue to operate in a degraded mode or reboot—with the failure automatically diagnosed and the relevant component

automatically configured out of the system. In addition, hot-serviceable hardware within these servers speeds service and allows for the simplified replacement or addition of components, without a need to stop the system.

Replaceable components fall into the following categories:

- **Hot-serviceable**. Hot-serviceable components can be removed and inserted while the server is running.
  - Hot-swappable components do not require any preparation prior to servicing.
  - Hot-pluggable components require preparation prior to servicing.
- **Cold-serviceable**. Cold-serviceable components require that the server be shut down for service. In addition, some service procedures require that the power cables be disconnected between the power supplies and the power source.

Table 8 lists the key hot-serviceable components in the SPARC M7 processor–based servers.

TABLE 8. KEY CUSTOMER HOT-SERVICEABLE COMPONENTS IN SPARC M7 PROCESSOR–BASED SERVERS.

Component	SPARC T7-1 Server	SPARC T7-2 Server	SPARC T7-4 Server	SPARC M7-8 Server	SPARC M7-16 Server
2.5-inch SFF Drive	✓	✓	✓	NA	NA
Power Supply	✓	✓	✓	✓	✓
Fan/Fan Module	✓	✓	✓	✓	✓
PCIe Card	–	–	✓ [1]	✓	✓
CMIOU Board	NA	NA	NA	✓ [2]	✓ [2]
Switch Unit (SWU)	NA	NA	NA	NA	✓ [3]
SP/SPP	–	–	–	✓	✓

✓ = Hot-serviceable, – = Gold-serviceable, NA = Not applicable

1. PCIe card in the card carriers. NVMe PCIe switch card in internal slots are cold-serviceable.
2. The PDom which the CMIOU board is in must be shut down. Other PDoms can be operational.
3. The SWU must be unconfigured, and system must have been booted without the SWU. All PDoms can be operational during the SWU hot-service. Reconfiguration to include the SWU requires rebooting of all PDoms.

Conclusion

Modern technology and business challenges require innovative solutions, significant compute power, and a range of server form factors to deliver on the promise of a rapidly evolving digital marketplace. Based on the strengths of Oracle's innovative SPARC M7 processor, Oracle's SPARC servers provide a new approach that can help organizations efficiently deliver their most demanding computational workloads in a resilient and highly available fashion. Through specialized acceleration engines on each SPARC M7 processor, innovative Software in Silicon technology provides Silicon Secured Memory, In-Memory Query Acceleration, In-Line Data Decompression, and accelerated encryption operations for enhanced security, which can make a dramatic difference for applications.

The extensive range of power that SPARC M7 processor–based servers provide lets organizations shift computing costs away from underlying infrastructure and focus on the productive business use of their workloads. For example, new levels of performance and efficiency let organizations run both OLTP and data analytics on the same system. Coupled with Oracle Solaris functionality, Oracle's layered virtualization lets organizations create the most efficient on-premises clouds, with superior capacity and scalability, helping to enable accelerated time to market and enhancing the bottom line.

For More Information

For more information, visit the resources listed in Table 9.

TABLE 9. RESOURCES.

Web Resources	**URL**
Oracle's SPARC Servers	oracle.com/us/products/servers-storage/servers/sparc/oracle-sparc and oracle.com/technetwork/server-storage/sun-sparc-enterprise
Oracle Solaris	oracle.com/solaris
Oracle Optimized Solutions	oracle.com/optimizedsolutions
Oracle Benchmark Results	oracle.com/benchmarks
Oracle Help Center Documentation	docs.oracle.com
Security	oracle.com/technetwork/topics/security
Virtualization	oracle.com/technetwork/topics/virtualization
Oracle Integrated Lights Out Manager	docs.oracle.com/en/hardware/?tab=4
Oracle Enterprise Manager Ops Center	oracle.com/technetwork/oem/ops-center
White Papers and Technical Articles	**URL**
"Oracle's SPARC T7 and SPARC M7 Server RAS"	oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/sparc-t7-m7-server-ras-2702879.pdf
"Oracle's SPARC T7 and SPARC M7 Servers: Domaining Best Practices"	oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/sparc-t7-m7-best-practices-2701865.pdf
"Consolidation Using Oracle's SPARC Virtualization Technologies"	oracle.com/technetwork/server-storage/sun-sparc-enterprise/technologies/consolidate-sparc-virtualization-2301718.pdf
"Best Practices for Oracle Solaris Kernel Zones"	oracle.com/technetwork/articles/servers-storage-admin/solaris-kernel-zones-best-practices-2400370.html
"Oracle VM Server for SPARC Best Practices"	oracle.com/technetwork/server-storage/vm/ovmsparc-best-practices-2334546.pdf
"Implementing Root Domains with Oracle VM Server for SPARC"	oracle.com/technetwork/server-storage/vm/ovm-sparc-rootdomains-wp-1914481.pdf
"Oracle Multitenant on SPARC Servers and Oracle Solaris"	oracle.com/technetwork/articles/servers-storage-admin/multitenant-on-sparc-solaris-2016889.html

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

ORACLE

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

Oracle's SPARC T7 and SPARC M7 Server Architecture
October 2015