

Oracle SPARC T8 和 SPARC M8 服务器的可靠性、可用性和可维护性

Oracle 白皮书 | 2017 年 10 月





| | |
|---|----|
| 目录 | |
| 引言 | 3 |
| 部署 RAS 架构..... | 2 |
| 基于 SPARC M8 处理器的服务器：融入到设计中的可用性 | 3 |
| 可靠性、可用性和可维护性特性..... | 3 |
| 错误检测、诊断和恢复 | 5 |
| 冗余组件 | 5 |
| 多 PCIe 根联合体..... | 6 |
| DIMM 热备 | 6 |
| 冗余服务处理器 | 7 |
| Oracle Solaris | 8 |
| Oracle Solaris Multipathing..... | 9 |
| Oracle Solaris Cluster | 9 |
| 虚拟化与 Oracle VM Server for SPARC..... | 10 |
| 面向高可用性基础设施的 Oracle 产品和解决方案 | 11 |
| Oracle Real Application Clusters (Oracle RAC) | 11 |
| Oracle WebLogic Server 集群..... | 11 |
| Oracle 优化解决方案和 Oracle 最高可用性架构 | 12 |
| 总结 | 13 |
| 更多信息..... | 13 |



引言

现代应用用户和企业对停机的容忍度很低，企业 IT 基础设施必须满足与业务需求相一致的应用可用性要求。对于某些应用，一定的停机时间是可以接受的。而对于另一些应用，任何停机都可能导致收入损失、公司声誉受损或错失商机。因此，组织需要确保其部署的系统能够满足严苛的要求，帮助他们满足服务级别协议 (SLA)，并提供与既定恢复点目标 (RPO) 和恢复时间目标 (RTO) 相一致的持续服务。

就此而论，可靠且可维护的系统是极其重要的。但是，真正的应用可用性需要完全集成的硬件和软件体系，还需要支持协同作业的系统组件。要取得成功，组织需要一种 *可用性架构*，将强健的冗余系统和存储技术与可靠的操作系统、中间件、虚拟化技术和数据库整合在一起。而集成所有这些技术并实现无缝协同作业，可能要花费很长的时间，付出高昂的成本，并且稍有配合不当就可能造成损失重大的灾难性故障和停机。

凭借基于 SPARC M8 处理器的服务器，Oracle 具有独特的优势，可通过完全集成的技术体系提供先进的系统架构。这些服务器提供高水平的可靠性、可用性和可维护性 (RAS) 特性，为构建卓越、高度安全的任务关键型计算系统奠定了基础。该集成体系包括 Oracle Solaris 操作系统和 Oracle 先进的中间件和数据库技术，集虚拟系统要素和物理系统要素于一体，创造了一个具有高水平 RAS 的应用运行平台。此外，Oracle 优化系统和 Oracle 最高可用性架构充分利用该平台独特的 RAS 特性，基于可靠的优秀实践提供一个经过测试、可扩展且高度可用的基础设施。

部署 RAS 架构

在设计高可用性 IT 基础设施时，通常只关注系统可靠性特性。事实上，高质量的应用服务依赖于整个硬件和软件体系的 RAS 功能。尽管最终重点是在应用可用性上，但在考量交付平台或数据中心解决方案的功能时，可靠性和可维护性同等重要。以下是要考虑的一些因素：

- » **可靠性。**可以采用多种方式来提高可靠性。例如，可以减少或最小化系统中的组件数量。组件数量的减少，甚至于组件之间连接数量的减少，可以带来更长的平均无故障时间 (MTBF)。对所有静态数据和传输中的数据实施数据完整性保障之后，系统将能够自动检测和纠正错误，避免因错误而损坏数据、导致应用或系统故障，从而显著提高可靠性。
- » **可用性。**可用性一般被定义为在底层组件故障期间提供不间断的应用访问的能力。此目标一般通过平台冗余、应用可用性特性、虚拟化和集群来实现。数据可用性包括各种机制，例如基于 SPARC M8 处理器的服务器提供的 DIMM 热备功能，以及内存通道热备和处理器互连通道热备等。数据可用性也可以通过软件来提高，例如 Oracle Solaris 或 Oracle Solaris ZFS 中的 I/O 多路径软件，可为磁盘设备提供先进的 RAID 保护。在许多情况下，可用性均在应用级提供。但万一在此级别没有提供时，可以通过集群软件在虚拟或物理系统之间提供故障切换，进一步增强平台的可用性。
- » **可维护性。**高水平的可维护性可以在极短的停机时间内对系统进行维护或修理，以此提高可用性。可维护组件可以在系统保持运转时进行热交换或热插拔。热交换组件可大幅减少中断，因为无需主机操作系统和服务器的协助或准备即可进行更换。热插拔组件也可以在系统继续运转时进行替换，但它们要求操作系统和平台在移除或插入组件之前做好准备。

任何应用可用性架构的强度均取决于其最薄弱的一环。而周详的设计能将系统硬件和软件各层中相辅相成的 RAS 特性整合到一个可用性架构（图 1）中，切实提高可靠性。通过这种方式，每一个要素均以下一层的优势为基础，共同在一个复杂系统中增强总体 RAS 能力。每一层的各个组件通过扬长避短打造了一个非常强健的系统架构。

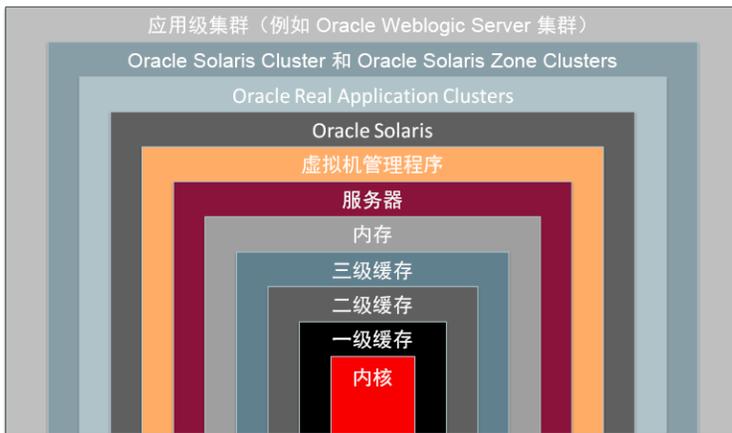


图 1. 将系统每层的硬件和软件元素的 RAS 功能整合在一起的可用性架构。

系统软件层为 RAS 功能做出了巨大贡献。举例来说，Oracle Solaris 操作系统和 Oracle Integrated Lights Out Manager (Oracle ILOM) 固件中作为 Oracle Solaris Fault Management Architecture (FMA) 的一部分实施的预测性自我修复功能，是确保硬件和应用软件高可用性的基本要素。此外，可以采用、中间件和集群软件（包括数据库集群）来进一步提高 RAS 级别。以下是 Oracle SPARC 服务器所提供的几个技术示例：

- » Oracle 的分层虚拟化技术可以帮助隔离系统映像 — 提供冗余和硬件独立性，并且便于在系统之间轻松迁移应用。
- » Oracle Solaris Cluster 软件在物理或虚拟系统之间提供系统级的冗余和故障切换。
- » Oracle WebLogic Server 为要求高性能、高扩展能力和高可靠性的应用提供了任务关键型云平台。
- » Oracle Real Application Clusters (Oracle RAC) 提供了一个有共享缓存架构的集群数据库，可实现高度可扩展和可用的数据库解决方案。

值得注意的是，有些提供 RAS 功能的特性是以牺牲其他特性为代价的。例如，将多个特性集成到单一芯片或者将所有芯片和 ASIC 安装到单一主板中有助于提高可靠性，不过会降低可用性（因为会延长平均维修时间）和可维护性。此外，热交换和热插拔组件可能需要额外的连接器，这样尽管提高了可用性，但从技术上降低了可靠性。因此，特性取舍始终是一个缜密的设计过程，因为要在系统成本与数据服务可用性和可靠性之间取得平衡。

基于 SPARC M8 处理器的服务器：融入到设计中的可用性

SPARC 服务器的设计初衷是用于支持要求严苛的企业环境和托管任务关键型应用。这些服务器设计了众多 RAS 特性，从先进的错误检测和纠正机制，到处理器和内存保护以及冗余组件。这些 SPARC 服务器和 Oracle Solaris 采用独有的架构设计，可用于部署支持预测性自我修复的服务。Oracle Solaris FMA 接收与硬件和软件错误相关的数据，自动诊断底层问题，并在可能时将故障组件置于离线状态作为回应。

可靠性、可用性和可维护性特性

从处理器、虚拟机管理程序到整个系统，SPARC 服务器提供了许多可靠性特性。这些特性互为补充、相互增强，使整个系统具有更强的恢复能力。基于 SPARC M8 处理器的服务器中实现的 RAS 特性包括：

- » Oracle SPARC M8 处理器
 - » 一级缓存标记、状态和数据：奇偶校验保护和出错重试
 - » 二级和三级缓存数据：SEC/DED 保护、内联纠正和缓存行热备
 - » 二级和三级缓存状态和目录：SEC/DED 保护和内联纠正
 - » 架构寄存器：SEC/DED 保护、准确陷阱和虚拟机管理程序纠正/重试
 - » 硬件中的消息重试
 - » 动态处理器核心取消配置
 - » 处理器每个象限的独立管理功率
 - » 链路级 RAS：自动帧重试、自动链路重新训练、链路级多路径和单通道故障切换
 - » 动态电压和频率调节

»主内存

- » SDRAM 错误保护：纠正单设备错误，以及跨设备执行三重错误检测
- » 内存通道互连：硬件中的消息重试、硬件中的通道停用、循环冗余检验 (CRC)
- » DIMM 热备支持通过停用全内存 DIMM 来防范硬错误。当发生持久的可恢复错误时，FMA 功能可以启动 DIMM 热备并停止使用可靠性出问题的 DIMM。DIMM 热备是自动的，不会导致应用服务中断。系统内存容量不变，错误保护保持不变，在执行 DIMM 热备之后不增加故障风险。
- » 如果持久的可恢复错误局限于某个特定的地址，系统将使用“分页停用”功能来防范硬错误。即使在对同一内存条应用 DIMM 热备之后，此功能仍保持可用。

» Fault Management Architecture 软件

- » SP 和 Oracle Solaris 上的诊断引擎
- » 在发生故障时自动重新配置
- » 软错误率判别 (SERD)
- » 错误分页停用
- » 操作系统和 SP 监视程序

» SPARC M8 虚拟机管理程序

- » 逻辑域 (LDom) 虚拟化和故障隔离
- » CPU、内存、IO 的动态分配
- » 动态 PCIe 总线分配
- » 为清除、纠正和收集错误提供处理器支持

» 系统

- » 初次故障时，在可更换单元 (FRU) 层面执行诊断
- » 冗余式可热维护服务处理器 (SP) 和服务处理器代理 (SPP)，支持自动故障切换（适用于 SPARC M8-8 服务器）
- » 可热插拔的内置 2.5 英寸 SAS 硬盘驱动器 (HDD) 和 SAS/NVMe 固态硬盘 (SSD)（适用于 SPARC T8-1、T8-2 和 T8-4 服务器）
- » 支持多个物理域 (PDom)（适用于 SPARC M8-8 服务器）
- » 独立的时钟源和冗余时钟（适用于 SPARC M8-8 服务器）。每个 CMIOU 主板和交换板都有双时钟源，能够在不中断应用服务的情况下从故障中恢复，并继续以单个时钟源运行。

» 系统 I/O

- » PCIe 端到端 CRC
- » PCIe 链路重试
- » 每个 I/O 控制器多达五个 PCIe 根联合体，实现冗余式根域和 I/O 多路径
- » CPU 至 I/O 控制器的链路通道热备
- » 热插拔半高 PCIe 卡（在 SPARC T8-4 和 M8-8 服务器中）

» 供电和散热

- » 高级电源管理
- » 冗余式热交换风扇和风扇模块
- » 冗余式热交换电源
- » 双网电源
- » 冗余式 CPU 核心电源（在 SPARC M8-8 服务器中）

错误检测、诊断和恢复

基于 SPARC M8 处理器的服务器提供了一些重要技术，可尽早纠正错误并保证边缘组件不会导致意外停机。服务器硬件子系统内的错误检测和恢复功能增强了架构方面的改进，可从根本上提高可靠性。从根本上说，下列特性协同工作，共同提高了应用可用性：

- » 端到端的数据保护可检测和纠正整个系统中的错误，从而确保全面的数据完整性。
- » 先进的故障隔离功能可帮助该服务器将错误隔离在组件边界之内，只将相关项目、芯片或组件子区置于离线状态。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、内存缓冲 ASIC 和 DIMM、连接链路、I/O 控制器 ASIC 以及 SP。
- » 持续环境监测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，包括每个域的操作系统。此特性还使用 Oracle ILOM 固件来触发错误通知和恢复功能。
- » 系统的 FMA 功能和 SPARC M8 处理器的动态 CPU 资源取消配置功能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。
- » 系统定期执行组件状态检查来检测潜在故障的迹象，然后触发恢复机制以防止系统和应用故障。
- » 错误日志记录、多级警报、电子 FRU 识别信息和系统故障 LED 指示有助于快速解决问题。

冗余组件

SPARC 服务器采用冗余配置的热交换电源和风扇单元，还可以选配多个存储设备、内存 DIMM 和 I/O 卡。所有基于 SPARC M8 处理器的服务器（除 SPARC T8-1 服务器之外）均支持多个 SPARC M8 处理器。管理员可以结合冗余式热插拔内部存储与 RAID 硬件及软件来增强数据可用性。SPARC M8-8 服务器还支持冗余式热插拔 SP。当发生故障时，这些冗余组件可以支持系统继续运行。视错误的组件和类型情况，系统可以继续以降级模式运行，也可以重新启动 — 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，由于无需中断或停止系统，这些服务器的可热维护硬件加快了维护速度，并简化了组件的更换或添加。

可更换组件分为以下几类：

- » **热维护**组件可以在服务器运行时进行移除和插入。
 - » 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**组件需要关闭服务器进行维护。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 1 列出了基于 SPARC M8 处理器的服务器中的主要热维护组件。

表 1. 基于 SPARC M8 处理器的服务器中的关键客户热维护组件。

| 组件 | SPARC T8-1 服务器 | SPARC T8-2 服务器 | SPARC T8-4 服务器 | SPARC M8-8 服务器 |
|---------------------------|----------------|----------------|----------------|----------------|
| 2.5 英寸小尺寸 (SFF) 驱动器 | ✓ | ✓ | ✓ | 不适用 |
| 电源 | ✓ | ✓ | ✓ | ✓ |
| 风扇/风扇模块 | ✓ | ✓ | ✓ | ✓ |
| PCIe 卡 | - | - | ✓ ¹ | ✓ |
| CPU、内存和 I/O 单元 (CMIOU) 主板 | 不适用 | 不适用 | 不适用 | ✓ |
| 服务处理器 (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 PCIe 交换卡仅支持冷维护。

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

多 PCIe 根联合体

基于 SPARC M8 处理器的各款服务器其基本 I/O 子系统采用相同的设计。每个 SPARC M8 处理器通过 I/O 链路连接至一个或多个 I/O 控制器 ASIC。PCIe 基础设施由 I/O 控制器 ASIC 提供，每个 ASIC 提供五个 PCIe 3.0 根联合体。PCIe 3.0 子系统的实现因服务器款型而异，但每款基于 SPARC M8 处理器的服务器均为专用的根联合体提供了多个扩展槽。此功能支持灵活的 I/O 虚拟化，由独立的根域提供冗余、高可用的网络和存储连接。

DIMM 热备

每个 SPARC M8 处理器通过 8 个板上缓冲器 (BoB) ASIC 支持多达 16 条 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。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。半配的内存配置为 8 路交错，每个处理器 16 条 DIMM 的全配配置则为 16 路交错。SPARC M8 处理器还支持 15 路交叉配置，可以动态执行从 16 路到 15 路交错的切换。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（图 2）特性。



图 2. DIMM 热备在不中断服务的情况下提供单个 DIMM 的自动故障切换。

DIMM 热备允许移除 16 条 DIMM 中的一条，并使用其余 15 条 DIMM 的全部容量来重新映射这条 DIMM 的内容。此过程的工作方式是，在初始配置中将每条 DIMM 容量的 1/16 保留为不使用状态。之后可以停用出错的 DIMM，并且其内容可以重新映射至其余 15 条 DIMM。当确定某个 DIMM 有故障时，会自动执行 DIMM 热备，不会造成服务中断。

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，无需关闭系统进行维护。实际的 DIMM 更换过程可以等到同一内存条中出现的后续故障造成使用故障时再进行。

基于 SPARC M8 处理器的服务器在采用全配内存配置（每个处理器 16 条 DIMM）时将启用 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。

冗余服务处理器

所有基于 SPARC M8 处理器的服务器均使用 Oracle ILOM 通过服务处理器 (SP) 进行平台管理。Oracle ILOM 提供了命令行界面 (CLI)、基于 Web 的图形用户界面 (GUI) 以及智能平台管理界面 (IPMI) 功能来帮助带外监视和管理。Oracle Enterprise Manager Ops Center 管理平台通过与 Oracle ILOM 通信来管理和监视基于 SPARC M8 处理器的服务器。所有系统遥测和运行状况诊断结果由 Oracle ILOM 记录并转发至 Oracle Enterprise Manager Ops Center，以便进一步分析和操作。如果认为有必要发布某个服务事件，那么 Oracle Enterprise Manager Ops Center 将与 Oracle Auto Service Request 协同，通告该服务需要采取措施。

SPARC M8-8 服务器旨在支持任务关键型企业负载，与此目标相一致的是，这些服务器配备了可自动故障切换的冗余热插拔 SP。在此配置中，单个 SP 发生故障不会影响监视和管理系统的能力。无论该服务器中配置了一个还是两个物理域，此架构都可以提供冗余和隔离的硬件管理支持。

为了确保可靠性和持续运行，SP 始终被配置为支持主用/备用故障切换的冗余对。SP 对之间的网络可为系统管理信息交换提供便捷。当发生故障切换时，备用 SP 随时可成为主用组件。每个 SP 上有一个或多个服务处理器模块 (SPM)。两种款型的 SPARC M8-8 服务器采用不同的 SPM 配置，每个 PDom 均连接一个冗余式 SP 或 SPP 对：

- » 在只有单个 PDom 的 SPARC M8-8 服务器上，两个 SP 各自只有一个 SPM 来管理系统中的单个 PDom。其中一个 SP 作为主用 SP 来管理平台，另一个作为备用 SP，以便在出现故障时接管主用 SP 的角色。
- » 在有二个 PDom 的 SPARC M8-8 服务器上，互为冗余的 SP 各有两个 SPM，服务器使用两对 SPM 来管理系统中的两个 PDom。第一对 SPM 提供主用 SP 和备用 SP 功能，并管理 CMI0U0 至 CMI0U3。第二对 SPM 管理 CMI0U4 至 CMI0U7（图3）。两对 SPM 协同作业，为所有服务器组件提供单一管理环境。

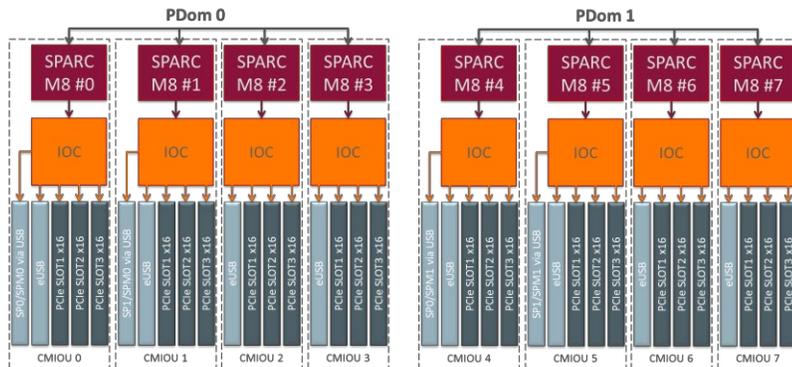


图 3. 具有两个物理域的 SPARC M8-8 服务器包含一对 SP，每个 SP 均有两个 SPM，可为两个 PDom 提供冗余和自动故障切换功能。

Oracle Solaris

在具备相应级别补丁的情况下，基于 SPARC M8 处理器的服务器支持 Oracle Solaris 11 和 Oracle Solaris 10。如果在控制域中使用或者需要裸机运行，则需要 Oracle Solaris 11.3。来宾域中支持 Oracle Solaris 10 1/13 或更高版本。Oracle Solaris 的以下特性可增强基于 SPARC M8 处理器的服务器的可靠性和可用性：

» **Fault Management Architecture.** Fault Management Architecture (FMA) 功能可自动诊断系统中的故障并发起自我修复操作，以帮助避免服务中断，提高可用性。该软件可以在故障发生之前将问题组件配置到系统之外，以此提高可靠性。在发生故障时，该特性会使用 Oracle Solaris 的 Service Management Facility 特性发起自动恢复和应用重启。

从传入错误流中观察到可辨别的模式时，FMA 诊断引擎会立即生成故障诊断。完成诊断之后，FMA 会将故障信息提供给知道如何响应特定故障的代理。尽管其他硬件和软件供应商也提供了类似技术，但这些解决方案都仅限于支持基于软件或基于硬件的故障检测。Oracle 全面的软件体系提供了一个故障检测和管理环境，可将待发硬件故障消息全面集成到操作系统服务中，这样操作系统服务可以根据需要调整硬件资源，从而显著减少停机。

通过在服务器的 SP 上运行 Oracle ILOM，系统支持发起 FMA 事件，提醒管理员存在平台故障。SP 有一个诊断引擎，可以独立于 Oracle Solaris 生成 FMA 事件。SP 会将事件发送到控制域，以便 Oracle Solaris 来宾系统对处理器或内存故障采取措施。Oracle Solaris 会将产生了过多可纠正错误的内核、线程和内存分页置于离线状态。在资源产生致命故障之前将其置于离线状态，有助于确保应用服务更长久地持续正常运行。与其他供应商的同类平台相比，在 SP 上集成 FMA 是基于 SPARC M8 处理器的服务器在可用性方面的一项优势。

» **Service Management Facility.** 借助 Oracle Solaris 的 Service Management Facility，系统管理员可以使用简单的命令行工具来轻松识别、查看和管理系统以及系统提供的服务。Service Management Facility 描述了发生故障的服务在哪些情况下会自动重启。如果这些服务被管理员无意间关闭、由于软件编程错误而终止，或者由于底层硬件问题而中断，那么它们将会自动重启。目前，市场上的其他操作系统要么使用整体式启动脚本，要么使用一系列按顺序执行的小脚本。这些系统无法在脚本之间建立相关性，也不能在纠正问题之后立即重启服务。Oracle Solaris 的预测性自我修复功能搭配 Fault Management Architecture 与 Service Management Facility 故障管理功能，可以将发生故障的处理器线程或内核置于离线状态、停用可疑的内存分页、记录 I/O 组件错误或故障以及处理服务器检测到的任何其他问题。

» **Oracle Solaris ZFS.** Oracle Solaris ZFS 文件系统可为存储提供极高的数据完整性、容量、性能和可管理性。它提供了许多具备高恢复能力的特性，如通过元数据记录确保系统发生故障时保持数据完整性和加快恢复。这种强大的数据完整性和恢复能力让 Oracle Solaris ZFS 从同类文件系统产品中脱颖而出。

Oracle Solaris ZFS 还采用了一些技术手段（如写入时复制和端到端的校验和等）来保持数据一致并消除静态数据损坏，从而显著简化文件系统管理并更加有效地避免管理错误。由于文件系统始终保持一致，因此当系统以不清洁方式关闭时，便不再需要使用 `fsck` 这样耗时的恢复过程。此外，系统会持续读取和检查数据来确保数据准确无误，自动修复从镜像池中检测到的任何错误，以避免数据丢失带来高昂的金钱和时间成本，并防范之前不能检测到的静态数据损坏。

通过实施 RAID-Z 也可以实现纠错 — RAID-Z 使用奇偶校验、条带化和原子操作来帮助重建损坏的数据。Oracle Solaris ZFS 提供了端到端的数据保护和一些数据自我纠错工具。它还提供了一个无需使用第三方卷管理器的简化管理模型。Oracle Solaris ZFS 具备扩展能力，可通过扩展来满足最高数据存储需求。Oracle Solaris ZFS 通过持续读取和检查数据来确保数据准确无误。如果在冗余式存储池（通过镜像、ZFS RAID-Z 或 ZFS RAID-Z2 提供冗余保护）中检测到错误，则 Oracle Solaris ZFS 会自动修复损坏数据。它通过基本的镜像、压缩和集成式卷管理功能维持数据冗余，提高了文件系统的可靠性。

» **Oracle Solaris DTrace.** Oracle Solaris 中的 DTrace 特性是一种动态跟踪框架，用于对生产系统中的系统性问题进行实时故障排除。DTrace 旨在快速识别系统性能问题的根源。它无需重启内核或重新编译（甚至无需重启）应用即可安全、动态地检测正在运行的操作系统内核和应用。这种设计大幅延长了服务的正常运行时间，并且让开发人员和操作人员可以了解动态性能问题。

Oracle Solaris Multipathing

利用 Oracle Solaris 多路径软件，组织可以定义和控制通向 I/O 设备（如存储设备和网络接口）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，服务器首先必须配置有冗余硬件（如冗余网络接口）或者连接到同一双端口存储阵列的两个主机总线适配器。

您可以选择使用的多路径软件有很多种，包括：

- » **Oracle Solaris IP Network Multipathing (IPMP)**。Oracle Solaris IP Network Multipathing 可为 IP 网络接口提供多路径和负载均衡功能。有关更多信息，请参阅相应的[Oracle Solaris 操作系统文档](#)。
- » **Oracle Solaris Datalink Multipathing (DLMP)**。Oracle Solaris Datalink Multipathing 聚合也称为 *中继*，支持管理员将系统上的多个接口配置为单个逻辑单元来增加网络通信的吞吐量。有关更多信息，请参阅相应的[Oracle Solaris 操作系统文档](#)。
- » **Oracle Solaris I/O Multipathing (MPxIO)**。借助 MPxIO，管理员可以为光纤通道设备配置多路径特性来控制所有受支持的光纤通道主机总线适配器。有关更多信息，请参阅相应的[Oracle Solaris 操作系统文档](#)。
- » **Oracle VM Server for SPARC Virtual Disk Multipathing**。借助 Virtual Disk Multipathing，管理员可以在来宾域配置虚拟磁盘通过多条路径来访问其后端存储。有关配置和管理 Oracle VM Server for SPARC Virtual Disk Multipathing 的说明，请参阅[文档库](#)中的相应 *Oracle VM Server for SPARC 管理指南*。
- » **Oracle Sun StorageTek Traffic Manager**。Sun StorageTek Traffic Manager 是完全集成在 Oracle Solaris 内（从 Oracle Solaris 8 版本开始）的架构，支持从 I/O 设备的单个实例通过多个主机控制器接口来访问 I/O 设备。有关更多信息，请参阅相应的[Oracle Solaris 操作系统文档](#)。

Oracle Solaris Cluster

基于 SPARC M8 处理器的服务器可提供高水平的 RAS，而 Oracle Solaris Cluster 可帮助组织提供高度可用的应用服务。为了限制由单点故障造成的中断，任务关键型服务需要在集群物理域或物理服务器中运行，因为这些服务器可以高效、顺畅地接管来自故障节点的服务，尽可能减少数据服务中断。基于 SPARC M8 处理器的服务器在硬件级提供冗余，而 Oracle Solaris Cluster 则为运行 Oracle Solaris 和应用的 SPARC 服务器提供了卓越的高可用性解决方案。

Oracle Solaris Cluster 主要用于在服务器内的区域与逻辑域之间以及服务器与服务器之间实现故障切换。Oracle Solaris Cluster 与 Oracle Solaris 紧密结合，可以几乎毫无延迟地（亚秒级延迟）检测到故障，与同类解决方案相比，能够更快地发出故障通知、执行应用故障切换和完成重新配置。凭借从应用层到存储层的故障切换保护，Oracle Solaris Cluster 可为当今复杂的解决方案体系提供高可用性。该方法大幅加快了在基于 SPARC M8 处理器的服务器上恢复 IT 服务的速度。Oracle Solaris Cluster 能够：

- » 与 Oracle Solaris Predictive Self Healing 框架紧密集成，支持 Oracle Solaris Zones 和逻辑域中由 Service Management Facility 控制的应用。
- » 广泛使用 Oracle 的存储管理和卷管理功能
- » 支持 Oracle Solaris ZFS 作为故障切换文件系统和引导文件系统，允许使用 ZFS 作为单个文件系统类型。
- » 充分利用 Oracle Solaris ZFS 特性，如池存储、内置冗余和数据完整性
- » 与 Oracle Enterprise Manager Ops Center 集成

虚拟化与 Oracle VM Server for SPARC

由于各企业力求将各种负载整合到数量更少、更强大的系统上，同时提高使用率，因此虚拟化技术已变得日益普及和重要。基于 SPARC M8 处理器的服务器专为虚拟化而设计，提供全部三个级别的分区和虚拟化技术：硬件分区、服务器虚拟化和操作系统虚拟化。

- » **物理域 (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。每个PDom运行一个独特的虚拟机管理程序 (Oracle VM Server for SPARC) 副本。SPARC M8-8 服务器有两种款型，一种款型提供一个PDom，另一种提供两个PDom。
- » **逻辑域 (LDom)** 通过 Oracle VM Server for SPARC 创建，用于对服务器或物理域进行虚拟化，以便托管多个分别运行自己的 Oracle Solaris 实例的虚拟机 (VM)。

Oracle Solaris Zones可实现操作系统虚拟化，使单个 Oracle Solaris 实例可以安全地将应用相互隔离，并为各个区域分配相应的系统资源。

每个基于 SPARC M8 处理器的系统均内置虚拟机管理程序固件，因此可通过逻辑域分区来实现软件故障隔离和资源分配，包括 I/O 虚拟化。虚拟机管理程序为清除、纠正和收集错误提供处理器支持。

借助虚拟机管理程序，Oracle VM Server for SPARC 提供可运行独立操作系统实例的完整虚拟机。每个逻辑域彼此完全隔离。单个平台上可以创建的虚拟机的最大数量取决于虚拟机管理程序的能力，而不是系统中安装的物理硬件设备数量。

Oracle VM Server for SPARC 还能在不同的域之间执行实时迁移。顾名思义，实时是指不再需要暂停或停止源域和应用。此功能允许将服务器上的逻辑域迁移至同一服务器、其他基于 SPARC M8 处理器的服务器或前几代服务器的另一个物理域上。

逻辑域还可以托管 Oracle Solaris Zones，获得这两种技术的隔离、灵活性和易管理性特性，或者使用 Oracle Solaris 标记区域来支持 Oracle Solaris 旧版本（图 4）。Oracle Solaris 实现了 Oracle VM Server for SPARC 与 SPARC M8 处理器的深度集成，进而提高了灵活性、隔离了负载处理并提高了充分利用服务器的潜力。

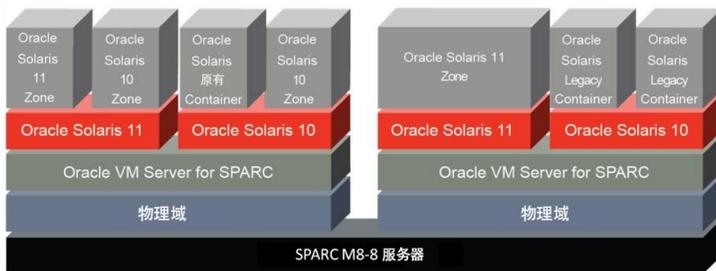


图 4. Oracle VM Server for SPARC 在物理域上运行，支持使用 Oracle Solaris Zones 和 Oracle Solaris 标记区域进一步进行系统虚拟化和隔离。

面向高可用性基础设施的 Oracle 产品和解决方案

以整个硬件和软件应用体系为依托，Oracle 在提高应用可用性方面具有得天独厚的优势，能够提供创新的解决方案。除了设计高可用性系统和软件，Oracle 还致力于推出将软件融入解决方案的解决方案和架构来解决关键应用和基础设施的实际问题。

Oracle Real Application Clusters (Oracle RAC)

Oracle Real Application Clusters (RAC) 是一个共享缓存的集群式数据库架构，突破了传统架构无共享或共享磁盘的局限。Oracle RAC 提供出色的数据库性能、扩展能力和可靠性，同时无需对现有的 Oracle 数据库应用进行修改。目前，已有数千家组织成功部署了 Oracle RAC，使用集群式数据库服务器在云上简便、高效、成功地交付数据库服务。Oracle RAC 能够：

- » **运行所有数据库负载。** Oracle RAC 可以用于联机事务处理、数据仓储应用或混合负载，而用户无需对应用进行修改。Oracle RAC 还可以与 Oracle Multitenant 和 Oracle Active Data Guard 等相辅相成的数据库技术一起部署。
- » **为云端数据库服务提供基础。** Oracle RAC 提供了所有必备的软件组件，可在服务器池上轻松部署 Oracle 数据库，从而充分利用集群技术提供的性能、扩展能力和可用性优势。Oracle RAC 利用 Oracle 数据库的 Oracle Grid Infrastructure 特性为 Oracle RAC 数据库系统奠定基础。Oracle Grid Infrastructure 包括 Oracle Clusterware 和 Oracle Automatic Storage Management (Oracle ASM)，可在高度可用和可伸缩的数据库云环境中高效地共享服务器和存储资源。
- » **支持按需扩展。** Oracle RAC 支持在集群服务器池上透明地部署 Oracle 数据库。这一能力让组织能够轻松地将单个服务器上的孤立 Oracle 数据库重新部署到数据库服务器集群上，从而充分利用集群数据库服务器提供的组合内存容量和处理能力。
- » **提供最高水平的数据库可用性。** Oracle RAC 可以避免因单个数据库服务器造成单点故障的情况，能够为组织提供最高水平的数据库可用性。在集群式服务器环境中，数据库本身在整个服务器池中共享。当服务器池中的任何服务器发生故障时，数据库将继续在正常的服务器上运行。
- » **提供经济高效的资源管理。** Oracle RAC 提供创新技术来管理集群内的负载，同时根据应用的配置和高可用性要求来提供高水平的应用吞吐量。

Oracle WebLogic Server 集群

无论在传统环境还是云环境中，Oracle WebLogic Server 12c 都是一款广受欢迎的应用服务器。借助 Oracle WebLogic Server，组织可在任务关键型云平台上交付下一代应用，通过原生云管理简化运营，并通过现代开发平台和集成工具加快市场投放速度。

一个 Oracle WebLogic Server 集群包含多个 Oracle WebLogic Server 实例，这些实例同时运行并协同作业，以此提高扩展能力和可靠性。对于客户端而言，一个集群就好像单个 Oracle WebLogic Server 实例。构成集群的服务器实例可以运行在同一个系统上，也可以位于不同的系统上。组织可以增加集群的容量，为此，可以在现有系统上向集群添加更多的服务器实例，也可以向集群添加系统以托管增加的服务器实例。集群中的每个服务器实例必须运行相同版本的 Oracle WebLogic Server。Oracle WebLogic Server 集群具有下列优势：

- » **可扩展性。**可以动态增加 Oracle WebLogic Server 集群中部署的应用的容量来满足需求。管理员可以在不中断服务的情况下将服务器实例添加到集群，此时，应用会继续运行而不会影响客户端和最终用户。
- » **高可用性。**在 Oracle WebLogic Server 集群中，如果某服务器实例出现故障，则应用处理可以继续。通过将应用组件部署在集群中的多个服务器实例上，可以实现应用组件的集群化。如果运行某组件的一个服务器实例出现故障，则另一个部署了该组件的服务器实例可继续进行应用处理。

Oracle 优化解决方案和 Oracle 最高可用性架构

Oracle 在 Oracle 优化解决方案和 Oracle 最高可用性架构中融汇使用了 Oracle 的系统、中间件和数据库技术来实现高可用性配置。Oracle 优化解决方案采用基于 SPARC M8 处理器的集群服务器，是经过精心设计、全面测试、文档详实且专为实现最佳性能和可用性进行了优化的架构。这些解决方案基于独特的匹配组件，包括 Oracle 的集成系统、服务器和存储设备，以及操作系统、虚拟化、数据库、中间件和企业应用。Oracle 针对各种领域提供了优化的解决方案，这些领域有：

- » **企业应用。**借助集成的可用性和安全功能，Oracle 针对企业应用的优化解决方案可为整个应用环境提供优化的性能，这些环境包括 Oracle JD Edwards EnterpriseOne 应用、Oracle E-Business Suite、Oracle FLEXCUBE、Oracle PeopleSoft 和 Siebel CRM 应用以及 SAP。
- » **中间件。**Oracle 针对中间件的优化解决方案可帮助组织轻松管理有效、安全和高性能的应用基础设施以及包含 Oracle WebLogic Server 在内的解决方案。
- » **数据管理。**Oracle 针对生命周期内容管理的优化解决方案、Oracle 针对 Oracle 数据库的优化解决方案以及 Oracle 针对企业数据库云的优化解决方案提供各种数据管理功能，运用成熟的优秀实践来控制成本、显著提高性能，从而满足不断增长的结构化和非结构化存储需求。
- » **核心基础设施。**Oracle 针对核心基础设施的优化解决方案包括经过完全测试的蓝图，可为打造更加经济高效、敏捷的数据中心奠定基础。这些解决方案包括 Oracle 针对备份和恢复的优化解决方案、Oracle 针对灾难恢复的优化解决方案、Oracle 针对企业云基础设施的优化解决方案以及 Oracle 针对分层存储基础设施的优化解决方案。

作为对 Oracle 优化解决方案的补充，Oracle 最高可用性架构还利用了 Oracle 的优秀实践蓝图、专家建议以及众多客户在成熟的 Oracle 高可用性技术方面的经验。每个 Oracle 最高可用性架构的目标都是以最低的成本和复杂性为用户提供最理想的高可用性。在本文截稿时，Oracle 最高可用性架构的优秀实践蓝图涵盖下列主题：

- » Oracle 数据库
- » Oracle 融合中间件
- » Oracle 应用无极限
- » Oracle 融合管理软件
- » Oracle Enterprise Manager
- » Oracle VM
- » Oracle 合作伙伴
- » Oracle 咨询和支持

总结

Oracle 基于 SPARC M8 处理器的全系列服务器是在全面了解企业在应用高可用性的严苛要求的基础上设计的。这些服务器与可靠的 Oracle Solaris 操作系统、创新的 SPARC M8 处理器特性、关键冗余组件以及 DIMM 热备等特性一道工作，通常在错误中断关键的应用或服务之前就能够检测出错误并从错误中恢复。Oracle 的分层虚拟化技术有助于隔离不同的应用和系统映像，而 Oracle WebLogic Server 和 Oracle RAC 等关键解决方案能够通过多个虚拟或物理系统来提供扩展能力和高可用性。此外，Oracle 优化解决方案和 Oracle 最高可用性架构提供经过广泛测试的配置，这些优秀实践将整个 Oracle 硬件和软件技术体系融入可靠的解决方案中，将帮助组织实现其最为严苛的可用性目标和 SLA。

更多信息

有关更多信息，请访问表 2 中列出的资源。

表 2 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 最高可用性架构 | oracle.com/technetwork/database/features/availability/maa-096107.html |
| Oracle Solaris Cluster | oracle.com/us/products/servers-storage/solaris/cluster/overview/index.html |
| Oracle Real Application Clusters (Oracle RAC) | oracle.com/technetwork/database/options/clustering/overview/index.html |
| 虚拟化 | https://www.oracle.com/technetwork/cn/topics/virtualization/whatsnew/index.html |
| Oracle Enterprise Manager Ops Center | https://www.oracle.com/technetwork/cn/oem/ops-center/index.html |
| Oracle 帮助中心（提供产品文档） | docs.oracle.com |

白皮书和技术文章**URL**

| | |
|--|---|
| “基于 SPARC T8 和 SPARC M8 处理器的 Oracle 服务器架构” | https://community.oracle.com/docs/DOC-1017902 |
| “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 |
| “数据库整合高可用性优秀实践” | oracle.com/technetwork/database/availability/maa-consolidation-2186395.pdf |
| “建立、配置和使用 Oracle WebLogic Server 集群” | https://www.oracle.com/technetwork/cn/articles/soa/vasiliev-wls-cluster-1867166.html |
| “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 |



ORACLE Redwood Shores, CA 94065, USA

Oracle Corporation, World Headquarters
500 Oracle Parkway

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。0615

Oracle SPARC T8 和 SPARC M8 服务器的可靠性、可用性和可维护性
2017 年 10 月