



Oracle SPARC T8 和 SPARC M8 服务器架构

软件芯片化：实现安全基础设施，助力实时企业

ORACLE 白皮书 | 2017 年 9 月



ORACLE®

目录

引言	39
特性对比	3
SPARC M8 处理器	5
SPARC M8 处理器架构	6
处理器内核和缓存架构	7
软件芯片化技术	8
内存中查询加速	9
内联解压缩	10
Java 流加速	10
Oracle Numbers 加速	11
加密加速	11
芯片保护的内存	12
基于 SPARC M8 处理器的服务器系列概述	14
内存子系统	14
I/O 子系统	15
I/O 控制器 ASIC	15
NVM Express 技术	15
嵌入式 USB 存储和 Oracle Solaris 引导池	16
PCIe 适配器卡	16
SPARC T8-1、T8-2 和 T8-4 服务器	16
SPARC T8-1 服务器	17
SPARC T8-2 服务器	19



SPARC T8-4 服务器	21
SPARC M8-8 服务器	24
服务器组件	24
CPU、内存和 I/O 单元主板	25
互连配件	26
服务处理器和服务处理器模块	26
系统机架和配电装置	26
SPARC M8-8 服务器架构	26
只有一个物理域的 SPARC M8-8 服务器	27
具有两个物理域的 SPARC M8-8 服务器	28
Oracle Solaris	30
虚拟化	32
系统管理	34
Oracle ILOM 和服务处理器	34
电源管理	35
Oracle Enterprise Manager Ops Center	36
可靠性、可用性和可维护性	36
高级可靠性特性	37
错误检测、诊断和恢复	37
冗余的热维护式组件	37
总结	38
更多信息	39



引言

现代科技创举正在推动 IT 基础设施朝着新的方向发展。在快速发展的数字市场中，大数据、社交业务、移动应用、云和实时分析都需要前瞻性的解决方案和充足的计算能力来提供所需性能。如今，在客户的推动下，业务速度日益加快，组织需要以客户的方式与他们互动。这就需要以高水平的安全性来管理敏感信息，随时捕获和分析海量数据并根据这些数据采取行动。

与过去几十年相比，这些挑战将极大地改变 IT 系统的设计、投资和运行方式。目前，数据库和 Java 已经成为编写现代云就绪应用的事实标准语言。数据的体量、种类和速度呈爆炸式增长，这让组织更加迫切地需要通过安全、有效的分析来更快地做出更加明智的决策。就在组织迫于压力亟需削减成本、提高运营效率和提供创新性技术来创造新收入来源的同时，复杂的 IT 基础设施正变得越来越难以维护，维护成本也越来越高，严重阻碍了这些目标的实现。此外，客户无法再将数据安全性视为能事后补救的事情，因为每年因计算机入侵造成的损失高达数十亿美元。



Oracle 新的基于 SPARC M8 处理器的服务器凭借其软件芯片化技术将服务器技术提升到了新的水平。这些服务器提供对数据库和 Java 应用的加速及安全性，提供数据分析加速 (DAX) 单元对压缩数据进行内存中分析，还提供芯片保护的内存 (Silicon Secured Memory) 技术和全速宽密钥加密。SPARC M8 处理器进一步提高了效率，其架构支持多达 256 个线程，而规格大小却只有 2U。此外，SPARC M8 处理器还提高了每线程性能，改善了可靠性、可用性和可维护性 (RAS) 功能，提高了能源效率，并且内存和 I/O 带宽均高于竞争对手的设计。

Oracle 基于 SPARC M8 处理器的服务器为实时企业带来了极其高效的平台，助力它们在当今的数字化市场中赢得竞争、节省资金和时间并提高盈利能力。同时，Oracle 的技术创新能够为组织创造价值，帮助它们降低成本和提高投资回报。依托于 SPARC M8 处理器的先进技术，Oracle 新推出的 SPARC 服务器系列（图 1）可以从 1 个 SPARC M8 处理器扩展至 8 个 SPARC M8 处理器。这些服务器构成了一组灵活且可扩展的产品系列，它们具有极高的集成度，有助于增强安全性、降低成本以及提高可靠性。优化的系统设计可支持所有类型的企业服务和应用。此外，统一管理界面和采用各种标准也有助于降低管理成本，创新性的机箱设计则将为现代数据中心赋予更高的密度、效率和经济性。



图 1. 基于 SPARC M8 处理器的 Oracle 服务器产品系列。



特性对比

表 1 提供了 SPARC T8-1、T8-2、T8-4 和 M8-8 服务器的特性对比。

表 1. 基于 SPARC M8 处理器的服务器特性对比

特性	SPARC T8-1 服务器	SPARC T8-2 服务器	SPARC T8-4 服务器	SPARC M8-8 服务器
外形规格	2U, 深度 737 毫米/29 英寸	3U, 深度 753 毫米/29.6 英寸	6U, 深度 835 毫米/32.9 英寸	系统机架： 宽度 600 毫米 深度 1200 毫米 高度 2 米/78.7 英寸 独立服务器： 10U, 深度 813 毫米/32 英寸
物理域	1		1 或 2 个（静态）	
处理器		32 核 5.0 GHz SPARC M8 处理器，64 MB 3 级共享缓存 每个处理器支持多达 256 个线程 芯片保护的内存 通过 32 个 DAX 引擎来支持内存中查询加速、内联解压缩和 Java 流加速 每个内核中的加密指令加速器直接支持以下 16 种行业标准加密算法，并能生成随机数：AES、Camellia、CRC32c、DES、3DES、DH、DSA、ECC、MD5、RSA、SHA-1、SHA-3、SHA-224、SHA-256、SHA-384、SHA-512		
处理器数量	1	2	2 或 4 个	2-8 个
最大内核数	32	64	128	256
最大线程数	256	512	1024	2048
内存	16 GB、32 GB 或 64 GB DDR4-2400 内存DIMM		32 GB 或 64 GB DDR4-2400 内存 DIMM	
	每个处理器 8 或 16 个 DIMM。DIMM 热备是一项标准特性，可提高系统可靠性，延长系统正常运行时间。 ¹			
内存容量 ¹	最大1024 GB 最小128 GB	最大2048 GB 最小256 GB	最大4096 GB 最小256 GB	最大8192 GB 最小512 GB
内置 2.5 英寸磁盘架	8	6	8	不适用
对内置 2.5 英寸磁盘架的 SAS 支持	1 个支持 RAID 0/1/10/1E 的集成式 SAS3 HBA，最多支持 8 个 2.5 英寸 SAS 硬盘驱动器 (HDD) 或固态驱动器 (SSD)	2 个支持 RAID 0/1/10/1E 的集成式 SAS3 HBA，最多支持 6 (2 + 4) 个 SAS HDD 或 SSD	2 个支持 RAID 0/1/10/1E 的集成式 SAS3 HBA，最多支持 8 (4 + 4) 个 SAS HDD 或 SSD	不适用
对内置 2.5 英寸磁盘架的 NVMe 支持	1 个厂配的可选 PCIe 交换卡，最多支持 4 个 2.5 英寸 NVMe SSD	1 个或 2 个厂配的可选 PCIe 交换卡，最多支持 4 个 2.5 英寸 NVMe SSD	2 个可选 PCIe 交换卡，最多支持 8 (4 + 4) 个 2.5 英寸 NVMe SSD	不适用

1. 原始内存容量。通过全配内存支持 DIMM 热备，预留 1/16 的内存容量。DIMM 热备可以自动停用一整个 DIMM，而不会中断系统运行，也不会导致内存容量损失或更改错误保护功能。

表 1. 基于 SPARC M8 处理器的服务器特性对比 (续)

特性	SPARC T8-1 服务器	SPARC T8-2 服务器	SPARC T8-4 服务器	SPARC M8-8 服务器
Oracle Flash Accelerator F640 PCIe 卡 (NVMe) 的最大数量	6	6	8	12
可移动介质	通过 USB 连接的外部 DVD			无 DVD (通过 USB 和 rKVMs 访问)
管理端口	1 个以太网 1000Base-T 端口 ² 1 个串行 RJ45 端口			2 个以太网 1000Base-T 端口 ² (主用/备用) 2 个串行 RJ45 端口 (主用/备用)
视频端口	1 个 HD-15 VGA 视频端口		2 个 HD-15 VGA 视频端口	不适用
USB 端口	2 个 USB 2.0 (前置) 端口 和 2 个 USB 3.0 (后置) 端口		4 个 USB 3.0 端口	不适用
以太网	4 个集成式 10GBase-T 端口 ³ 1 个集成式以太网控制器			通过 PCIe 适配器卡
PCIe 3.0 半高插槽	6 个插槽 6 个 x8 插槽, 或 2 个 x16 和 2 个 x8 插槽 4 个 PCIe 根联合体 提供支持	8 个插槽 4 个 x8 和 4 个 x16 插槽 8 个 PCIe 根联合体 提供支持	16 个热插拔插槽 8 个 x8 和 8 个 x16 插槽 12 个 PCIe 根联合体 提供支持	至多 24 个热插拔插槽 每个处理器 3 个 x16 插槽 每个插槽一个 PCIe 根联合体
PCIe 根联合体 总数	5	10	20	至多 32 个
冗余电源	2 个冗余 (1+1) 热交换交流 1200 W 电源	2 个冗余 (1+1) 热交换交流 2000 W 电源	4 个冗余 (N+N) 热交换交流 3000 W 电源	6 个冗余 (N+N) 热交换 AC 3000 W 电源
N+1 个冗余热交换风扇	4 个双风扇模块, 顶置式	6 个风扇, 顶置式	5 个双风扇模块, 后置式	8 个双风扇模块, 前置式
操作系统	Oracle 建议采用 Oracle Solaris 11.3 或更高版本, 以获得增强的性能和功能, 包括软件芯片化技术支持的特性 控制域、根域和 I/O 域: Oracle Solaris 11.3 SRU 24 或更高版本 ⁴ 来宾域中支持下列版本: » Oracle Solaris 11.3 SRU 24 或更高版本 ⁴ » Oracle Solaris 10 1/13 ⁵ 仅通过 Oracle Solaris 8 或 Oracle Solaris 9 认证的应用, 可在 Oracle Solaris 10 来宾域的 Oracle Solaris 8 或 Oracle Solaris 9 标记区域中运行。			

2. 1000Base-T 可自动协商到 10 Mb/秒、100 Mb/秒和 1 Gb/秒, 仅支持全双工。

3. 10GBase-T 可自动协商到 100 Mb/秒、1 Gb/秒和 10 Gb/秒, 仅支持全双工。支持最大 15500 字节的巨型帧。

4. 基于 SPARC M8 处理器的服务器不支持 11.3 SRU 23 之前的 Oracle Solaris 11 版本。

5. 外加所需补丁。



SPARC M8 处理器

凭借其第二代软件芯片化功能和新的内核设计，Oracle SPARC M8 处理器可提供创世界纪录的处理速度，并对恶意软件和软件错误提供创新性防御。

SPARC M8 处理器纳入了可加速特定软件功能或基元的硬件单元。这些片上数据分析加速器 (DAX) 单元可分流数据库查询处理、执行实时数据解压缩以及加速 Java 流。相比于其他处理器，内存中查询加速将性能提高了 7 倍之多。内联解压缩特性让同样的内存空间可以存储多达原先两倍的数据，并且丝毫不影响性能。SPARC M8 处理器的新内核还包含 Oracle Numbers 单元，可加快对 Oracle 数据库中的实数的处理。

SPARC M8 处理器的芯片保护的内存特性（图 2）提供实时数据完整性检查，可防范指针相关的软件错误和恶意软件。该特性以低开销的硬件监视取代了成本高昂的软件方法。利用芯片保护的内存特性，应用能够识别错误或未经授权的内存访问，诊断原因并采取相应的恢复操作。SPARC M8 处理器还在各个处理器内核中直接集成了各种增强加密指令加速器。这些加速器能高速运行 16 种行业标准加密和哈希算法，从而消除了安全计算中常见的性能和成本障碍。

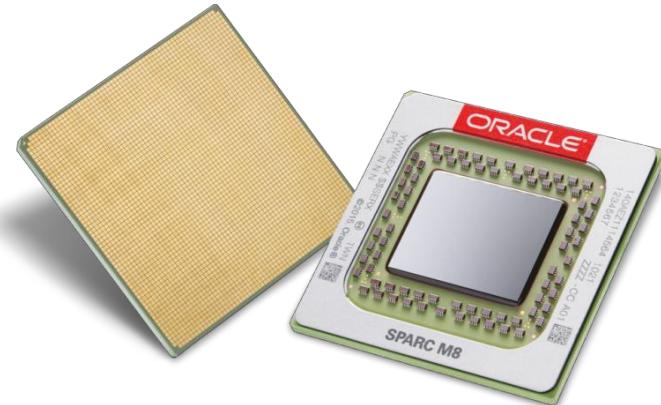


图 2.SPARC M8 处理器在 32 个内核中结合了第二代软件芯片化特性，提高了应用和数据库性能。

由于采用了新的内核和增强的片上二级和三级缓存设计，还提高了处理器频率，因此每线程性能得以提高。64 MB 三级缓存实现了完全共享，热缓存行会移至最近的分区，从而显著降低延迟并提高性能。内核和缓存的架构非常适用于服务器虚拟化和可插拔数据库。因为这种设计大大减少了逻辑域之间或数据库之间的交互，使系统管理和性能调优变得更容易。处理器可以运行多达 256 个线程，因此能够通过降低每线程的性能来换取更高的吞吐量；处理器也可以对每个线程投入更多的资源，减少运行线程数量，提高运行线程性能。这种灵活性使得系统可在总吞吐量与每个线程的性能之间取得平衡，实现最佳效果。

32 核 SPARC M8 处理器与早期的 SPARC 处理器保持二进制兼容。它非常适用于虚拟化云计算环境，可支持大量的虚拟机并提供卓越的多线程性能。借助该处理器，组织能以极高的效率和可预测性快速扩展新网络服务的供应能力。

表 2 提供了 Oracle SPARC M8、SPARC M7、SPARC M6 和 SPARC T5 处理器之间的对比。

表 2. SPARC M8、SPARC M7、SPARC M6 和 SPARC T5 处理器特性对比。

特性	SPARC M8 处理器	SPARC M7 处理器	SPARC M6 处理器	SPARC T5 处理器
CPU 频率	5.0 GHz	4.13 GHz	3.6 GHz	3.6 GHz
乱序执行	支持	支持	支持	支持
指令发送宽度	4	2	2	2
数据/指令预取	支持	支持	支持	支持
SPARC 内核	第五代	第四代	第三代	第三代
每处理器内核数	32	32	12	16
每内核线程数	8	8	8	8
每处理器线程数	256	256	96	128
系统中的插槽数	至多 8 个	至多 16 个	至多 32 个	至多 8 个
每处理器内存	至多 16 个 DDR4 DIMM	至多 16 个 DDR4 DIMM	至多 32 个 DDR3 DIMM	至多 16 个 DDR3 DIMM
缓存	32 KB 一级四路指令缓存 16 KB 一级四路数据缓存 共享 256 KB 二级四路 指令缓存（每四核） 128 KB 二级八路数据 缓存（每核） 共享 64 MB（三级）缓存	16 KB 一级四路指令缓存 16 KB 一级四路数据缓存 共享 256 KB 二级四路 指令缓存（每四核） 共享 256 KB 二级八路 数据缓存（每双核） 共享 64 MB（三级）缓存	16 KB 一级四路指令缓存 16 KB 一级四路数据缓存 128 KB 二级八路缓存 共享 48 MB 三级十二路 缓存	16 KB 一级四路指令缓存 16 KB 一级四路数据缓存 128 KB 二级八路缓存 共享 8 MB 三级十六路 缓存
大分页支持 ¹	16 GB	16 GB	2 GB	2 GB
电源管理粒度	二分之一芯片	四分之一芯片	整个芯片	整个芯片
技术	20 纳米技术	20 纳米技术	28 纳米技术	28 纳米技术

1. Oracle Solaris 11.3 提供大分页支持

SPARC M8 处理器架构

为了以适当的吞吐水平支持商用负载，SPARC M8 处理器采用了新的内核和缓存层次结构，并做了其他方面的改进，可提供高达同类产品两倍的处理速度。电源管理仍然是提高系统中性能的关键，因此处理器中提供了动态电压频率调节 (DVFS)。



图 3 展示了 SPARC M8 处理器的架构。该处理器包含 32 个内核，采用完全共享的三级缓存。它提供了 4 个内存控制器单元 (MCU)，每个 MCU 均通过高速链路连接到板载缓冲器 (BoB) ASIC。BoB 包含 2 个 DDR4 通道，每个通道均连接到 1 个内存 DIMM。这样，每个 SPARC M8 处理器总共可支持高达 16 个 DDR4 DIMM。

SPARC M8 处理器能够利用内含目录进行八路无缝伸缩。它提供了 8 个互联链路 (CL)，用于与其他 SPARC M8 处理器互联互通。2 个 I/O 链路 (IL) 连接到基于 SPARC M8 处理器的服务器中的 I/O 控制器 ASIC。因此，无需其他逻辑链路即可在单个对称式多处理 (SMP) 系统中连接多达 8 个 SPARC M8 处理器。

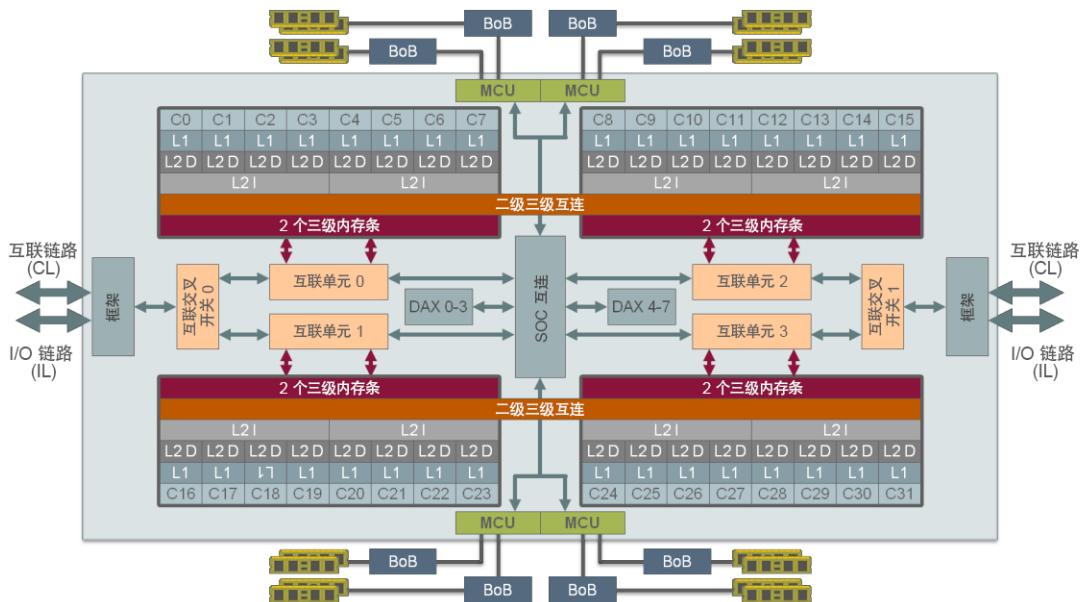


图 3. SPARC M8 处理器配 32 个内核（分为两个分区）、4 个内存控制器单元 (MCU) 和 8 个数据分析加速器 (DAX) 单元。

处理器内核和缓存架构

SPARC M8 内核采用四宽发 (four-wide issue)、乱序设计，支持多达 192 条动态指令以及 8 个硬件线程。这种内核采用动态线程技术进行了优化，可提供超强的每线程性能。软件可通过关键线程优化在每个内核上激活多达 8 个硬件线程。这样处理器硬件可以在活动的硬件线程间动态、无缝地分配内核资源。

处理器内核和最末级缓存被组织为两个分区，每个分区包含 16 个内核和 32 MB 三级缓存。每个内核都有自己的 32 KB 一级指令缓存、16 KB 一级数据缓存和 128 KB 二级数据缓存。4 个内核共享一个 256 KB 二级指令缓存。每个处理器总共支持 64 MB 三级缓存，这些三级缓存完全共享、采用 16 路集关联且包括所有内部缓存。任何三级分区均可处理 SPARC M8 处理器 32 个内核中的任何一个发出的请求。热缓存行移至最近的三级缓存分区，以此提高性能。

内核中的内存管理单元 (MMU) 提供了一个硬件表通道 (HWTW)，支持 8 KB、64 KB、4 MB、256 MB 和 16 GB 页面。

软件芯片化技术

大多数处理器芯片开发都侧重于改进和加快通用处理。几年前，Oracle 启动了一个创新项目，旨在将内存中数据库功能直接移到芯片上，同时对内存中的数据提供硬连接保护。通过在处理器、系统和应用层面不断创新，这种方法让 Oracle 在优化应用性能方面有着得天独厚的优势。Oracle 在 SPARC M7 处理器中引入了软件芯片化功能，首次运用了这一能力。在此基础上，SPARC M8 处理器采用了第二代软件芯片化技术。

SPARC M8 处理器集成了片上加速器，可分流内存中数据库查询处理负载，执行实时数据解压缩以及加快 Java 流。每个处理器内核中直接集成了各种增强加密指令加速器和新的 Oracle Numbers 单元。这些软件芯片化特性共同实现了许多重要安全性、性能和效率优势：

- » **内存中查询加速：**该特性由 DAX 单元提供，可将性能提高到其他处理器的 7 倍之多。
- » **内联解压缩：**支持在同样的内存空间中存储多达之前两倍的数据，并且丝毫不影响性能。
- » **Java 流加速：**允许 Java 8 应用使用新的 Stream 库来利用 DAX 单元，从而改进流分析并显著提升 Java 流操作的性能。
- » **Oracle Numbers 单元：**利用独有的片上加速器来提高对 Oracle Number 数据类型执行算术运算的性能，后者是 Oracle 数据库独有的基本数据类型。
- » **芯片保护的内存：**提供实时数据完整性检查来防范指针相关的软件错误和恶意软件，以低开销的硬件监视取代了成本高昂的软件方法。利用芯片保护的内存特性，应用能够识别错误或未经授权的内存访问，诊断原因并采取相应的恢复操作。
- » **加密加速：**有助于消除安全计算中常见的性能和成本障碍 — 这对于现代业务运营越来越重要。

除了在每个内核中提供各种加密指令加速器和 Oracle Numbers 单元之外，SPARC M8 处理器还包含 8 个第二代 DAX 单元，每个单元有 4 个管道（引擎）。这些引擎可以处理 32 个独立的数据流，分流处理器内核以执行其他任务。DAX 引擎可以处理查询功能，例如解压缩、扫描、筛选和连接。

DAX 单元使用极低开销的进程间通信和极快的原子操作。例如，位于不同处理器上的 DAX 单元可以交换消息并访问远程内存位置，无需使用 CPU 即可交换锁。要利用该功能，需要随带 Oracle Database In-Memory 选件的 Oracle Database 12c 以及 Oracle Solaris 11.3 或更高版本。下面几节将介绍片上加速器实现的软件芯片化特性。

通过关联正确的 Oracle Solaris 库并在测试环境中进行验证，用户无需对现有应用进行重新编译即可启用芯片保护的内存。软件开发人员可通过开放式 Oracle Solaris API 来利用芯片保护的内存和 DAX 特性技术。

内存中查询加速

内存中查询加速特性旨在与 Oracle Database In-Memory 配合使用，后者的主要设计宗旨是提供快速的分析响应。在数据库中存储和访问数据的传统方法是采用行格式。这种方法既适用于会经常插入和更新的事务负载，也适用于报告式查询。然而，数据分析在列格式下运行速度最快。得益于 Oracle Database In-Memory，该特性可以采用双格式架构，能同时提供行格式（适用于联机事务处理 OLTP 操作）和列格式（适用于分析操作）。¹

Oracle Database In-Memory 将数据填充到内存中列存储，自动对要存储在内存中列中的数据运行一系列压缩算法，从而节省存储空间。此外，当运行查询时，它扫描和筛选压缩格式的数据，无需对数据解压缩。内存中列存储创建了许多内存中压缩单元 (IMCU)，如图 4 所示。内存中列数据被分散到这些小型的 IMCU 中，以便对整个数据运行查询时可以进行并行化处理。

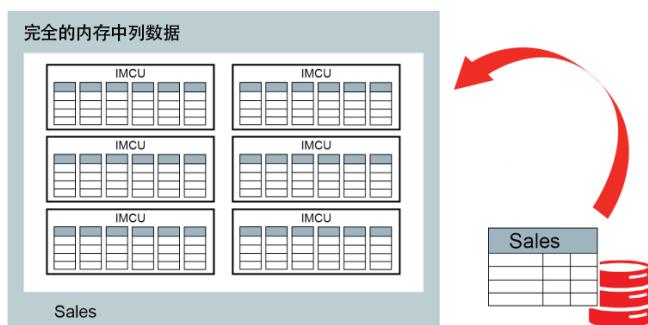


图 4. 内存中列数据分散到这些小型的 IMCU 中，以便可以进行并行化处理。

当 SPARC M8 处理器中的一个内核收到数据库查询时，可以将该查询分流到片上加速器。得到加速的数据库操作包括：

- » 选择：筛选，以减少列
- » 扫描：搜索（“where”子句）
- » 提取：解压缩
- » 转换：进行查找，以加速大到小连接

将查询作业进行分流后，可以将内核释放出来用于执行其他作业，如更高级的 SQL 函数。在此期间，加速器运行查询并将查询结果放在三级缓存中供内核快速访问。当相关内核获悉查询完成后，它便拾取查询结果。

除了能加速操作之外，这种查询分流机制的另一个优势是可以通过每个 SPARC M8 处理器内的 32 个加速器实现大规模并行处理。处理器 32 个内核均可以访问所有这些加速器引擎，并可以同时使用它们，以完全并行的方式运行单个查询。

这种并行机制是通过处理器实现的，不需要应用代码或数据库执行任何额外的操作。加速器可通过处理器的极高带宽接口直接从内存子系统获取数据流。因此，对内存中的数据执行查询时，查询速度由内存接口决定，不受连接至处理器内核的缓存架构的控制。

1 基于 SPARC M8 处理器的服务器支持内存中查询加速，但必需满足下列条件：Oracle Solaris 11.3 或更高版本、Oracle Database 12c 12.1.0.2 补丁包以及 Oracle Database In-Memory 选件。

内联解压缩

毫无疑问，压缩是将更多数据放入内存和存储的关键因素。对于通常读取操作多于写入操作的数据库应用来说，解压缩的速度至关重要。遗憾的是，尽管如今处理器的解压缩性能对于磁盘访问来说足够了，但是对于闪存来说仍然稍慢，因此内存中数据库应用还存在一个巨大的瓶颈。

为了解决这一难题，DAX 单元中实现了一个称为内联数据解压缩的特性，作为查询过程不可或缺的一个步骤。这样，加速器在一个步骤中完成解压缩数据和运行查询函数两项作业，不必多次读取和写入操作，实现以内存速度运行（大于 120 GB/秒）而不影响性能的内联解压缩。其解压缩步骤为：

- » 处理器内核将查询工作分流到加速器，由加速器读取完全压缩的数据（使用 OZIP 压缩）。
- » 加速器动态解压缩数据，并在同一步骤中计算查询，无需任何额外的读取或写入操作。
- » 然后处理器内核将最终查询结果以未压缩的数据格式写出。

Java 流加速

Java 编程语言一直在面向对象的编程领域以及分析应用和框架执行领域开拓进取。如今，许多 Java 应用都在使用 Collections 这个 Java 类，利用预定义的数据结构和方法以程序员友好的方式高效分组和处理大量数据。但是，Collections 类中的许多方法是计算密集型的，在分析大量内存中数据时尤其如此，因为集合中的每一个元素在添加到集合中之前都需要进行计算。

作为 Collections 的替代方法，Java 8 引入了 Java 流的概念。从概念上说，Java 流是一种固定的数据结构，它根据需要来计算元素，并不修改源数据。使用 Java 流允许应用以声明方式（类似于 SQL 语句）处理数据。

使用 Java 流具有许多优势。首先这可以简化编码和操作执行，因为流只在必要时才会对数据执行计算，从而避免不必要的内存复制。此外，得益于其底层设计，Java 流能够利用多线程和多核架构。

流的特点包括：

- » 元素顺序：流会按顺序提供一组特定类型的元素。流会根据需要获取/计算元素。它从不存储元素。
- » 数据源：流使用集合、数组或 I/O 资源作为其输入源。
- » 聚合操作：流支持聚合操作，例如筛选、映射、限制、缩减、查找、匹配等。
- » 管道化：大多数流操作都会返回一个流，这可以实现管道化。这些操作称为中间操作，其功能是获取输入、加以处理并将输出返回给下一个流操作。
- » 自动迭代：与需要显式迭代的集合相比，流操作在内部对所提供的源元素进行迭代。

SPARC 处理器的 DAX 单元旨在以极高速度执行专用功能，包括扫描、选择、提取、填充和转换。它们基于 Java 流的实现和功能，非常适合利用 DAX 单元的独有特性。

Oracle 发布了一个新的 Stream API 来充分利用 DAX 单元。该库随 Oracle Solaris 软件包提供，让 Java 程序员只需对源文件中的 `import` 语句进行细微更改就能够在利用与标准 Stream API 相同的接口的同时使用 SPARC M8 处理器的 DAX 单元。此外，该库只有在能产生效益的情况下才会将整数流函数分流至 DAX 单元，在其他情况下则执行常规流操作。

DAX 单元在 Java 流处理方面具备显著的优势：某些操作的性能提升高达 20 倍 — 具体提升倍数取决于数据源的大小和对数据所执行的操作类型。

Oracle Numbers 加速

Oracle 数据库具有独特的基本数据类型。SQL 语句中的每个列值和常量都有一个数据类型，该数据类型与特定的存储格式、约束和有效值范围有关。应始终为 Oracle 数据库中的表的每一列指定一个数据类型。

具体到固定数和浮点数，Oracle 数据库有一种称作 Oracle Numbers 的独有数据类型，用于存储固定数和浮点数。该数据类型可以存储几乎任何数量级的数字，并且可以在运行 Oracle 数据库的不同系统之间移植，精度高达 38 位。

SPARC M8 处理器中的 32 个内核均包含一个 Oracle Numbers 单元，专门用于加速 Oracle Numbers 算术性能。Oracle Numbers 数据类型提供了四个新指令（ONadd、ONsub、ONmul 和 ONdiv），原生支持 22 个字节以内所有长度的固定数和浮点数。

处理 Oracle Numbers 数据的算术操作是一项计算密集型任务，因为每个操作都需要多个指令，而每个任务的数据库算术过程可能都需要数百万次算术运算。SPARC 处理器的新 Oracle Numbers 单元可加快处理速度，每个操作只需一条指令。这大幅提升了性能：大型 Oracle Numbers 位长度（超过 16 个字节）的性能可提高 10 倍，从而减少处理操作所需的时间并释放内核用于执行其他任务。

对于算术密集型负载（例如报告和数据仓库）来说，该特性可显著节省资源并减少部署所需的计算节点总量。

加密加速

如今，增强安全性比以往更加重要，而 SPARC 处理器和系统在提供基于处理器的加密加速方面历史悠久。SPARC M8 处理器中的 32 个内核均包含一个加密指令加速器，支持最全面的加密密码、散列、密钥操作和校验和。这些内核支持 16 种行业标准加密算法，并能生成随机数。加速加密则通过 Oracle Solaris 的加密框架支持。

SPARC M8 处理器允许访问有加密密码的硬件实现，支持的算法包括：AES、Camellia、CRC32c、DES、3DES、DH、DSA、ECC、MD5、RSA、SHA-1、SHA-3、SHA-224、SHA-256、SHA-384 和 SHA-512。它在相应的管线内部实现加密，而不是在处理器内部加密。这种基于硬件的加密实现方法更高效，而且无需更改权限，从而大幅提高加密算法计算的效率。

此外，数据库操作可以更加高效地使用在指令流水线本身内实现的各种加密算法。通过在 Oracle 体系的每一层中应用 SPARC M8 处理器的内置加密功能，数据安全性提高了，而且几乎不会对性能产生任何影响。

芯片保护的内存

SPARC M7 和 SPARC M8 处理器中的芯片保护的内存特性通过在硬件中置入动态指针检查来提供基于硬件的内存保护。芯片保护的内存可检测并报告内存引用错误，阻止对内存中数据的无意或恶意的访问。

C 和 C++ 等编程语言仍然易于发生由软件错误引起的内存损坏。这些类型的内存引用错误极难发现，受影响者通常在损坏发生很久后才会发现损坏的数据。而数据库和应用可能包含数千万行代码，可能由数千名开发人员开发完成，这让问题变得更复杂。重要的是，缓冲区溢出等错误是让组织面临风险的安全漏洞的主要根源。

现代应用使用许多占用大量共享内存段的线程。这些应用中的错误或指针问题可能会导致高度不可预测的行为，需要应用开发人员花费大量时间来进行问题排除和诊断。静默数据损坏和缓冲区溢出就是其中两个难以诊断的问题。对于这两个问题，芯片保护的内存特性可以显著缩短应用开发人员排除内存引用错误所花费的时间。就静默数据损坏而言，芯片保护的内存特性可支持应用立即采取应对措施，从而无需执行成本高昂的恢复工作。

图 5 展示了静默数据损坏问题，其中两个应用线程（A 和 B）意外访问同一个内存位置。图中用颜色区分每个线程各自本应访问的内存区域。然而，由于软件编程错误，可能出现这样的情况：线程 A 错误地写入线程 B 的红色轮廓区域。这种错误往往不会立即被发现，只有当线程 B 读取内存时才有可能发现。此时线程 B 包含被线程 A 静默损坏的数据，而损坏根源往往极难追踪。这个问题很难诊断，常常表现为会带来严重后果的软件错误。

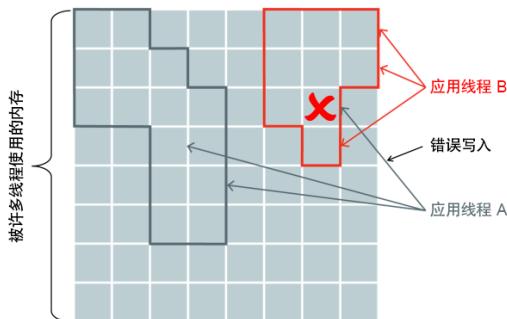


图 5. 当两个线程错误地写入同一个内存位置时，发生了静默数据损坏问题。



缓冲区溢出是应用中可能会发生的另一个问题。简单地说，缓冲区溢出意味着应用错误地开始在其分配区域以外写入数据（图 6）。发生这个错误后，敏感数据可能会泄漏到其他内存位置，而应用却不知道这一情况。于是，具有恶意企图的应用可能会读取所有这些敏感信息。缓冲区溢出可能会带来灾难性的安全噩梦，在当今的环境下通常以恶意病毒攻击的形式出现。

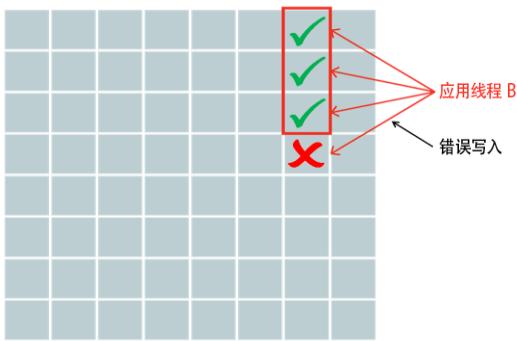


图 6. 缓冲区溢出可能会带来巨大的安全风险。

芯片保护的内存特性将各（内存）指针的键作为内存版本，解决了这些问题。在内存分配过程中，该特性会将一个相应的代码作为内存版本写入到内存中。当任何指针访问该内存时，硬件会对键和代码进行比较。如果两者相匹配，则访问是合法的。如果两者不匹配，则存在内存引用错误，系统会立即发现该错误。

基于 SPARC M8 处理器的服务器系列概述

基于 SPARC M8 处理器的服务器专为需要高水平的安全性、性能和效率的云基础设施而设计。这些 SPARC 服务器适用于数据库、Java、中间件和企业应用，可提供卓越的吞吐性能和内存带宽。该系列服务器产品支持从 1 个 SPARC M8 处理器扩展至 8 个，因此支持非常广泛的应用、功能和容量。

这些服务器所共有的硬件特性包括：

- » SPARC M8 32 核 5.0 GHz 处理器（采用第二代软件芯片化技术）
- » 32 GB 和 64 GB DDR4-2400 内存 DIMM（SPARC T8-1 和 T8-2 服务器也可使用 16 GB DIMM）
- » PCIe 3.0 x16 扩展插槽
- » 支持 NVMe (NVMe) 闪存设备
- » SPARC T8-1、T8-2 和 T8-4 服务器中提供板载 12 Gb/秒 SAS3 I/O 控制器
- » 嵌入式 USB (eUSB) 存储设备，支持通过 InfiniBand 网络启动

内存子系统

每个 SPARC M8 处理器通过 8 个板载缓冲器 (BoB) ASIC 支持多达 16 个 DDR4 内存 DIMM。每个处理器配 有 16 个 64 GB DIMM 时支持高达 1 TB 内存，每个 DIMM 均由专用内存通道提供支持。SPARC M8 处理器的内存链路原始带宽为 374 GB/秒，足以满足 16 个 DDR4-2400 通道对 307 GB/秒的原始聚合带宽的需求。此外，支持半配和满配内存配置。详细的配置策略将在本文稍后介绍各款服务器的章节中讨论。

由内存 DIMM 提供、但由单个 SPARC M8 处理器控制的物理地址空间采用交错技术来尽可能提高性能。半配内存配置采用八路交错，全配内存配置（每个处理器 16 个 DIMM）采用 16 路交错。SPARC M8 处理器还支持 15 路交错配置，可以动态地从 16 路切换至 15 路配置。以此功能为基础，基于 SPARC M7 处理器的服务器中首次引入了 DIMM 热备特性。DIMM 热备可减少 DIMM 更换维护工作，从而延长系统正常运行时间。

DIMM 热备功能能够将一个故障 DIMM 从配置中移除，从而防止该故障 DIMM 引起意外的系统中断。该特性将每个 DIMM 容量的 1/16 保留闲置，之后可以停用故障 DIMM，而将其内容重新映射到其余的 15 个 DIMM。在确定某个 DIMM 出现故障时会自动执行 DIMM 热备，不会中断应用服务。通过这一过程，执行 DIMM 热备后，系统内存容量不会发生变化，错误保护也保持不变。系统仍继续运行，且无任何容量损失，也不会增加故障风险。因此，不需要为了维护硬件而让系统停机。实际的 DIMM 更换过程可以等到同一个内存条中的第二个 DIMM 需要更换时进行。

当采用全配内存配置（每个处理器 16 个 DIMM）时，基于 SPARC M8 处理器的服务器支持 DIMM 热备。采用半配内存配置不支持 DIMM 热备。您可以在全配内存配置下禁用 DIMM 热备，但不建议这样做。

I/O 子系统

基于 SPARC M8 处理器的服务器中的 I/O 子系统采用相同的基本设计。每个 SPARC M8 处理器均通过 I/O 链路 (IL) 连接至 1 个或 2 个 I/O 控制器 ASIC。SPARC M8 处理器和 I/O 控制器 ASIC 中均有 2 个 IL。

SPARC M8 处理器与 I/O 控制器 ASIC 之间的连接方式有两种。在 SPARC T8-1 和 SPARC M8-8 服务器中，每个 SPARC M8 处理器使用全部 2 个 IL 连接到单个 I/O 控制器 ASIC。SPARC T8-2 和 SPARC T8-4 服务器则使用交叉连接模式，这样，当某个处理器或 IL 不可用时也能连接到 2 个 I/O 控制器 ASIC（和 PCIe 设备）。在交叉连接模式中，处理器中的一个 IL 连接到一个 I/O 控制器，另一个 IL 连接到另一个 I/O 控制器。有关更多详细信息，请参见本文介绍各款服务器的章节。

I/O 控制器 ASIC

PCIe 基础设施是通过 I/O 控制器 ASIC 提供的。每个 ASIC 提供 5 个 PCIe 3.0 根联合体和 72 GB/秒聚合带宽。由于 I/O 控制器 ASIC 包含整个 PCIe 结构，因此在添加或删除处理器时该结构保持不变。这样，PCIe 设备路径不会发生变化，因为与设备的连接在 I/O 控制器 ASIC 提供的根联合体中是固定的。基于 SPARC M8 处理器的服务器中所使用的 I/O 控制器 ASIC 具有下列重大创新：

- » 2 个 x16 IL 连接至 SPARC M8 处理器，每个 IL 包含 2 个 x8 连接。
- » 每个 x8 IL 连接支持单通道故障。
- » 每个 IL 均参与硬件缓存一致性。
- » SPARC T8-2 和 T8-4 服务器中使用双主机处理器故障切换。
- » 提供 SR-IOV 合规性。
- » 提供每个 DMA 流的地址转换。
- » 提供每个 DMA 流的宽松数据包排序。
- » 4 个 PCIe 3.0 x16 端口，可分为四路，可以作为 1 个 x16、4 个 x4 或 2 个 x8 PCIe 端口来实现。
- » 1 个 PCIe 3.0 x8 端口，可分为两路，可以作为 1 个 x8 或 2 个 x4 PCIe 端口来实现。
- » 5 个 PCIe 3.0 端口中，每一个都是一个独立的根联合体。

NVM Express 技术

基于 SPARC M8 处理器的服务器支持 NVM Express (NVMe) 这一新兴的闪存存储技术。NVMe 规范为固态驱动器 (SSD) 定义了一个基于 PCIe 的优化的接口。利用非易失性内存，基于 NVMe 的 SSD 比基于 SAS 或 SATA 的 SSD 延迟更低、吞吐性能更高。NVMe 使用 PCIe 信令，对每个驱动器提供一个 8 GT/秒的 x4 接口，这样可在全双工状态下为驱动器提供约 4 GB/秒的带宽。

所有基于 SPARC M8 处理器的服务器均支持 Oracle Flash Accelerator F640 PCIe 卡 — 可在一个半高 PCIe 卡上提供基于 NVMe 的 SSD 设备。SPARC T8-1、T8-2 和 T8-4 服务器支持在精选的磁盘架中内置 2.5 英寸的小尺寸 (SFF) NVMe SSD，这种磁盘架也支持基于 SAS 的 HDD 和 SSD。如果使用 SFF NVMe 驱动器，则需要厂配的 NVMe PCIe 交换卡和缆线。该交换卡使用一个 x8 PCIe 3.0 接口，最多可以为 4 个 x4 下游链路（每个 NVMe 驱动器一个）提供扇出和电子重定时功能。

NVMe 设备支持热插拔，但必须遵循 OS 特定的热插拔过程。管理员可以使用 `nvmeadm` 命令收集驱动器运行状况和固件级别信息，检查温度，获取错误日志，访问 SMART 数据以及执行安全擦除和低级格式化。

嵌入式 USB 存储和 Oracle Solaris 引导池

基于 SPARC M7 和 SPARC M8 处理器的服务器支持一种特定的引导过程，能够从更多类型的设备进行引导。传统的引导过程要求引导设备可由系统固件进行访问。举例来说，出于此原因，以前不支持通过 InfiniBand 进行网络引导。

备选的 Oracle Solaris 引导过程涉及一个称为 **引导池** 的概念。引导池是一个用于存储引导归档文件的引导设备。在基于 SPARC M8 处理器的服务器中，将一个或多个嵌入式 USB (eUSB) 存储设备放在一起，形成一个引导池，使得 OpenBoot PROM 固件可访问引导池。eUSB 存储是一个内置的 USB 闪存设备，出厂前就已安装在系统中。SPARC T8-1、T8-2 和 T8-4 服务器包含一个 eUSB 设备。在 SPARC M8-8 服务器中，每个 CPU、内存、I/O 单元 (CMIOU) 机箱主板都有一个 eUSB 设备，这样引导池可包含多个条带化的 eUSB 存储设备。

本地引导池的存在使得 OpenBoot PROM 固件能够加载引导归档文件，随后使用基于 IP over InfiniBand (IPoIB) 的 iSCSI 在根池中挂载根文件系统。此外，基于 SPARC M8 处理器的服务器还为这个新的引导过程提供了回退机制，当基于 eUSB 的引导归档文件不可用时就可以使用这种机制。系统服务处理器 (SP) 的闪存中有一个引导归档文件。这个引导归档文件在出厂时就已加载到 SP 中，旨在仅当其他引导方式不可用时才使用。

PCIe 适配器卡

基于 SPARC M8 处理器的服务器同时配有 PCIe 3.0 x8 和 x16 扩展卡插槽。支持的选件和要求因服务器型号而异。在本文发布时，Oracle 提供的适配器卡包括：

- » Oracle Flash Accelerator F640 PCIe 卡：6.4 TB，NVMe PCIe 3.0
- » Oracle 存储双端口 16 Gb 或 32 Gb 光纤通道 PCIe HBA
- » Sun 存储 16 Gb FC PCIe 通用 HBA
- » Oracle 四端口 10GBase-T 适配器
- » Sun 四端口 GbE PCIe 2.0 半高适配器，UTP
- » Oracle 四端口 10Gb 或 双端口 40Gb 以太网适配器
- » Sun 双端口 10GbE SFP+ PCIe 2.0 半高适配器
- » Oracle 存储 12 Gb SAS PCIe HBA，外部：8 端口
- » Oracle 双端口 QDR InfiniBand 适配器 M3

SPARC T8-1、T8-2 和 T8-4 服务器

SPARC T8-1、T8-2 和 T8-4 服务器旨在提供突破性的安全性和性能，同时尽可能提高可靠性并尽可能降低能耗和复杂性。这些系统非常适合横向扩展应用，因为这种环境通过对系统及支持系统间故障切换的应用进行复制来提供高可用性。SPARC T8-1、SPARC T8-2 和 SPARC T8-4 服务器拥有通用、共享的组件和子系统，但用不同的设计来优化系统，从而实现其特定的设计点和功能。这些服务器采用强健的机箱、组件、子配件设计，具有增强的系统和组件可维护性，并且尽可能减少了布线，加大了空气流动。

- » SPARC T8-1 服务器采用单处理器，是一款入门级服务器。然而，这款服务器提供 256 个硬件线程和 1 TB 内存，因此性能优于双处理器同类产品。与该服务器系列中的其他款型一样，SPARC T8-1 服务器提供全套的软件芯片化特性，包括利用芯片保护的内存特性来保证安全操作，以及利用内存中数据库查询加速和内联数据解压缩特性来提高性能。
- » SPARC T8-2 服务器提供两倍于 SPARC T8-1 服务器的容量和更多的资源。双处理器 SPARC T8-2 服务器还包含其他可用性特性，例如处理器与 I/O 控制器 ASIC 之间创新的交叉连接、2 个集成式 SAS 控制器，以及支持内置 NVMe SSD 的 2 个选配 NVMe PCIe 交换卡。

» SPARC T8-4 服务器最多支持 4 个处理器，这些处理器安装在前置的处理器模块上。根据需要，客户可以轻松地从双处理器这种入门级配置扩展为四处理器。此外，SPARC T8-4 服务器还提供由大量专用 PCIe 根联合体支持的可以逐个热插拔的 PCIe 卡，因此，这款服务器非常适用于将负载整合到虚拟化私有云中。

SPARC T8-1 服务器

SPARC T8-1 服务器是一款入门级服务器，配有单个 SPARC M8 处理器，采用 2U 机柜设计。其标准配件包括 8 个板载内存 DIMM 插槽，利用选配的双夹层卡（内存扩展卡）可扩展至总共 16 个，从而最多可提供 1 TB 的系统内存。图 7 展示了 SPARC T8-1 服务器的前视图和后视图。这款服务器包括 6 个半高 PCIe 3.0 扩展插槽，可从系统后部触及。

一个板载 12 Gb/秒 SAS HBA（提供 RAID 0、1、10 和 1E 保护）支持全部 8 个 2.5 英寸的小尺寸 (SFF) 前置式磁盘架。Oracle Solaris ZFS 提供更高级别的 RAID 支持。该服务器还提供了一个厂配的选件，可支持多达 4 个 2.5 英寸 SFF NVMe SSD。该选件包括一个 NVMe PCIe 交换卡，它安装在 PCIe 插槽 3 中并通过内部布线连接到磁盘笼。支持混合使用 SAS 和 NVMe 驱动器。SPARC T8-1 服务器的其他标准配件包括：

- » 系统后部提供 4 个 10GBASE-T 以太网端口 (RJ45)，并通过板载网络接口控制器 (NIC) 支持这些端口。端口自动协商至 100 Mb/秒、1 Gb/秒和 10 Gb/秒，仅支持全双工。支持最高 15500 字节的巨型帧。
- » 2 个热交换 1+1 电源单元 (PSU)，200VAC–240VAC 时功率为 1200 W，从系统后部插入。
- » 提供 4 个顶置式热交换风扇模块，每个模块带有 2 个对转风扇。
- » 前置 2 个 USB 2.0 端口，后置 2 个 USB 3.0 端口和 1 个 VGA 视频端口 (HD15)。
- » 提供 1 个外部 DVD，通过 USB 端口连接。

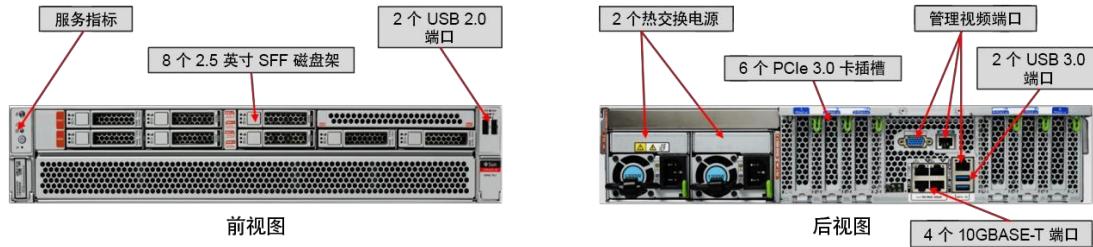


图 7. SPARC T8-1 服务器前视图和后视图。

图 8 是 SPARC T8-1 服务器的原理图。SPARC M8 处理器连接到板载内存插槽以及选配的内存扩展卡，这样最多可支持 16 个 DDR4 DIMM 插槽。SPARC M8 处理器直接连接至 I/O 控制器 ASIC（提供 PCIe 根联合体）。选配的 NVMe PCIe 交换卡支持将最多 4 个 2.5 英寸 NVMe 设备置于居中的 4 个 SFF 磁盘架。



服务处理器 (SP) 上运行的 Oracle Integrated Lights Out Manager (Oracle ILOM) 远程控制台提供远程键盘、视频和鼠标 (rKVM) 功能，提供单独的串行管理端口 (RJ45) 和以太网管理端口 (1000Base-T, RJ45) 来与 SP 接口，还可以使用板载 10GBASE-T 网络端口（启用边带特性）以高达 10 Gb/秒的速度连接至 SP。6 个 PCIe 插槽中的 2 个（插槽 3 和 4）在相邻插槽（分别为插槽 2 和 5）空置时可支持 x16 连接。

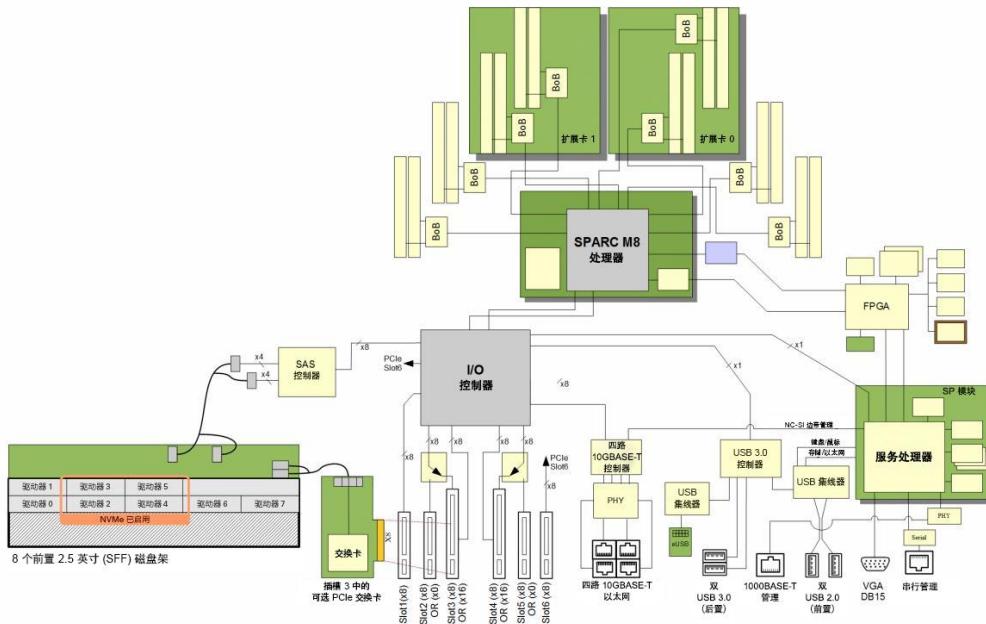


图 8.SPARC T8-1 服务器搭载单个 32 核 SPARC M8 处理器和 I/O 控制器 ASIC。

表 3 描述了 PCIe 设备如何共享 SPARC T8-1 服务器的 I/O 控制器 ASIC 上的 5 个根联合体。

表 3. SPARC T8-1服务器根联合体的对应信息。

根联合体	目标	速度
根联合体 0	PCIe 插槽 6 SAS 控制器	x8 x8
根联合体 1	PCIe 插槽 4 PCIe 插槽 5	x8 或 x16 x8 或 x0
根联合体 2	PCIe 插槽 1 四路 10GbE 控制器	x8 x4
根联合体 3	PCIe 插槽 2 PCIe 插槽 3	x8 或 x0 x8 或 x16
根联合体 4	USB 控制器 服务处理器 (板载显卡)	x4 端口中的 x1 x4 端口中的 x1



SPARC T8-2 服务器

SPARC T8-2 服务器搭载双 SPARC M8 处理器，采用 3U 架装式机柜设计。每个扩展卡配有 2 个或 4 个内存 DIMM，8 个扩展卡上最多装有 2 TB 的系统内存。图 9 展示了 SPARC T8-2 服务器的前视图和后视图。这款服务器包含 8 个半高 PCIe 3.0 扩展插槽，可从系统后部触及。

2 个板载 12 Gb/秒 SAS HBA（提供 RAID 0 和 1 保护）支持 6 个 2.5 英寸 SFF 前置式磁盘架（通过 2 个 HBA 拆分成 2 组，其中 2 个驱动器为一组，另外 4 个驱动器为一组）。带有 4 个磁盘架的 SAS HBA 还支持 RAID 10 和 1E。Oracle Solaris ZFS 提供更高级别的 RAID 支持。该服务器还提供了厂配的选件，可支持多达 4 个 2.5 英寸 SFF NVMe 驱动器。这些选件包括 1 个 NVMe PCIe 交换卡，它通过内部布线连接至磁盘笼。当配有下列选件时，支持混合使用 SAS 和 NVMe 驱动器：

- » **单 NVMe PCIe 交换卡。** 该交换卡安装在 PCIe 插槽 1 中，支持全部 4 个具备 NVMe 功能的磁盘架（上面 4 个磁盘架）。
- » **双 NVMe PCIe 交换卡。** 这两个交换卡安装在 PCIe 插槽 1 和 2（具有各自的根联合体）中，每个交换机支持 2 个 NVMe 磁盘架。

SPARC T8-2 服务器的其他标准配件包括：

- » 系统后部提供 4 个 10GBASE-T 以太网端口 (RJ45)，并通过板载 NIC 支持这些端口。端口自动协商至 100 Mb/秒、1 Gb/秒和 10 Gb/秒，仅支持全双工。支持最高 15500 字节的巨型帧。
- » 2 个热交换 1+1 电源单元 (PSU)，200 VAC–240 VAC 时功率为 2000 W，从系统后部插入。
- » 6 个热交换风扇模块，从机箱顶部装载。
- » 前置 2 个 USB 2.0 端口，后置 2 个 USB 3.0 端口。
- » 机箱后部提供 1 个 VGA 视频端口 (HD15)。
- » 提供 1 个外部 DVD，通过 USB 端口连接。

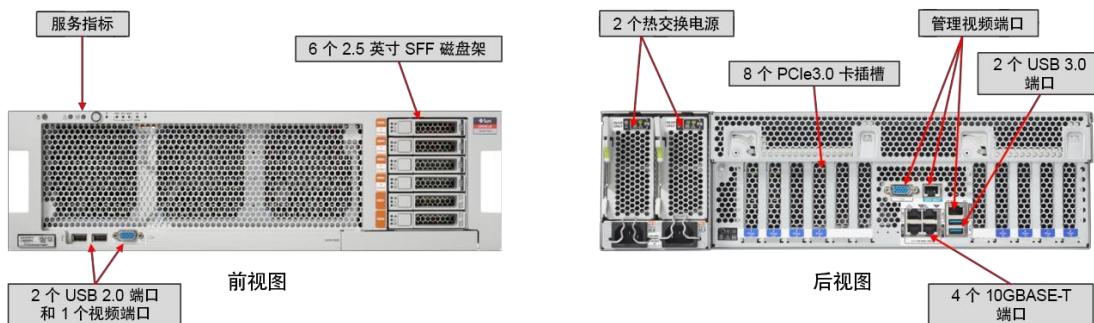


图 9. SPARC T8-2 服务器前视图和后视图。

图 10 是 SPARC T8-2 服务器的原理图。SPARC M8 处理器的 2 个插槽连接到内存扩展卡，这样最多可支持 32 个 DDR4 DIMM 插槽（每个处理器插槽支持 16 个）。处理器通过冗余互联链路相连接。2 个 I/O 控制器连接至 2 个 SPARC M8 处理器。这种交叉连接设计的好处在于，即使一个处理器出现故障，系统也仍然能够访问所有 I/O 设备。如果系统仅通过一个处理器引导，则 PCIe 设备路径和全部 10 个 PCIe 根联合体将保持不变。6 个 PCIe 插槽中的 4 个支持 x16，4 个连接起来以支持 x8。对于高吞吐量设备至关重要的是，支持 x16 的 PCIe 插槽具有专用的根联合体，没有其他设备共享这些根联合体。



服务处理器 (SP) 上运行的 Oracle ILOM 远程控制台提供远程键盘、视频和鼠标 (rKVM) 功能，提供单独的串行管理端口 (RJ45) 和以太网管理端口 (1000Base-T, RJ45) 来与 SP 接口，还可以使用板载 10GBASE-T 网络端口（启用边带特性）以高达 10 Gb/秒的速度连接至 SP。

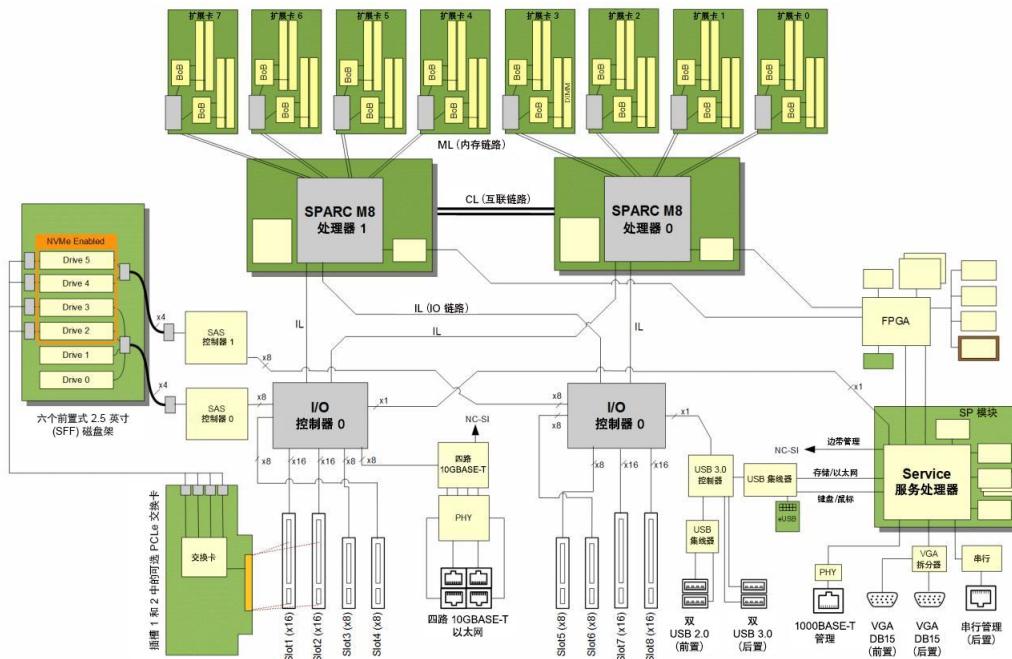


图 10.SPARC T8-2 服务器搭载双 SPARC M8 处理器，并通过交叉连接 2 个 I/O 控制器 ASIC 来提高可用性。

表 4 描述了 PCIe 设备如何共享每个 I/O 控制器上的根联合体。

表 4. SPARC T8-2 服务器 I/O 控制器和根联合体的对应信息。

I/O 控制器	根联合体	目标	PCIe 速度
I/O 控制器 0	根联合体 0	PCIe 插槽 4 SAS 控制器 0	x8 x8
	根联合体 1	PCIe 插槽 3 四路 10GbE 控制器	x8 x8
	根联合体 2	PCIe 插槽 2	x16
	根联合体 3	PCIe 插槽 1	x16
	根联合体 4	服务处理器 (板载显卡)	x4 端口中的 x1
I/O 控制器 1	根联合体 0	PCIe 插槽 6 SAS 控制器 1	x8 x8
	根联合体 1	PCIe 插槽 8	x16
	根联合体 2	PCIe 插槽 7	x16
	根联合体 3	PCIe 插槽 5	x8
	根联合体 4	USB 控制器	x4 端口中的 x1



SPARC T8-4 服务器

SPARC T8-4 服务器搭载 4 个 SPARC M8 处理器，内存高达 4 TB，采用 6U 架装式机柜设计。图 11 展示了 SPARC T8-4 服务器的前视图和后视图。这款服务器包含 8 个 x8 和 8 个 x16 PCIe 3.0 扩展插槽（装备有热插拔 PCIe 支架）— 可从系统后部触及。

8 个 2.5 英寸 SFF 前置式磁盘架均匀分布在 2 个板载 12 Gb/秒 SAS HBA（提供 RAID 0、1、10 和 1E）上。Oracle Solaris ZFS 提供更高级别的 RAID 支持。除了 SAS HDD 和 SAS SSD 之外，该服务器还提供了厂配的选件，以支持多达 8 个 2.5 英寸 SFF NVMe SSD 驱动器。这些选件最多包含 2 个 NVMe PCIe 交换卡，交换卡通过内部布线连接至磁盘笼。每个 NVMe PCIe 交换卡允许使用 4 个磁盘架来支持 NVMe 驱动器。支持混合使用 SAS 和 NVMe 驱动器。

SPARC T8-4 服务器的其他标准配件包括：

- » 系统后部提供 4 个 10GBASE-T 以太网端口 (RJ45)，并通过板载 NIC 支持这些端口。端口自动协商至 100 Mb/秒、1 Gb/秒和 10 Gb/秒，仅支持全双工。支持最高 15500 字节的巨型帧。
- » 通过 4 个 I/O 控制器 ASIC 支持 16 个 PCIe 半高热插拔支架 I/O 插槽，包括
 - » 8 个 PCIe 3.0 x8 插槽
 - » 8 个 PCIe 3.0 x16 插槽
- » 4 个热交换 N+N 冗余 PSU，200VAC–240VAC 时功率为 3000 W，从系统前面插入。
- » 5 个热交换风扇模块，从机箱顶部装载。
- » 提供本地（前置和后置 USB 和视频）和远程键盘、视频和鼠标 (KVM) 功能。
- » 提供单独的串行和网络管理端口，用于与板载服务处理器接口。
- » 提供一个外部 DVD，通过 USB 端口连接。

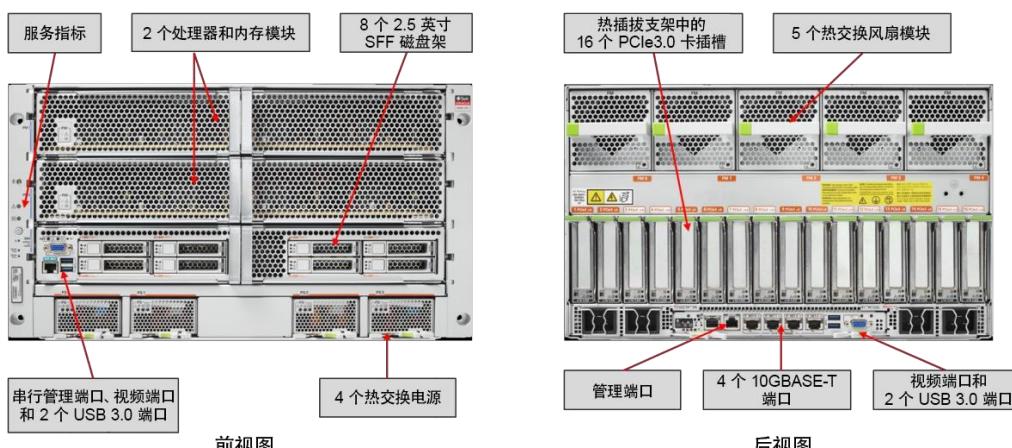


图 11.SPARC T8-4 服务器前视图和后视图。



图 12 是 SPARC T8-4 服务器的原理图。该服务器最多有 2 个专用处理器/内存模块，每个模块提供 2 个 SPARC M8 处理器插槽和 32 个 DDR4 DIMM 插槽（每个处理器插槽 16 个 DIMM 插槽，每台服务器 64 个 DIMM 插槽）。一块中板以冗余的方式将每个处理器插槽的互联链路连接到另一个处理器/内存模块上的一个处理器插槽，冗余 IL 将每个处理器插槽连接至 4 个 I/O 控制器中的 2 个，由 4 个 I/O 控制器为系统中的所有 PCIe 插槽提供 PCIe 根联合体，从而提供大量 I/O 功能。SAS HDD、SAS SSD 或 NVMe SSD 可安装到任意磁盘架中。选配的 NVMe PCIe 交换卡不占用机箱后部的任何热插拔 PCIe 支架插槽，这样为额外的 I/O 设备留下了更多的容量。

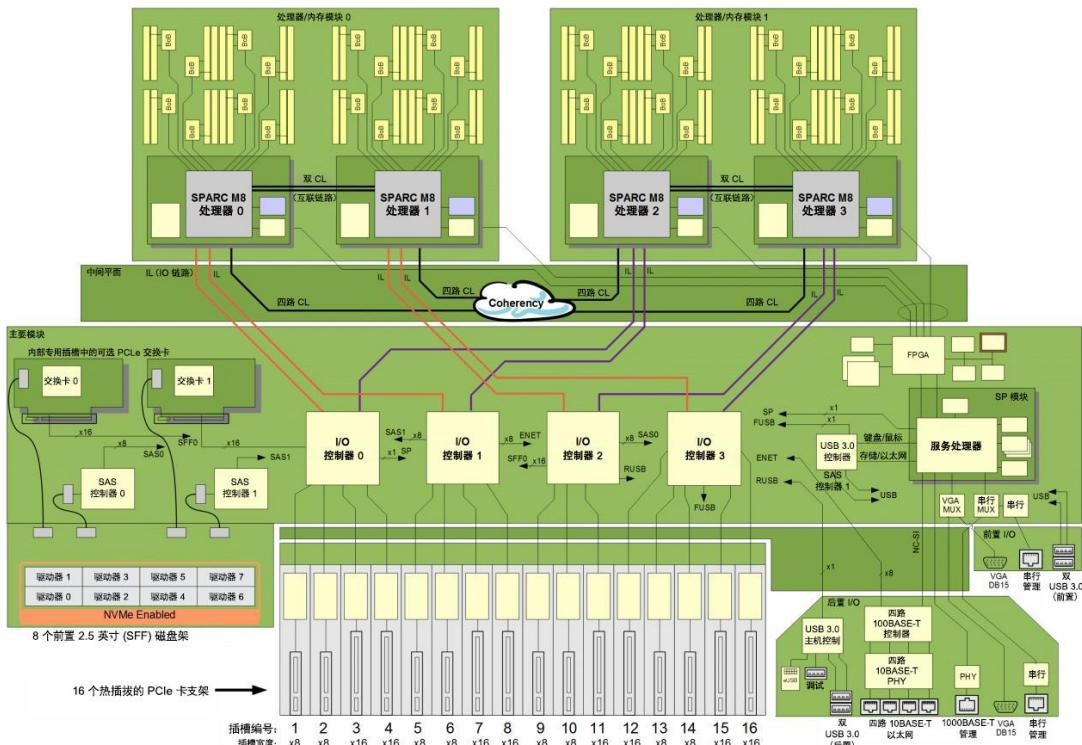


图 12. SPARC T8-4 服务器配有 2 个双处理器/内存模块和 4 个 I/O 控制器 ASIC。



表 5 描述了 PCIe 设备如何共享 4 个 I/O 控制器上的根联合体。

表 5. SPARC T8-4 服务器 I/O 控制器和根联合体的对应信息。

I/O 控制器	根联合体	目标	PCIe 速度
I/O 控制器 0	根联合体 0	PCIe 插槽 1 PCIe 插槽 2	x8 x8
	根联合体 1	NVMe PCIe 交换卡 1	x16
	根联合体 2	PCIe 插槽 4	x16
	根联合体 3	PCIe 插槽 3	x16
	根联合体 4	服务处理器（板载显卡）	x4 端口中的 x1
I/O 控制器 1	根联合体 0	PCIe 插槽 5 PCIe 插槽 6	x8 x8
	根联合体 1	SAS 控制器 1	x16
	根联合体 2	PCIe 插槽 8	x16
	根联合体 3	PCIe 插槽 7	x16
	根联合体 4	四路 10 GbE 控制器	x8
I/O 控制器 2	根联合体 0	PCIe 插槽 9 PCIe 插槽 10	x8 x8
	根联合体 1	SAS 控制器 0	x8
	根联合体 2	NVMe PCIe 交换卡 0	x16
	根联合体 3	PCIe 插槽 11	x16
	根联合体 4	后置 USB 3.0 控制器	x4 端口中的 x1
I/O 控制器 3	根联合体 0	PCIe 插槽 13 PCIe 插槽 14	x8 x8
	根联合体 1	PCIe 插槽 12	x16
	根联合体 2	PCIe 插槽 16	x16
	根联合体 3	PCIe 插槽 15	x16
	根联合体 4	前置 USB 3.0 控制器	x4 端口中的 x1



SPARC M8-8 服务器

SPARC M8-8 服务器专为现代云基础设施而设计。对于需要大规模任务关键型计算环境具有的运营效率、可靠性和可伸缩性的数据库和商业应用来说，这款服务器是理想之选。SPARC M8-8 服务器具有很大的内存空间和出色的内存带宽，因而非常适用于将数据库和企业应用完全部署在内存中来提供极致的性能和容量。

依托于创新式的 CPU、内存和 I/O 单元 (CMIOU) 机箱和系统互连，SPARC M8-8 服务器可提供更高的可伸缩性、更多的数据中心特性和更加丰富的功能，因此非常适用于要求严苛的业务应用和数据中心应用。这款服务器可通过有效扩展来支持大规模的处理器容量和内存容量，同时还提供了广泛的可靠性和可用性特性。SPARC M8-8 服务器经过专门设计，可以选择划分为两个单独的物理域 (PDom)。该服务器有两种厂配的款式。用户可选购其中一个或两个 PDom。

服务器组件

以下小节将简要介绍 SPARC M8-8 服务器的主要组件。

CPU、内存和 I/O 单元机箱

SPARC M8-8 服务器装载在一个 CMIOU 机箱内 — 该机箱中装有处理器/内存板、服务处理器 (SP)、用于互连配件的连接器，以及电源和散热风扇。图 13 显示了 CMIOU 机箱的前后视图及其主要组件。

机箱前部配有热交换风扇模块和电源单元 (PSU) 以及互连配件（用于将 CMIOU 主板连接成一个系统）。机箱前部包括：

- » 8 个热交换风扇模块
- » 6 个热交换 N+N 冗余 PSU，200 VAC–240 VAC 时功率为 3000 W

CMIOU 机箱后部包括：

- » 最多 8 个 CMIOU 主板
- » 2 个冗余 SP

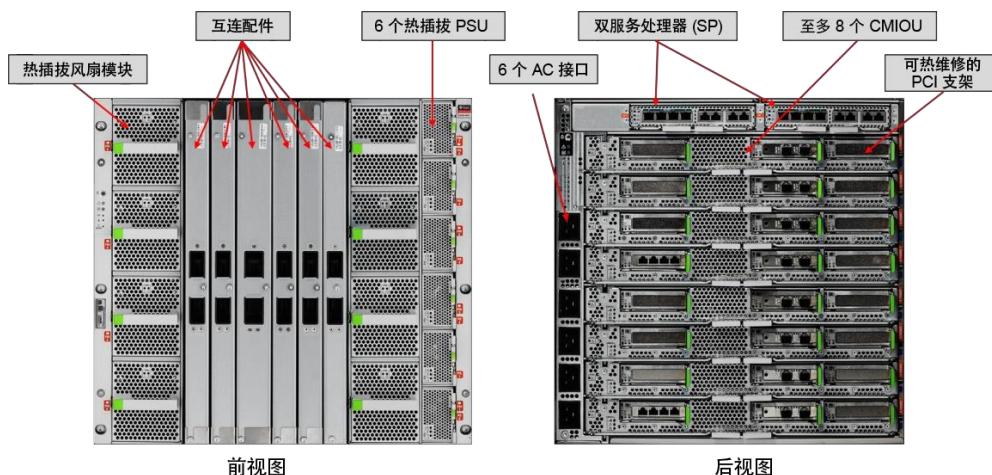


图 13.SPARC M8-8 CMIOU 机箱的前视图和后视图。



CPU、内存和 I/O 单元主板

每个 CMIOU 主板配件在一个夹层板上包含一个 SPARC M8 处理器，还包含相关的内存和 I/O 组件。全部 16 个内存 DIMM 插槽均位于主板上。一个 I/O 控制器 ASIC 为三个 PCIe 3.0 (x16) 插槽提供专用根联合体。主板上随带 PCIe 热插拔支架。图 14 是 CMIOU 主板的原理图，图 15 是以自上而下的俯拍视图。插入到 CMIOU 机箱后，CMIOU 主板与互连配件相连接，互连配件则用于连接 CMIOU 主板与 SP。

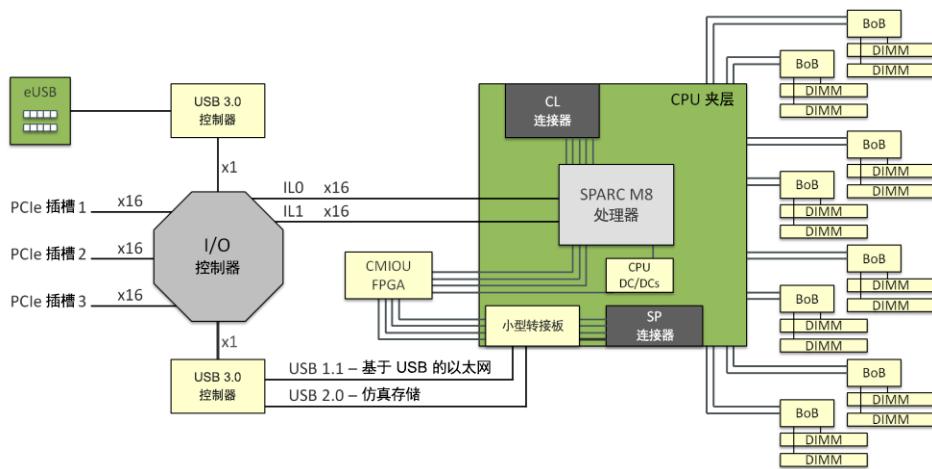


图 14. 单个 SPARC M8-8 CMIOU 主板的架构图。

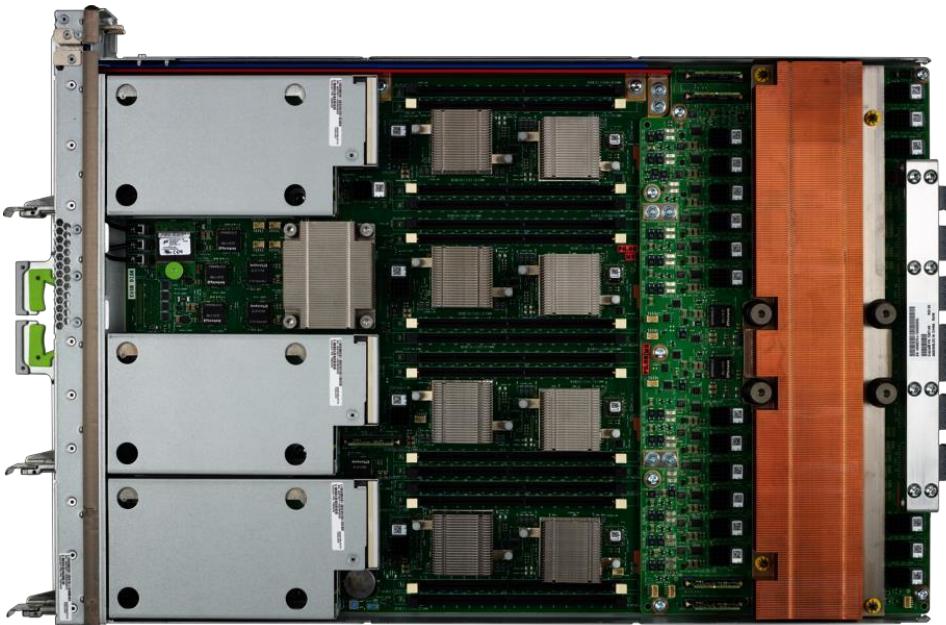


图 15. 每块 CMIOU 主板包含 1 个 SPARC M8 处理器、16 个 DDR4 DIMM 插槽、1 个 I/O 控制器 ASIC 和 3 个 PCIe x16 热插拔支架。

互连配件

SPARC M8-8 服务器通过互连配件实现系统互连。互连配件包含在 CMIOU 机箱的 10U 空间内。5 个互连配件在 CMIOU 主板之间提供互联链路 (CL) 连接（无缝系统互连）。一个互连配件用于 CMIOU 主板与机箱顶部的双 SP 之间的通信。互连配件中的布线在两种款式的 SPARC M8-8 服务器中是不同的，以便支持一个或两个电子隔离的 PDom。

服务处理器和服务处理器模块

SPARC M8-8 服务器配有冗余的热插拔服务处理器 (SP)。CMIOU 机箱中有 2 个 SP，每个 SP 包含 1 个或 2 个串行管理端口 (RJ45) 和 1 个 1000BASE-T 管理端口（自动协商至 10/100/1000 Mb/秒）。SP 通过 SP 互连配件与 CMIOU 主板进行通信。

每个 SP 包含 1 个或 2 个服务处理器模块 (SPM)。SPM 组件负责运行 Oracle ILOM 软件以及为服务器系统提供 SP 功能。为了始终具备冗余的 SP 功能和故障切换功能，在某些配置中，每个 SP 中均包含双 SPM。下面为各款服务器之间的差异：

- » 在只有单个 PDom 的 SPARC M8-8 服务器中，2 个 SP 中的每一个只包含一个 SPM。这足以提供冗余性。其中一个 SP 作为主用 SP 来管理平台，另一个作为备用 SP，以便在出现故障时担任主用 SP 的角色。
- » 在具有两个 PDom 的 SPARC M8-8 服务器中，每个 SP 包含两个 SPM，服务器使用两对 SPM 来管理系统中的两个 PDom。两个 SPM 提供主用 SP 的功能，其余两个作为备用。主用 SPM 和备用 SPM 分别位于单独的 SP 中。

系统机架和配电装置

SPARC M8-8 服务器通常作为出厂配置安装在 Oracle Rack Cabinet 1242 中。该机架包含 2 个配电装置 (PDU)，构成了一个三相双网电源。PDU 提供单相输出，用于给 CMIOU 机箱以及同一机架内的其他设备供电。PDU 包含一个电源监视模块，该模块带有一个串行 (RJ45) 端口和一个 100Base-T (RJ45) 管理端口。

Oracle 建议订购安装于机架中的厂配 SPARC M8-8 服务器。不过，客户也可以选择购买现场架装的独立系统。独立式 SPARC M8-8 服务器机柜为 10 个单元 (10U) 高。此外，还在机架底部或顶部留出 3 个单元的空间来容纳三相电源线。PDU 垂直安装在机柜中，不占用机架单元。当 Oracle Rack Cabinet 1242 中安装了单个 SPARC M8-8 服务器时，它还可以为其他设备提供 29 个单元。

SPARC M8-8 服务器架构

SPARC M8-8 服务器的两种款式均支持多达 8 个处理器。只有一个 PDom 的 SPARC M8-8 服务器配置为单个对称多处理 (SMP) 服务器，具有两个 PDom 的 SPARC M8-8 服务器则包含两个电子隔离的硬件分区。所有 SPARC M8-8 服务器硬件都装载在一个 CMIOU 机箱中 — 该机箱连同双 PDU 在出厂前就已安装到机架上，Oracle 建议采用这种厂配系统。对于 SPARC M8-8 服务器的两种款式来说，其大部分硬件组件是相同的。不过，互连配件中的布线是不同的，一种布线用于支持一个多达八路的 PDom，另一种布线用于支持两个相互隔离的 4 路 PDom。布线为出厂设置，之后无法更改。此外，服务处理器硬件也随支持的 PDom 数而变化，以便实现冗余性。



表 6 列出了 SPARC M8-8 服务器的主要特点。

表 6. SPARC M8-8 服务器特性。

特性	描述
最大处理器数	8 或 2×4
最大线程数	2048 或 2×1024
最大内存容量	8 TB 或 2×4 TB (基于 64 GB DIMM)
最大PCIe 3.0 (x16) 插槽数	24 或 2×12
服务处理器 (SP)	一个 PDom: 2 个热插拔 SP, 每个包含 1 个 SP 模块 (SPM) 从而为一个 PDom 提供冗余 两个 PDom: 2 个热插拔 SP, 每个包含 2 个 SPM, 从而为两个相互隔离的 PDom 提供冗余
风扇	8 个热交换 N+1 风扇模块, 每个模块带有 2 个对转风扇 前置式
电源	6 个热交换 N+N 冗余电源单元 (每个电源单元在 200 VAC–240 VAC 时功率为 3000 W); 前置式
PDU	系统机架中配有 2 个 PDU。每个 PDU 三条 3 相电源线, 用于双网格配置。 使用 6 条单相电源线将 PDU 连接至系统机箱。

只有一个物理域的 SPARC M8-8 服务器

只有一个 PDom 的 SPARC M8-8 服务器装载在一个 CMIOP 机箱中。这款服务器配置为单 PDom，最多支持 8 个 SPARC M8 处理器。SPARC M8-8 服务器中的每个 SP 均只有一个服务处理器模块 (SPM)，因为只需要一对 SPM 为 SP 功能提供冗余。默认情况下，第一个 CMIOP 主板 (插槽 0) 连接至 SP0/SPM0，作为主用设备。插槽 1 中的 CMIOP 主板提供与 SP1/SPM0 的备用连接。图 16 展示了 SPARC M8-8 服务器的 I/O 设备总图，其中包括 I/O 控制器 ASIC 的专用根联合体支持下的总共 24 个 PCIe 3.0 (x16) 插槽。

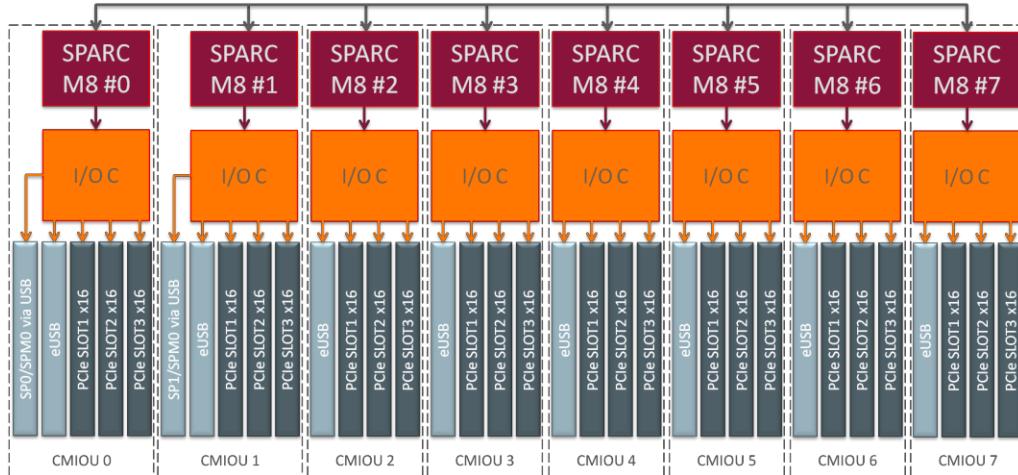


图 16.这款 SPARC M8-8 服务器包含 8 个 SPARC M8 处理器，出厂配置为单 PDom。



在八路无缝 CPU 间互连结构中，每个处理器通过 7 条 x16 互联链路 (CL) 连接至其余的 7 个处理器（图 17）。

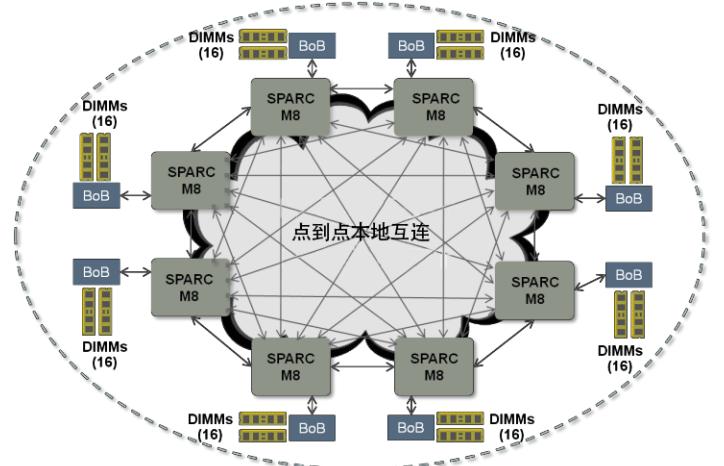


图 17.SPARC M8-8 服务器的八路完全互连结构。

具有两个物理域的SPARC M8-8 服务器

具有两个 PDom 的 SPARC M8-8 服务器同样装载在一个 CMIOU 机箱中，但出厂前电子分区为两个 PDom。其互连配件的布线与只有一个 PDom 的 SPARC M8-8 服务器不同。

两个 SP 各包含两个 SPM，以便为两个 PDom 提供冗余的服务处理器功能。一个 PDom 与每个服务处理器中的一个 SPM 相关联，于是两个 SPM 形成一个单活 (active/passive) 对。与 SPM 的连接通过每个 PDom 中的前两个 CMIOU 主板来实现。如果主用 SPM 或 SP 出现故障，则 SP 单活功能切换到备用连接。

图 18 是具有两个 PDom 的 SPARC M8-8 服务器的 PCIe 设备总图。该图展示了 PDom 与 SP 之间的连接。默认情况下，对于 PDom 0，插槽 0 中的 CMIOU 主板与 SP0/SPM0 的连接为主用连接。插槽 1 中的 CMIOU 主板提供与 SP1/SPM0 的备用连接。同样，对于 PDom 1，插槽 4 和插槽 5 中的 CMIOU 主板分别提供与 SP0/SPM1 的主用连接以及与 SP1/SPM1 的备用连接。该设备图还展示了与 I/O 控制器 ASIC 的专用根联合体支持的 PCIe 3.0 插槽的 x16 连接。

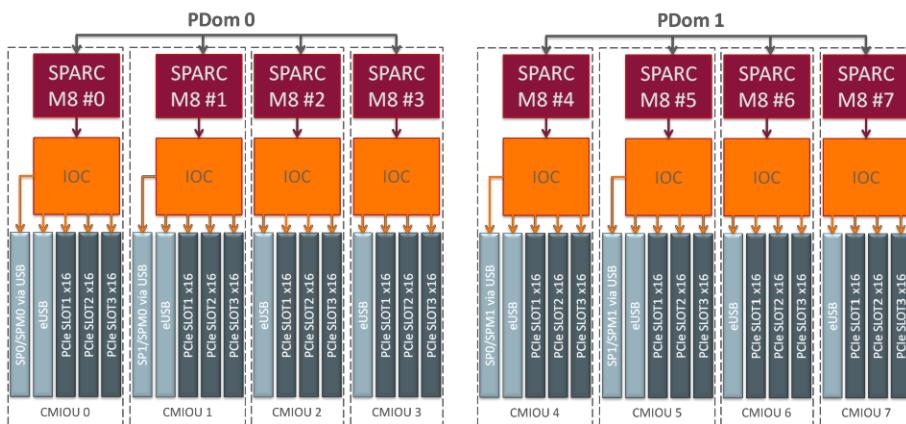


图 18.可以订购出厂配置为两个电子隔离 PDom 的 SPARC M8-8 服务器。



图 19 展示了具有两个 PDom 的 SPARC M8-8 服务器中的无缝系统互连。这种系统配置包含两个单独的四路完全互连结构。在这种四路 CPU 间互连结构中，每个处理器通过六条 x16 互联链路 (CL) 连接至所有其他处理器。

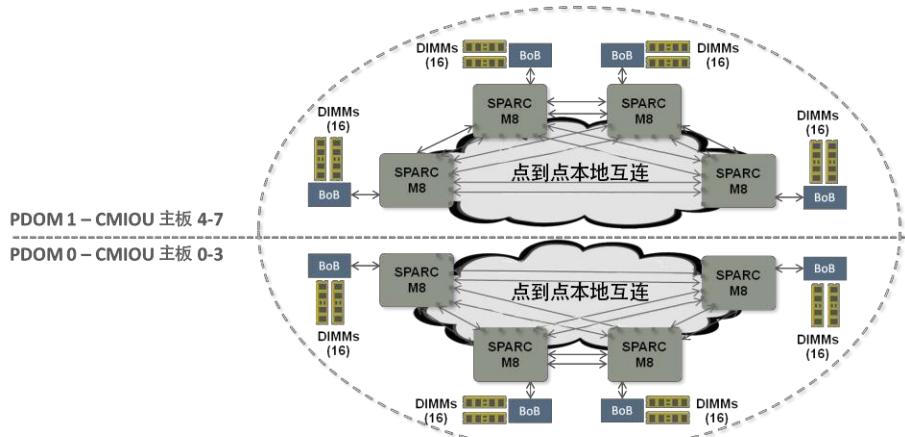


图 19. 具有两个 PDom 的 SPARC M8-8 服务器包含两个完全互连结构。

Oracle Solaris

无论是纵向还是横向伸缩环境，Oracle Solaris 都能为它们提供用于虚拟化的关键功能、极高的利用率、高可用性、强大的安全性以及超强性能。Oracle Solaris 可以在各种基于 SPARC 和基于 x86 的系统上运行，并确保与现有应用兼容。经过专门设计的 Oracle Solaris 11.3 版本可充分利用基于 SPARC M7 和 SPARC M8 处理器的服务器的大量资源，包括一整套高级软件芯片化特性。只有独占式使用服务域（例如，控制域、根域或 I/O 域）提供的虚拟磁盘、控制台和网络设备的来宾域才提供对 Oracle Solaris 10 的支持。来宾域不包含物理 I/O 设备，而是依赖服务域。表 7 列出了基于 SPARC M8 处理器的服务器中的各种域所支持的 Oracle Solaris 11 和 10 版本。

表 7. 基于 SPARC M8 处理器的服务器支持的 ORACLE SOLARIS 版本。

Oracle Solaris 版本	控制域	根域	I/O 域	来宾域
Oracle Solaris 11.3 SRU 24 ¹	✓	✓	✓	✓
Oracle Solaris 10 1/13 ²				✓

1. 基于 SPARC M8 处理器的服务器不支持 11.3 SRU 23 之前的 Oracle Solaris 11 版本

2. 外加所需补丁

基于 SPARC M8 处理器的系统极具吸引力的一个特点是，在 Oracle Solaris 及其所支持的应用看来，这些系统是它们所熟悉的 SMP 系统。此外，Oracle Solaris 引入了许多特性来提高 Oracle 多核/多线程架构上的应用性能，例如：

» **Oracle Solaris 软件定义网络 (SDN)。** Oracle Solaris 11 增强了 Oracle Solaris 现有集成式软件定义网络技术，无需昂贵网络硬件的额外开销即可提高应用敏捷性。该技术现在支持：基于一组完全分布式系统的应用驱动的多租户云虚拟网络；与物理网络基础设施的分离；应用级网络服务级别协议 (SLA) — 所有这些都作为平台的一部分内置于平台中。增强特性和新特性包括：

- » 通过虚拟 NIC (VNIC) 和虚拟交换实现网络虚拟化
- » 通过网络资源管理和集成的服务质量 (QoS) 对 VNIC 和流量流施加带宽控制
- » 云就绪。这得益于 Oracle Solaris 11 中随带的 OpenStack 发行版的一项核心特性
- » 通过 Oracle Solaris 弹性虚拟交换和虚拟可扩展 LAN (VXLAN) 实现应用驱动的多租户云虚拟网络
- » 通过应用驱动的 SDN 提供应用级 QoS
- » 与 Oracle Solaris Zones 和 Oracle Solaris 10 Zones 紧密集成

» **生命周期管理。** Oracle Solaris 11 包含一组全面、集成的技术，可用于管理平台的软件生命周期。Oracle Solaris 11 支持通过 Oracle Solaris Automated Installer 实现安全的端到端供应，支持通过 Oracle Solaris 映像打包管理系统实现有故障保护的软件更新，支持 ZFS 引导环境，并支持通过 Oracle Solaris Unified Archives 快速部署应用和全面的合规性框架，因此有助于提高工作效率，减少人为错误以及大大降低 IT 成本。

» **加速加密。** 与 SPARC M8 处理器一样，Oracle Solaris 中的加密框架也提供了对加速加密的支持。SPARC M8 处理器允许访问加密算法的硬件实现。这是第一次通过用户级指令，在相应的流水线本身中实现加密，而不是采用协处理器来实现。这意味着，基于硬件的加密实现不仅更高效，而且无需更改权限，因此大幅提高了加密算法计算的效率。此外，数据库操作可以更加高效地使用在指令流水线本身内实现的各种加密算法。

- » **关键线程优化。** Oracle Solaris 11 和 Oracle Solaris 10 允许用户或程序员通过使用命令行界面 (CLI) 或调用系统函数来将线程优先级提高到 60 或以上，从而让 Oracle Solaris 调度程序能够识别出关键线程。Oracle Solaris 调度程序通常在所有的系统内核间平均分配线程。然而，当某线程在 FX60 模式下运行时，调度程序会尝试让这个线程能够独占式地访问整个内核，而将其余的可运行线程分配给其他可用内核。
- » **多核/多线程感知。** Oracle Solaris 11 和 Oracle Solaris 10 了解 SPARC M8 处理器的层次结构，因此 Oracle Solaris 调度程序能够有效地在所有可用流水线间均衡分布负载。尽管 Oracle Solaris 将每个处理器均视为 256 逻辑处理器，但它了解内核集群之间、二级和三级缓存层次结构之间以及内核与其所支持的线程之间的相关性，从而可以提供快速、高效的线程执行。
- » **细粒度管理。** 对于 SPARC M8 处理器，Oracle Solaris 11 和 Oracle Solaris 10 能够分别启用或禁用各个内核和线程（逻辑处理器）。此外，标准的 Oracle Solaris 特性，如处理器集，提供了定义一组逻辑处理器并在其上调度进程或线程的功能。
- » **绑定接口。** Oracle Solaris 提供了相当大的灵活性，因为进程和各个线程可以根据需要绑定到一个处理器或一个处理器集。
- » **支持虚拟化网络和 I/O。** Oracle Solaris 提供了相应的技术来支持和虚拟化 SPARC M8 处理器上的组件和子系统。作为高性能网络架构的一部分，Oracle 多核/多线程感知型设备驱动程序让运行于虚拟化框架中的应用能够有效共享 I/O 和网络设备。
- » **Oracle Solaris 中的非一致性内存访问 (NUMA) 优化。** 内存由各个 SPARC M8 处理器分别管理，这些实现表示这是一种 NUMA 架构。在 NUMA 架构中，处理器访问自己内存所需的时间比访问其他处理器管理的内存所需的时间略短。Oracle Solaris 提供了以下技术来专门帮助减小 NUMA 对应用的影响并提高 NUMA 架构上的性能：
 - » **内存布置优化 (MPO)。** Oracle Solaris 使用 MPO 来改善整个服务器物理内存的内存布置，从而提高性能。借助 MPO，Oracle Solaris 可确保内存尽可能靠近访问它的处理器，同时又保持整个系统的平衡。因此，MPO 大大加快了许多数据库应用的运行。
 - » **分层 Lgroup 支持 (HLS)。** HLS 可帮助内存延迟层次较复杂的系统提高性能，从而优化 Oracle Solaris 中的 MPO 特性。借助 HLS，Oracle Solaris 可以辨别内存与处理器的远近程度，从而为应用分配延迟尽可能低的资源。如果默认情况下特定应用无法获得本地资源，则 HLS 会帮助 Oracle Solaris 分配相距最近的资源。
- » **Oracle Solaris ZFS。** Oracle Solaris ZFS 是全球独有的 128 位文件系统，在数据管理、自动化和整合复杂存储管理的理念以及提供无限可扩展性方面取得了巨大进步。Oracle Solaris ZFS 基于一种事务处理对象模型，该模型消除了对 I/O 发送指令的大部分传统约束，从而大幅提升性能。此外，Oracle Solaris ZFS 还提供数据完整性，通过 64 位校验和来检测和纠正静默数据损坏，从而保护所有数据。
- » **多路径软件。** 借助 Oracle Solaris 中的多路径软件，组织能够定义和控制通往存储设备和网络接口等 I/O 设备的冗余物理路径。如果通向设备的活动路径变为不可用，则该软件可以自动切换（故障切换）至备用路径以维持可用性。要利用多路径功能，服务器必须配有冗余硬件，如冗余网络接口或两个连接至同一双端口存储阵列的主机总线适配器。
- » **OpenStack 云管理。** Oracle Solaris 11 包括完整的 OpenStack 发行版，这样管理员能够通过单一管理界面集中共享和管理数据中心资源，包括其他供应商提供的基础设施和虚拟化资源。Oracle Solaris 上的 OpenStack 已集成到 Oracle Solaris Zones、ZFS 文件系统、统一存档和全面的 SDN 等核心技术基础中，因此可提供自助式计算，从而让 IT 组织能够在数分钟而非数周内提供服务，并提供企业级可靠性、安全性和性能。
- » **安全强健的企业级环境。** 现有 SPARC 应用可继续在基于 SPARC M8 处理器的系统上运行，且无需进行任何更改，从而保护了企业的软件投资。经过认证的多级安全性可以保护 Oracle Solaris 环境免遭入侵。Oracle Solaris 中的故障管理器架构可确保 Oracle Solaris 预测性自我修复等特性能够与硬件直接通信，从而减少计划停机和意外停机。Oracle Solaris DTrace 等高效的工具可帮助组织调优应用，从而充分利用系统资源。



虚拟化

由于各企业力求将各种负载整合到数量更少、功能更强大的系统上，同时提高利用率，因此虚拟化成为了一项重要的技术。基于 SPARC M8 处理器的服务器包含内置虚拟化功能：多个 PDom（具有双 PDom 的 SPARC M8-8）、Oracle VM Server for SPARC，此外 Oracle Solaris Zones 提供基于 OS 的虚拟化形式的第三层支持。

- » PDom 用于将一个硬件系统划分为多个安全隔离和故障隔离的服务器或硬件分区。
- » 逻辑域 (LDom) 通过 Oracle VM Server for SPARC 创建，用于对服务器或 PDom 进行虚拟化，以便可以托管多个分别运行自己的 Oracle Solaris 实例的虚拟机 (VM)。Oracle VM Server for SPARC 是一项免费特性，包含在所有 Oracle SPARC 服务器中。
- » Oracle Solaris Zones 可实现操作系统虚拟化，即只需一个 Oracle Solaris 实例即可安全隔离不同的应用，同时为各个区域分配系统资源。这实质上允许在单个 Oracle Solaris 操作系统实例内创建多个 VM。

这些虚拟化技术互为补充。事实上，优秀实践通常包括分层虚拟化，即部署两种或全部三种技术来实现卓越的安全性、可用性、性能和可管理性。此外，无需增加成本即可获得 Oracle 虚拟化技术。图 20 展示了基于 SPARC M8 处理器的服务器中提供的虚拟化技术，本文稍后将详细介绍此方面的信息。

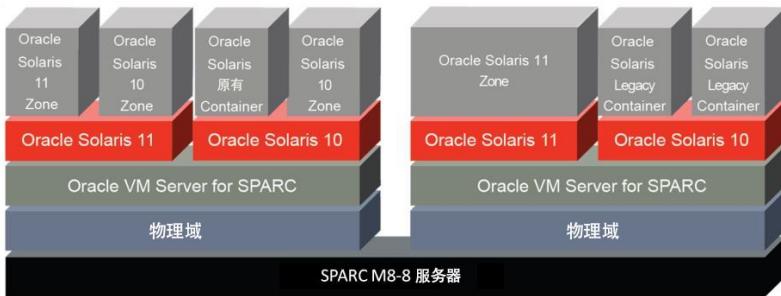


图 20.SPARC M8-8 服务器可支持双 PDom，SPARC T8-1、T8-2 和 T8-4 服务器则提供单个 PDom。

借助 SPARC M8-8 服务器上的 PDom, IT 组织能够将一个硬件系统划分为多个安全隔离和故障隔离的服务器。客户可以订购出厂配置为两个 PDom 的 SPARC M8-8 服务器。通过合理配置, 一个域中的软件故障或软件故障将保持隔离, 不会影响其他域的运行。

与前几代 SPARC 处理器一样, SPARC M8 处理器也支持 Oracle VM Server for SPARC, 这是一个虚拟机管理程序, 可用于在单个 PDom 中, 或者 SPARC T8-1、T8-2、T8-4 和 M8-8 服务器中的单个服务器中创建逻辑域 (LDom)。与以往的虚拟机管理程序不同, 该虚拟机管理程序是一个小型固件层, 可提供一个与处理器紧密集成的稳定虚拟机架构。由于虚拟机管理程序直接与底层多核/多线程处理器交互, 因此多线程至关重要。所有采用 Oracle 多核/多线程技术的服务器均支持 Oracle VM Server for SPARC, 后者可提供运行独立操作系统实例的完全虚拟机。这些完全虚拟机可配置为根域 (在这种情况下, 会为其分配根联合体, 以便直接访问 I/O 设备), 也可以配置为来宾域 (在这种情况下, 它们将访问虚拟化 I/O 设备)。每一个操作系统实例均包含处理器、内存、存储、控制台和加密设备。

Oracle Solaris 11 提供了一个名为 Oracle Solaris Zones 的独有 OS 容器技术。该技术可用于创建隔离和安全的环境来运行应用。Oracle Solaris Zones 是在单个 Oracle Solaris 实例中创建的一个高度轻量级虚拟操作系统环境。Oracle Solaris Zones 可用于将应用和进程与系统的其余部分隔离开来。这种隔离可确保不同 Oracle Solaris 区域中的进程互不干扰, 因此有助于提高安全性和可靠性。多处理器系统中的虚拟 CPU (或 SPARC M8 处理器中的线程) 可以分配为动态或固定 CPU 资源, 从而实现非常高效或强分区的配置, 或两者的混合。该共享资源或专用资源可以配置为 Oracle Solaris Zone 设置的一部分, 或者也可以通过动态重新配置来适应业务需求。Oracle Solaris Zones 还可以在 Kernel Zones 模式中运行, 该模式中可以运行完全不同的 Solaris 11 版本, 这些 Solaris 11 版本的修补时间可以不同于托管 OS 全局区域的修补时间。此外, Oracle Solaris 11 还提供了 Immutable Zones 特性, 可以实现一个适用于 Oracle Solaris Zones 和托管 OS 全局区域以及任何 Oracle VM Server for SPARC 域的只读根文件系统来锁定区域。除非作为特定的维护操作执行, 否则阻止对系统二进制文件或系统配置进行修改。

Oracle 的分层虚拟化技术可用于创建弹性的高可用性系统。例如, Oracle 集成系统和 Oracle 最高可用性架构利用这些虚拟化技术和虚拟化优秀实践来实现极高的可靠性、可用性和可维护性 (RAS)。要了解有关虚拟化的更多信息, 请参阅第39页的表 9 中列出的资源。

系统管理

对于大多数组织来说，为服务器系统提供需动手操作的本地系统管理已不再现实。全天候系统运行、灾难恢复热站和地理上分散的组织需要远程系统管理。在 Oracle 服务器的众多优势之中，一项优势就是支持无人值守的数据中心，这让宝贵的支持人员能够在网络通达的任意位置开展工作。基于 SPARC M8 处理器的服务器采用强大的服务处理器 (SP) 运行 Oracle Integrated Lights Out Manager (Oracle ILOM) 软件；该软件与 Oracle Enterprise Manager Ops Center 软件共同帮助管理员远程执行和控制不需触碰硬件的几乎所有任务。这些管理工具和远程功能可减轻管理负担，帮助组织节约时间并降低运营成本。

Oracle ILOM 和服务处理器

每个 SP 上运行的 Oracle ILOM 软件为 SPARC 服务器提供核心的远程监视和管理功能。SP 包含一个专用处理器，该处理器独立于服务器系统，运行 Oracle ILOM 软件包。大型的 SPARC M8-8 服务器提供冗余的 SP，这些 SP 能够自动故障切换并支持热维护性，从而确保持续运行。只要服务器未断电，即使所有域都处于非活动状态，SP 也会持续监视系统。

SP 会定期监视环境传感器、提供潜在错误情况的预先警告，并根据需要执行主动系统维护过程。例如，在可能导致物理系统损坏的温度状况下，SP 会触发服务器关闭。在 SP 上运行的 Oracle ILOM 软件包可以帮助管理员远程控制和监视物理域、虚拟机以及硬件平台本身。

借助 SP 的网络或串行连接，操作员可以在网络上的任何位置有效管理服务器。SP 的远程连接独立于操作系统并且提供了对系统控制台的完全控制和授权。Oracle ILOM SP 可充当系统控制器，帮助进行远程管理。SP 功能完备，其实施方式与 Oracle 的其他服务器中所使用的 SP 类似，因此，服务器可与现有管理基础设施轻松集成。Oracle ILOM SP 对有效开展系统管理至关重要，其作用包括：

- » 部署一个符合 IPMI 2.0 标准的 SP，为服务器固件、操作系统和应用提供 IPMI 管理功能，同时为通过 ILOM 4.0 以太网管理接口访问该 SP 的、基于 IPMI 的管理工具提供 IPMI 管理功能。该 SP 还可洞察服务器模块上以及机箱中其他位置的环境传感器。
- » 管理服务器设备清单和环境控制，包括处理器、DIMM、风扇和电源，并提供对这些数据的 HTTPS、CLI 和 SNMP 访问。
- » 提供远程文本控制台界面。
- » 提供一种将升级软件下载到所有系统固件的方法。

此外，Oracle ILOM 和 SP 还让管理员能够远程管理服务器，这不受该平台上运行的操作系统约束，也不会干扰任何系统活动。Oracle ILOM 可通过电子邮件发送硬件故障警报、警告以及与服务器相关的其他事件。其电路系统独立于服务器运行，使用服务器的备用电源。因此，当服务器操作系统离线或服务器关机时，Oracle ILOM 固件和软件将继续运行。

Oracle ILOM 可监视以下服务器状况：

- » CPU 温度状况
- » 硬盘驱动器状态
- » 机柜热状态
- » 风扇转速和状态
- » 电源状态
- » 电压状况
- » Oracle Solaris 预测性自我修复、启动超时和服务器自动重启事件

电源管理

降低不断攀升的服务器电力和散热成本是企业数据中心面临的头号挑战。扩展数据中心时面临的电力和空间限制迫使客户密切关注服务器的电源效率。他们与电力公司签署的合同规定，用电量超出规定将受到处罚，这就要求服务器要能将功耗限制在客户的控制范围内。此外，电源效率和碳排放量也已成为客户对服务器的重要考量因素。

除了 Oracle 多核/多线程设计固有的高效性，SPARC M8 处理器还在处理器内核和内存这两个级别引入了的电源管理特性。这些特性包括降低指令速率，寄存闲置线程和内核，以及能够关闭内核和内存中的时钟以降低能耗。此外，以前的 SPARC 处理器支持在芯片级进行电源管理，而 SPARC M8 处理器支持进行子芯片电源管理。SPARC M8 处理器分为两个区域，每个区域包含 16 个内核。每个分区可以独立于其他分区设置不同的耗电水平。这样便可为每个分区设置不同的性能级别，重要的是，当在不同分区中部署 LDom 时可以设置不同的 SLA。

除了在 Oracle ILOM 中提供电源管理支持以外，Oracle Solaris 11.3 及更高版本还提供了一个支持基于 SPARC M8 处理器的服务器的电源管理程序。Oracle Solaris 可根据 poweradm 设置来确定启用哪些节电特性。poweradm 设置由平台基于系统 (Oracle ILOM) 策略设定，不过可由 Oracle Solaris 管理员覆盖。其在多个领域实现了重大创新：

- » 限制推测，如不采用条件分支
- » 数据路径、控制块和阵列中广泛的时钟门控
- » 功耗节流，允许将额外的停滞周期注入到译码阶段

在使用 Oracle VM Server for SPARC 的虚拟化环境中，电源管理程序在管理 LDom 来宾系统时可执行以下任务：

- » 根据电源管理策略确定启用哪些节电特性
- » （对于不属于 Oracle Solaris 11.3 来宾系统的资源）通知电源管理引擎更改其资源的电源状态以调整功率或达到电能利用率，或者通知虚拟机管理程序启用或禁用虚拟机管理程序/硬件托管的电源状态。只有 Oracle Solaris 11.3 来宾系统具备电源管理对。



Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center 为基于 SPARC M8 处理器的服务器提供整合的硬件管理解决方案，可在整个基础设施体系内进行统一管理。Oracle Enterprise Manager Ops Center 提供先进的虚拟化管理和报告功能、应用到磁盘的管理、智能配置管理等，可帮助 IT 管理人员降低复杂性和简化基础设施管理。所有 SPARC 服务器中均包含 Oracle Enterprise Manager Ops Center，因此数据中心管理员从一个界面即可监视和管理存储、网络、服务器、Oracle Solaris 和虚拟化环境。这不仅有助于改善运营效率，而且还将降低运营成本。

Oracle Enterprise Manager Ops Center 是专为 Oracle 服务器和 Oracle 集成系统基础设施量身打造的全面的管理解决方案。Oracle Enterprise Manager Ops Center 提供一个控制台来管理多个服务器架构和大量操作系统，可使用资产发现、固件和操作系统供应、自动化补丁管理、补丁和配置管理、虚拟化管理和综合合规性报告来管理基于 SPARC M8 处理器的服务器中的组件（图 21）。

Oracle Enterprise Manager Ops Center 可通过基于策略的管理来实现工作流自动化并加强合规性，所有这些任务均可在一个直观界面中完成。采用 Oracle Enterprise Manager Ops Center，IT 员工可以实施和加强数据中心标准化和优秀实践，加强合规性和安全性策略，同时有效部署基础设施，从而满足业务需求。

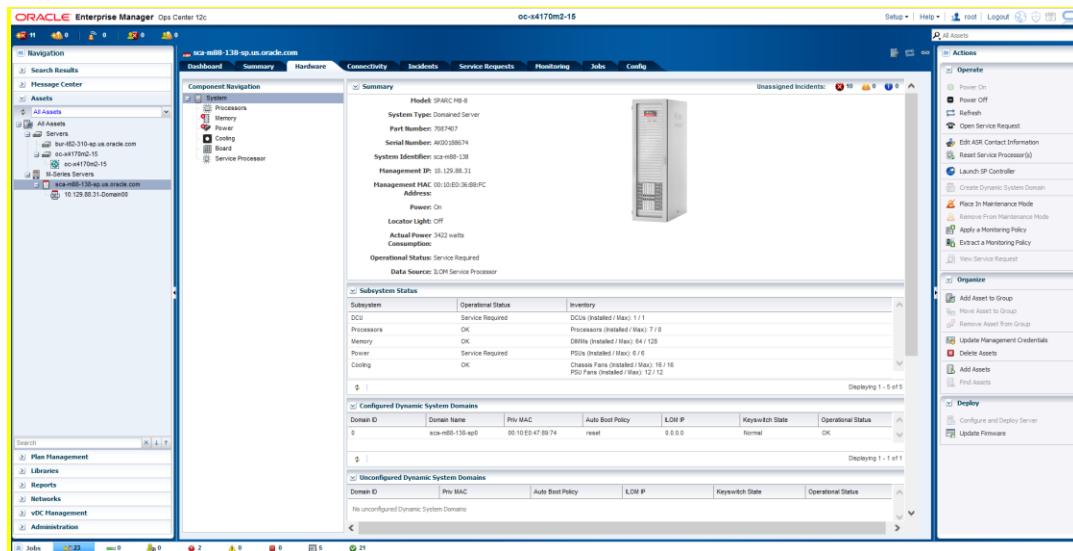


图 21.Oracle Enterprise Manager Ops Center 为 SPARC 服务器提供详细的管理功能。

可靠性、可用性和可维护性

减少计划停机和意外停机是 IT 服务的关键。系统设计中采纳的机制必须支持故障恢复、快速修复甚至迅速扩展，而又不影响关键服务可用性。为了支持复杂的网络计算解决方案和满足严格的高可用性 (HA) 需求，基于 SPARC M8 处理器的服务器的设计中考虑了冗余的热维护式系统组件，整体融入了诊断和错误恢复特性，并内置了远程管理特性。这些可靠的服务器所采用的先进架构能提高应用可用性，支持系统迅速从多种类型的硬件故障中恢复，从而帮助企业简化系统运营和降低成本。

高级可靠性特性

基于 SPARC M8 处理器的服务器组件的高级可靠性特性可提高服务器的整体稳定性。首先，服务器架构内组件数量和复杂性的降低有助于提高可靠性。其次，SPARC M8 处理器为自发错误恢复提供了高级 CPU 集成和有保障的数据路径完整性，从而缩短启动纠正措施的时间并延长正常运行时间。另外，关键组件中的冗余性支持故障切换；热插拔和热交换组件有助于提高可维护性；在基于 SPARC M8 处理器的服务器中，高级 I/O 控制器包含 PCIe 根联合体，这让它们在添加或删除处理器时能够保持不变。这些特性共同形成了一个非常强健的系统设计。

在固件和 Oracle Solaris 预测性自我修复软件中实现的故障管理架构进一步增强了 SPARC 服务器的可靠性。它可以提供对处理器、内存和 I/O 设备的持续监视。根据错误的性质，它可以自动将线程、内核或整个处理器置于离线来解决持久的 CPU 软错误。此外，内存分页退出功能支持主动将内存分页置于离线状态来响应针对特定内存 DIMM 的多项数据纠正。如前所述，基于 SPARC M8 处理器的服务器支持 DIMM 热备，DIMM 热备支持在不中断系统运行的情况下退出整个 DIMM，同时内存容量和错误保护功能不受影响。

错误检测、诊断和恢复

基于 SPARC M8 处理器的服务器采用了一些重要技术，可尽早纠正故障，防止边际组件导致重复性停机。服务器硬件子系统中的错误检测和恢复功能对架构进行了增强，进一步提高了可靠性。归根结底，服务器通过下列特性提高应用可用性：

- » 端到端的数据保护可检测和纠正整个系统中的错误，确保全面的数据完整性。
- » 先进的故障隔离功能可帮助服务器隔离组件边界内的错误，仅将相关芯片或芯片的一部分置于离线状态，而非让整个组件离线。将错误隔离到最小的实体有助于提高稳定性，并确保持续提供最大计算能力。此特性适用于处理器、内存 ASIC 和 DIMM、交换 ASIC、连接链路、I/O 控制器 ASIC 以及 SP。
- » 持续环境监视功能可提供相关环境和错误状况的全面历史记录。
- » 主机监视程序特性定期检查软件操作，包括域操作系统。此特性还使用 SP 固件触发错误通知和恢复功能。
- » 系统的故障管理器架构和 SPARC M8 处理器的动态 CPU 资源取消配置功能可支持强大的隔离和恢复。如有必要，系统可动态退出处理器资源（如内核），且不会中断运行中的应用。
- » 通过执行定期组件状态检查来确定多种系统设备的状态，检测潜在故障的信号。系统会触发恢复机制来防范系统和应用故障。
- » 错误日志记录、多级警报、电子可更换单元 (FRU) 识别信息和系统故障 LED 指示有助于快速解决问题。

冗余的热维护式组件

如今，IT 组织面临业务运营节奏的挑战 — 不能中断业务运营。在互联的全球经济环境中，创收机遇随时存在，这就要求尽量减少计划内停机，或者在某些情况下完全消除停机。为了满足这些需求，基于 SPARC M8 处理器的服务器采用了内置、冗余且支持热插拔和热交换的硬件，可避免因个别组件故障或系统配置更改而发生中断。实际上，这些系统往往能够从硬件故障中恢复 — 而且往往不会对用户或系统功能造成影响。



基于 SPARC M8 处理器的服务器采用冗余配置的热交换式电源和风扇单元，还可以选配多个处理器、内存 DIMM 和 I/O 卡。管理员可以结合使用热插拔磁盘驱动器和磁盘镜像软件来打造冗余式内部存储。SPARC M8-8 服务器还支持冗余式热插拔服务处理器。当发生故障时，这些冗余组件可以支持服务器继续运行。根据组件和错误的类型，系统可继续以降级模式运行，也可重新引导 — 自动诊断故障并自动配置为将相关组件排除到系统之外。此外，由于无需停止系统，这些服务器采用的热维护式硬件可加快维护工作并简化组件更换或添加操作。

可更换组件分为以下几类：

» **热维护**。可在服务器运行过程中移除和插入热维护式组件。

» 热交换组件在维修之前不需要做任何准备。

» 热插拔组件在维修之前需要做好准备。

» **冷维护**。冷维护式组件需要服务器停机才可进行维修。此外，一些维修过程还要求断开电力供应与电源之间的电源电缆。

表 8 列出了基于 SPARC M8 处理器的服务器中的主要热维护式组件。

表 8. 基于 SPARC M8 处理器的服务器中的关键客户热维护组件。

组件器	SPARC T8-1 服务器	SPARC T8-2 服务器	SPARC T8-4 服务	SPARC M8-8 服务器
2.5 英寸 SFF 驱动器	✓	✓	✓	不适用
电源	✓	✓	✓	✓
风扇/风扇模块	✓	✓	✓	✓
PCIe 卡	—	—	✓ ¹	✓
CMIQ 主板	不适用	不适用	不适用	✓ ²
SP	—	—	—	✓

✓ = 热维护，— = 冷维护，NA = 不适用

1. 卡支架中的 PCIe 卡。内部插槽中的 NVMe PCIe 交换卡是冷维护式组件。

2.CMIQ 主板所在的 PDom 必须关闭。其他 PDom 可以继续运行。

总结

要应对现代技术挑战和业务挑战，企业需要具备创新性的解决方案、强大的计算能力和各种外形规格的服务器才能实现快速发展的数字市场广阔前景。依托 Oracle 创新式 SPARC M8 处理器的强大优势，Oracle SPARC 服务器为组织提供了一种新方法，帮助组织以灵活、高度可用的方式高效应对要求极为严苛的计算负载。通过每个 SPARC M8 处理器上的专用加速引擎，创新式软件芯片化技术提供了芯片保护的内存特性、内存中查询加速特性、内联解压缩特性和用于增强安全性的加速加密特性，为应用带来颠覆性的性能增强。



基于 SPARC M8 处理器的服务器提供了各种强大功能，让组织能够将计算成本从底层基础设施转移出来，投入到高产能业务负载上。例如，更高水平的性能和效率让组织能够在同一系统上同时运行 OLTP 和数据分析。与 Oracle Solaris 功能相结合，Oracle 的分层虚拟化可支持组织创建高效的本地部署云，并提供卓越的容量和可伸缩性，从而缩短市场投放时间，提高盈利能力。

更多信息

有关更多信息，请访问表 9 中列出的资源。

表 9. 资源。

Web 资源	URL
Oracle SPARC 服务器	https://www.oracle.com/cn/servers/sparc/ 和 https://www.oracle.com/technetwork/cn/server-storage/sun-sparc-enterprise/overview/index.html
Oracle Solaris	https://www.oracle.com/cn/solaris/solaris11/
Oracle 优化的解决方案	https://www.oracle.com/cn/it-infrastructure/
Oracle 基准测试结果	oracle.com/benchmarks
Oracle 帮助中心文档	docs.oracle.com
安全性	https://www.oracle.com/technetwork/cn/topics/security/whatsnew/index.html
虚拟化	https://www.oracle.com/technetwork/cn/topics/virtualization/whatsnew/index.html
Oracle Integrated Lights Out Manager	https://www.oracle.com/technetwork/cn/server-storage/servermgmt/tech/integrated-lights-out-manager/index.html
Oracle Enterprise Manager Ops Center	https://www.oracle.com/technetwork/cn/oem/ops-center/index.html
白皮书和技术文章	URL
“Oracle SPARC T7 和 SPARC M7 服务器：引领优秀实践”	oracle.com/technetwork/server-storage/sun-sparc-enterprise/documentation/sparc-t7-m7-best-practices-2701865.pdf
“使用 Oracle SPARC 虚拟化技术进行整合”	oracle.com/technetwork/server-storage/sun-sparc-enterprise/technologies/consolidate-sparc-virtualization-2301718.pdf
“Oracle Solaris Kernel Zones 优秀实践”	oracle.com/technetwork/articles/servers-storage-admin/solaris-kernel-zones-best-practices-2400370.html
“Oracle VM Server for SPARC 优秀实践”	oracle.com/technetwork/server-storage/vm/ovmsparc-best-practices-2334546.pdf
Web 资源	URL
“通过 Oracle VM Server for SPARC 实现根域”	oracle.com/technetwork/server-storage/vm/ovm-sparc-rootdomains-wp-1914481.pdf
“在运行 Oracle Solaris 的 SPARC 服务器上使用 Oracle Multitenant”	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
电话: +1.650.506.7000
传真: +1.650.506.7200

关注我们

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

Integrated Cloud Applications & Platform Services

版权所有 © 2017, Oracle 和/或其关联公司。保留所有权利。本文档仅供参考，内容如有更改，恕不另行通知。本文档不保证绝对正确，也不受其他任何口头表达或法律暗示的担保或条件的约束，包括对特定用途的适销性或适用性的暗示担保和条件。我们特别声明拒绝承担与本文档有关的任何责任，本文档不直接或间接形成任何契约义务。未经预先书面许可，不允许以任何形式或任何方式（电子或机械的）、出于任何目的复制或传播本文档。

Oracle 和 Java 是 Oracle 和/或其关联公司的注册商标。其他名称可能是其各自所有者的商标。

Intel 与 Intel Xeon 是 Intel Corporation 的商标或注册商标。所有 SPARC 商标均为 SPARC International, Inc. 的商标或注册商标，需经许可方可使用。AMD、Opteron、AMD 徽标以及 AMD Opteron 徽标是 Advanced Micro Devices 的商标或注册商标。UNIX 是 The Open Group 的注册商标。0917

Oracle SPARC T8 和 SPARC M8 服务器架构
2017 年 9 月