

# Oracle Exadata

## 数据库云平台

### X8M-8

Oracle Exadata 数据库云平台通过集成设计为 Oracle 数据库带来了卓越的性能、极佳的成本效益和极高的可用性。Exadata 采用支持云的现代化架构，内含可横向扩展的高性能数据库服务器、可横向扩展的智能存储服务器（配有先进的 PCI 闪存）、使用持久性内存的先进存储缓存以及连接所有服务器和存储的云级 RDMA over Converged Ethernet (RoCE) 内部结构。Exadata 采用独有的算法和协议，在存储、计算和网络中实现了数据库智能，因此，与其他平台相比，能以更低的成本提供更高的性能和容量。它适用于所有类型的现代数据库负载，包括联机事务处理 (OLTP)、数据仓库 (DW)、内存分析、物联网 (IoT) 以及高效整合的混合负载。Exadata 数据库云平台部署简单、快速，能为您极其重要的数据库提供全面的支持和保护。Exadata 既可以作为私有数据库云的理想基础平台购置并部署在内部环境中，也可以采用订阅模式购买并部署在 Oracle 公有云或公有云本地化环境中而由 Oracle 负责全部基础设施管理工作。



Oracle Exadata 数据库  
云平台 X8M-8

Exadata 数据库云平台 X8M-8 是一个高端系统，和 X8M-2 一样具有超强的性能、存储能力和网络结构，区别在于它配备了大规模 8 槽 SMP 服务器，而不是 X8M-2 的 2 槽服务器。X8M-8 每台 8 槽服务器总共有 192 个处理器内核和 3 至 6 TB 的 DRAM，因此能够很好地支持高端联机事务处理 (OLTP)、多机架数据仓库和内存中数据库负载，是需要能够纵向扩展且具有大容量内存的大型服务器的应用的理想选择。

### 集成设计，支持快速可靠部署

Oracle Exadata 数据库云平台是运行 Oracle 数据库的成本效益极佳、性能超强的平台。Exadata 易于部署，甚至适用于要求最严苛的任务关键型系统，其数据库服务器、存储服务器和网络均由 Oracle 专家预先配置、预先调优和预先测试。该系统经过了广泛的端到端测试和验证，可确保所有组件（包括数据库软件、操作系统、固件、驱动器等）均能无缝协同工作，消除性能瓶颈或单点故障。

所有 Oracle Exadata 数据库云平台均采用完全一致的配置方式，因此客户将受益于从数以千计的其他客户部署中不断积累的丰富经验。此外，客户所使用的机器与 Oracle 支持部门用于发现和解决问题的机器、Oracle 开发部门用于开发和测试 Oracle 数据库的机器以及 Oracle 用于运行其公有云服务的机器也是完全一样的。**Exadata 为运行 Oracle 数据库进行了完全地测试和调优。**

如今使用 Oracle 数据库的任何应用都可无缝迁移至 Oracle Exadata 数据库云平台，无需对应用进行任何更改。同样，客户还可以轻松将任何 Oracle 数据库从 Exadata 中迁出，完全不必担心“供应商依赖”。

### 通过弹性配置提供超强的系统可扩展性和扩容能力

Exadata 数据库云平台对数据库服务器和存储服务器均采用了一种可横向扩展的架构。随着负载增加，客户可以添加数据库 CPU、存储和网络，从而扩展系统，同时不会造成瓶颈。**这种架构可从小规模配置扩展至极大规模的配置，可适应任何规模的负载。**

全面的高带宽、低延迟、100 Gb/秒的 **RDMA over Converged Ethernet (RoCE) 网络结构**将 Oracle Exadata 数据库云平台中的所有组件连接在一起。与常规通信协议相比，专有数据库网络协议延迟更低、带宽更高，可**更加快速地响应 OLTP 操作并提高分析负载的吞吐量**。与 Oracle Exadata 数据库云平台的外部连接则通过标准 10 Gb/秒或 25Gb/秒以太网来提供。

Exadata X8M-8 是 Oracle Exadata 系列中的高端产品，特别适合需要在少量大型计算节点上运行的大型 OLTP 数据库、需要超大内存容量的内存中数据库负载、需要尽量减少数据库节点数量的大型多机架配置，以及需要高 I/O 吞吐量和超大缓冲区缓存的数据库。

### 主要特性

- 每机架多达 576 个 CPU 内核和 18 TB 内存用于数据库处理
- 每机架多达 576 个 CPU 内核专用于在存储中进行 SQL 处理
- 每个机架高达 21 TB 的持久性内存加速
- 100 Gb/秒 RoCE 网络
- 完全冗余，确保高可用性
- 每机架 2 到 3 台数据库服务器
- 每机架 3 到 14 台存储服务器
- 每机架高达 720 TB 闪存容量（物理）
- 每机架高达 2.3 PB 磁盘容量（物理）

### 主要优势

- 该系统经过预先配置、预先测试并且针对所有数据库应用进行了优化
- 每个全机架未压缩 SQL I/O 带宽高达 560 GB/秒
- 每个全机架每秒可执行多达 900 万次 8K 数据库读取 I/O 操作和 510 万次 8K 闪存写入 I/O 操作
- 轻松添加计算服务器或存储服务器，可满足任意规模应用的需求
- 可连接多个 Exadata 数据库云平台 X8M-8 机架或 Exadata X8M 存储扩展机架进行扩展。只需增加 RoCE 电缆和内部交换机即可连接多达 18 个机架。通过外部 RoCE 交换机创建更大规模的配置

**Exadata 数据库云平台是最通用的数据库平台。** Exadata 数据库云平台 X8M-8 采用了强大的数据库服务器，每台服务器均配有 8 个 24 核 x86 处理器和 3 TB 内存（可扩展至 6 TB）。Exadata 还采用了可横向扩展的智能存储服务器，这些存储服务器有两种配置 — 大容量 (HC) 或极速闪存 (EF)。HC 存储服务器包含 4 个 NVMe PCI 闪存卡，每个闪存卡配有 6.4 TB（原始）Exadata 智能闪存缓存和 12 个 14 TB 7200 转 RPM 磁盘。EF 存储服务器采用全闪存配置，包含 8 个 NVMe PCI 闪存驱动器，每个闪存驱动器配有 6.4 TB（原始）存储容量。Exadata X8M HC 和 EF 存储现已配备持久性内存，可进一步提高容量和性能。HC 和 EF 服务器引入 12 个 128 GB Intel® Optane™ DC 持久性内存模块，作为 DRAM 与闪存之间的新层。Exadata 结合持久性内存与**绕过网络和 I/O 体系的创新性 RDMA 算法**，可消除成本极高的 CPU 中断和上下文切换，最终可将延迟时间减少 10 倍，即从 200μs 缩短至不到 19μs。

Exadata 数据库云平台的最低配置包含两台数据库服务器和三台存储服务器，用户可以将其扩展到弹性配置，在同一个机架内添加更多的数据库服务器和/或存储服务器。利用弹性配置，用户可以灵活、高效地满足任何规模的业务需求。

除了在机架内部扩展以外，还可以使用集成式 RoCE 网络结构连接多个 Exadata X8M 机架，构成更大规模的配置。例如，由四个机架构成的系统在能力上相当于单一机架的四倍：它提供四倍的 I/O 吞吐量、四倍的存储容量和四倍的处理能力。对于多机架连接，用户可将其配置为单个系统，也可以通过逻辑分区整合多个数据库。横向扩展非常容易，因为 Oracle Real Application Clusters (RAC) 可以动态增加处理能力，自动存储管理 (ASM) 则可动态增加存储容量。

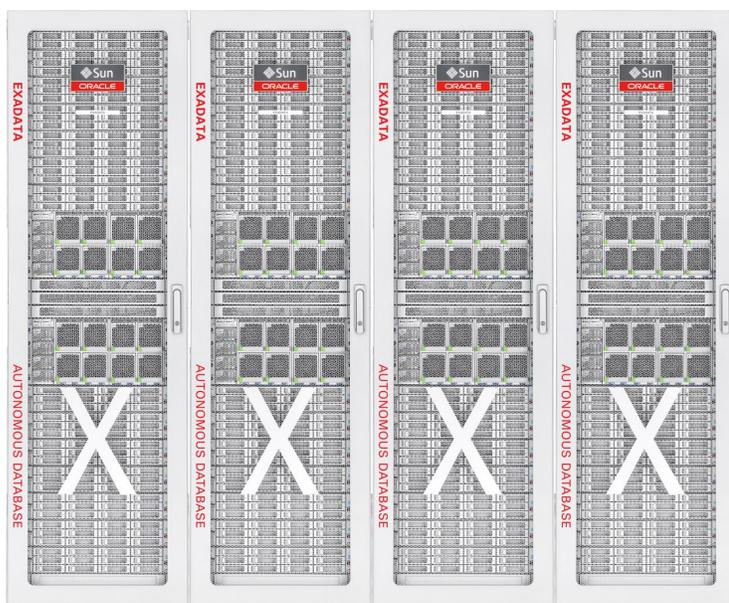


图 1. 灵活地横向扩展为多机架 Exadata

#### 相关产品

- Oracle 数据库 Exadata 云服务
- Oracle 数据库 Exadata 公有云一体机
- Oracle Exadata 数据库云平台 X8M-8
- Oracle Exadata 存储扩展机架 X8M
- Oracle Exadata 存储服务器 X8M-2 + 网络结构
- Oracle Exadata 数据库服务器 X8M-8 + 网络结构
- Oracle Database 11g、12c、18c 和 19c
- RAC
- 分区
- 多租户
- Database In-Memory
- 高级压缩
- 高级安全
- 活动数据卫士 (ADG)
- GoldenGate
- RAT
- OLAP
- 高级分析
- BI
- Enterprise Manager
- Oracle Linux
- Oracle Virtual Machine

#### 相关服务

Oracle 提供下列服务：

- 高级客户服务
- Oracle 标准系统支持服务
- Oracle 白金服务
- Oracle 咨询服务
- Oracle 大学课程

当需要大存储容量时，可以利用 **Oracle Exadata X8M 存储扩展机架**进行扩展。客户可以使用 Exadata 存储扩展机架来为任何 Oracle Exadata 数据库云平台增加存储容量和带宽。该存储扩展机架专为拥有海量数据的数据库部署而设计，这些数据包括历史或归档数据、备份、文档、图像、XML、JSON 和 LOB。该存储扩展机架使用集成式 RoCE 网络结构连接至 Exadata 数据库云平台，无需设置 LUN 或挂载点，因此只需使用几个简单的命令即可完成配置。Oracle Exadata 存储扩展机架的初始配置包含 4 台存储服务器，用户可添加更多存储服务器来对其进行扩展。

### EXADATA X8M 的新特性：基于 RDMA 的突破性网络结构

Exadata X8M 版本提供下一代超高速云级网络结构，即 RDMA over Converged Ethernet (RoCE)。RDMA（远程直接内存访问）可允许一台计算机在不占用操作系统或 CPU 资源的情况下直接从另一台计算机访问数据，从而实现高带宽和低延迟。网卡直接读取/写入内存，无需额外的复制或缓冲过程，而且延迟非常低。RDMA 是 Exadata 高性能架构中的一个重要组成部分，并且在过去十年中得到了持续调优和增强，可支持多种 Exadata 专有技术，例如 Exafusion Direct-to-Wire 协议和 Smart Fusion Block Transfer。由于 RoCE API 基础架构与 InfiniBand 的基础架构相同，因此**可以在 RoCE 上使用所有现有的 Exadata 性能特性**。

Exadata X8M 版本采用 100 Gb/秒的 RoCE 网络结构，这使其成为全球最快的数据库云平台。在采用新的**共享持久性内存加速器**的 Exadata X8M-2 上运行的实际数据库负载已经超越了 Exadata X8 之前创下的 670 万读取 IOPS 的基准测试记录，达到超过 1600 万 OLTP 读取 IOPS (8K IO)。

### EXADATA X8M 的新特性：共享持久性内存加速

作为 Exadata X8M 中的新特性，存储服务器在闪存缓存前端引入了持久性内存 (PMEM) 数据和提交加速器，可将访问远程存储数据的延迟降低若干个数量级。持久性内存是一项全新的芯片技术，可在 DRAM 与闪存之间添加一个极具性价比的大容量存储层。由于持久性内存物理上存在于存储服务器的内存总线上，因此可直接以内存速度执行读取操作，其速度相比闪存大幅提升。与 DRAM 不同，写入操作是持久性的，在断电后仍然会保留。通过利用 RDMA 远程访问持久性内存，**Exadata 智能 PMEM 缓存**可绕开网络、I/O 软件、中断和上下文切换，与之前的 Exadata 相比，其延迟可降低 10 倍以上，减少至不到 19 微秒。Exadata 智能系统软件还可以确保跨存储服务器建立数据镜像，从而提供额外的容错能力。Exadata 在 Oracle 数据库与 Exadata 存储之间实现了独有的端到端集成，可自动识别要存储的热度最高的数据块，同时确保数据库、持久性内存和闪存缓存不会多次保存同一数据块，从而提高整个存储层的效率。将持久性内存添加到存储层意味着此新缓存层的聚合性能可以由任何服务器上的任何数据库动态使用。与通用存储架构相比，这是一项显著的优势，可以避免跨服务器共享。

此外，新的智能 Exadata 系统软件特性还可以提高日志写入性能。日志写入延迟对于 OLTP 性能至关重要，更快的日志写入意味着更快的提交速度。相反，日志写入速度的任何下降都可能会导致数据库中断。作为 Exadata X8M 中的一项独有特性，**Exadata 智能 PMEM 日志**可以自动让数据库向持久性内存发起单向 RDMA 日志写入。在 RDMA 和持久性内存技术的支持下，无需确认即可执行日志写入操作，而且智能软件会将写入操作分布在多个服务器上，从而确保弹性。这将日志写入性能提高了 8 倍。

**“我建议使用 Exadata 平台，因为它能提供卓越的性能、可靠性和便捷的支持。这三个方面平日里对我非常重要。”**

**Richard Ewald**  
Sprint 高级技术架构师

**“借助 Oracle Exadata，我们无需进行重大修改就无缝、快速地迁移了我们基于云的集成式业务应用，并且降低了成本。利用 Oracle 产品，我们的批处理速度提高了 241 倍，总体系统性能提高了 3 倍，从而提高了我们日常关键任务的运营效率，改善了客户服务。”**

**Kyoji Kato**  
Daiwa House  
执行官兼总经理

这个新层还实现了自动化安全和管理流程。持久性内存会在安装时自动完成配置，而不需要任何用户交互。硬件监视也是现成配置好的。持久性内存只能由使用数据库访问控制的数据库访问，从而确保数据的端到端安全性。在 Exadata X8M 中部署持久性内存是一个非常简单的透明过程。

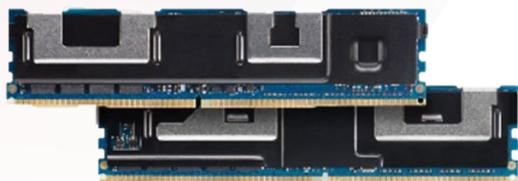


图 2. Intel® Optane™ DC 持久性内存模块

### 极速闪存存储服务器：破纪录的 I/O 性能

Exadata **极速闪存 (EF) 存储服务器**最初在 Exadata X5 中引入，是数据库优化的全闪存 Exadata 数据库云平台的基础。每台 EF 存储服务器包含 8 个 6.4 TB Flash Accelerator F640v2 NVMe PCI 闪存驱动器，可提供 51.2 TB 的原始闪存容量。这种先进的闪存可提高速度和电源效率，对于大多数数据库负载，其预期使用寿命可达到 8 年甚至更久。此外，Exadata 将这些闪存设备直接置于高速 PCI 总线上（而不是在低速磁盘控制器和管理器的后方），可提供超强的性能。Exadata X8M 在闪存前端添加了共享持久性内存加速层，12 个 128 GB Intel® Optane™ DC 持久性内存模块，可进一步提高性能。

Exadata X8M 结合使用了横向扩展存储、RDMA over Converged Ethernet 网络、数据库分流、持久性内存加速器和 PCI 闪存，可提供极高的内存和闪存性能。单一机架配置的数据库云平台 X8M-8 配有 3 台数据库服务器和 14 台极速闪存存储服务器，使用新的持久性内存，可实现高达**每秒 900 万次的随机 8K 数据库读取 I/O 操作和 510 万次的随机 8K 闪存写入 I/O 操作 (IOPS)**。

这些数据是在单机架 Exadata 系统中以标准 8K 数据库 I/O 规模运行 SQL 负载时测定的真实的端到端性能结果，而一般存储供应商给出的性能数据则通常基于较小的 I/O 规模和低级 IO 工具，因此比运行实际 SQL 负载时的性能虚高很多倍。Exadata 处理实际数据库负载的性能比传统存储阵列架构高若干数量级，也远非当今的全闪存存储阵列（其架构瓶颈会限制闪存吞吐量）可比。



图 3. 闪存加速器 PCIe 卡

**“Oracle Exadata 数据库云平台正在帮助我们实现业务转型。我们的 SAP 环境是全球巨型环境之一，现在，它能以更高的稳定性支持高达以往两倍的吞吐量。”**

**Milt Simonds**  
企业平台交付总监  
AmerisourceBergen 公司

## 大容量存储服务器：分层式磁盘闪存和持久性内存以磁盘的成本提供共享内存性能

另一种 Exadata 存储选件是**大容量 (HC) 存储服务器**。该服务器包含 12 个 14 TB SAS 磁盘驱动器，磁盘原始总容量为 168 TB。它还包含 4 个闪存加速器 F640v2 NVMe PCIe 卡，可提供 25.6 TB 的闪存物理总容量。Exadata X8M 在闪存前端添加了共享持久性内存加速层，12 个 128 GB Intel® Optane™ DC 持久性内存模块，可进一步提高性能。使用智能软件 **Exadata Smart PMEM Cache** 进行部署时，只有热度最高的数据库块会自动缓存到此新层中。可通过 RDMA 直接从数据库访问，从而以极低的延迟提供极高的 I/O 速度。

HC 存储服务器中的闪存可直接用作闪存盘，不过一般都是将其配置为 PMEM 缓存背后的磁盘存储前端的闪存缓存（**Exadata 智能闪存缓存**）来提供最佳性能。Exadata 智能闪存缓存与 PMEM 缓存同步运行，可自动缓存频繁访问的数据，而将不常访问的数据保留在磁盘中。这既提供了闪存的高 I/O 速率和快速响应速度，还兼具磁盘的大容量和低成本优势。Exadata 以独有方式理解数据库负载，**知道何时不用缓存那些会对整体性能产生负面影响的数据**。例如，如果由备份或大表扫描引起的大规模写入 I/O 可能会中断更高优先级的 OLTP 或扫描操作，则这些大规模 I/O 将绕开闪存缓存并直接进入磁盘。否则，Exadata 系统软件会缓存这些 I/O，以便于利用额外的备用闪存容量和 I/O 带宽来优化性能。除了自动缓存以外，管理员还可以选择提供 SQL 指令来确保将特定的表、索引或分区优先保存于闪存缓存中。

在实际的数据库负载处理中，Exadata 智能闪存缓存的命中率通常可以超过 95% 甚至达到 99%，产生的有效闪存容量可达到物理闪存的许多倍。例如，一个配有 8 台数据库服务器和 14 台大容量存储服务器的全机架系统提供的有效闪存容量接近于 700 TB 的可用磁盘容量。

此外，Exadata 智能闪存缓存还可以使用 Exadata 写回闪存缓存技术来缓存数据库块写入。写入缓存可消除大规模 OLTP 和批处理负载的磁盘瓶颈。一个配有 3 台数据库服务器和 11 台大容量存储服务器的全机架 Exadata 数据库云平台 X8M-8 的闪存写入容量**每秒超过 510 万次 8K 闪存写入 I/O 操作**。Exadata 写入缓存具备透明性、持久性和完全冗余性，其性能相当于数十个由数千磁盘驱动器组成的企业磁盘阵列。

与其他基于闪存的解决方案相比，Exadata 的另一大优势是支持在 RAM、持久性内存和磁盘之间自动进行数据分层。许多存储供应商已经开发了全闪存阵列，可实现比传统阵列更高的性能。这些全闪存式阵列在性能上有提升，而 Exadata 在磁盘与闪存之间实现的智能数据分层则具有强大的成本优势，这是因为闪存相对比较昂贵，其大小限制了可从中受益的总数据大小。此外，这些闪存阵列也无法从 Exadata 独有的数据库感知的存储优化技术中获益。某些闪存阵列提供的通用重复数据删除技术在 Virtual Desktop Infrastructure 环境中效果出众，但对数据库不起作用。

Exadata 不仅在容量上比一般全闪存式阵列大得多，其性能也更加出色。Exadata 采用经过优化的集成式架构，基于 100 Gb/秒的 RDMA over Converged Ethernet 实现完全横向扩展的网络结构，配备极速 PCI 闪存，将数据密集型操作分流至存储并且所有算法都专门针对数据库进行了优化，其吞吐量是全闪存式存储阵列所无法企及的。

**“我们选择 Oracle Exadata 是因为它提供了一个全面的解决方案……我们生成日财务报表的速度提高了 4 倍，生成流动性风险报告的速度提高了 7 倍，从而可以始终如一地满足我们的服务级别协议要求。另外，我们还改善了信贷风险管理，减少了我们数据中心的占地面积。”**

**Vaibhav Samant**  
HDFC Bank Ltd. IT  
高级副总裁

**“任何报表的生成都不会超过 10 分钟。以前一般要花 3-4 个小时，而现在 3 分钟内即可完成。这听起来不可思议，但事实确实如此。”**

**财务用户**  
Turkcell

## 容量扩展存储服务器：针对不常用数据的低成本 EXADATA 存储

Exadata X8 新引入了第三种 Exadata 存储选项 — **Extended (XT) 存储服务器**。每台 Exadata XT 存储服务器包含 12 个 14 TB SAS 磁盘驱动器，原始总磁盘容量为 168 TB。为了降低成本，该存储服务器不包含闪存，并且存储软件是可选的。对于 X8M 版产品，XT 扩展存储服务器也将受益于新增的 100Gb/秒网络。

此存储选项可将 Exadata 的运营和管理优势扩展到很少访问但又必须保持联机的数据。Exadata 扩展 (XT) 存储服务器具备以下特性：

- 高效 — XT 存储服务器和 HC 存储服务器一样提供很高的容量，并且具备混合列压缩功能
- 简单 — XT 存储服务器既能为 Exadata 增添容量，同时又能保持对应用和 SQL 透明，并能保持当前的工作模式
- 安全 — XT 服务器支持客户将用于联机数据的 Exadata 安全模式和加密功能扩展到不常用的数据
- 快速、可扩展 — 与其他针对不常用数据的存储解决方案不同，XT 存储服务器已经集成到 Exadata 结构中，因而支持快速访问，并且易于横向扩展
- 兼容 — XT 服务器只是另一种款式的 Exadata 存储服务器，客户可以将 XT 存储服务器添加到任何 Exadata 机架中

借助 Exadata 扩展 (XT) 存储服务器，企业可以一直使用值得信赖的并且持续经过验证的 Exadata 解决方案来满足其长期数据保留合规性需求，无需在多个平台上管理信息生命周期的运营风险和成本。

## 利用智能系统软件加快数据库处理

随着数据量持续快速增长，传统的存储阵列难以按所需的速率将数据从磁盘和闪存快速传输至数据库服务器以充分发挥 CPU 的处理能力。搭载数十个 CPU 内核的现代服务器每秒可以处理数十至数百 GB 的数据量，这种速度远远超出了传统存储阵列通过其存储控制器和存储网络可以达到的水平。

**Exadata 系统软件**在 Exadata 存储服务器上实施了独有高效、针对数据库进行了优化的存储基础架构，从而消除了传统存储阵列方案的所有瓶颈问题，实现了卓越的性能。每台 Exadata 存储服务器具有两个 16 核 x86 处理器，用于分流数据库处理工作。一个 Exadata 数据库云平台机架在其存储服务器中总共可以包含多达 576 个处理器内核，可用于数据库服务器分流。存储服务器中的 CPU 不是取代数据库 CPU，而是用于加速数据库密集型负载，就像显卡加速图像密集型负载一样。

Exadata 系统软件有许多独有特性，其中一个就是**智能扫描技术**，用于**将数据密集型 SQL 操作从数据库服务器直接分流至存储服务器**。通过将 SQL 处理推送到存储服务器，从磁盘和闪存中读取数据时就可以立即在所有存储服务器上并行执行数据筛选和处理。而且，**只将与查询直接相关的行和列发送至数据库服务器**。

例如，如果某查询要查找三月份订单额超过 1000 美元的客户，那么 Exadata 系统的处理方式为：将表扫描分流至 Exadata 存储，过滤掉金额小于 1000 美元的所有订单，过滤掉三月份以外的订单，仅提取相关的客户信息。

“现在，Oracle Exadata 数据库云平台是我们整个体系的核心和灵魂。使用 Oracle Exadata，我们能够**将查询时间从数天缩短为几分钟，而过去花费数分钟的查询能够缩短为几秒钟。**”

Dunnhumby  
企业架构师  
Chris Wones

“[利用 Exadata] 我们能更快地处理每天高达 650 亿次的数据库计费事务，同时还为客户问询提供实时信息，提高了客户满意度，降低了成本。”

Jin Hyung Lee  
网络工程部 ICT 团队经理  
SK Telecom

这将传输至数据库服务器的数据量减少了数个量级。智能扫描可大幅加快查询执行速度，消除了瓶颈，并显著降低数据库服务器的 CPU 开销。

**存储索引**是 Oracle Exadata 系统软件的另一项**独特**、强大的功能，该功能有助于避免不必要的 I/O 操作，从而提高总体性能。存储索引在存储服务器的内存中维护，对于在该存储服务器上的一个存储区中包含的表列，它跟踪其汇总信息。如果查询指定了 WHERE 子句，则 Exadata 系统软件将使用布隆过滤器查看存储索引，以便确定存储服务器的某个磁盘区中是否可能存在具有指定列值的行。如果不存在该列值，则可避免该查询在这一区域中的扫描 I/O。存储索引使许多 SQL 操作的运行速度显著提高，因为大量 I/O 操作自动由几个内存中查找操作取代。

除了 Exadata 系统软件的内在功能之外，Oracle 数据库软件、Exadata 系统软件和 Exadata 基础设施三者强强结合，实现了更多独有功能，从而能够为 **OLTP** 负载提供卓越的性能水平。例如，**Exafusion Direct-to-Wire 协议**采用一种**独特**的方式，支持数据库进程使用**远程直接内存访问 (RDMA)**通过极速 RoCE 网络直接读取和发送 Oracle 真正应用集群 (Oracle RAC) 消息，从而避免了操作系统内核和网络软件开销。这提高了 Oracle Exadata 数据库云平台上 Oracle RAC OLTP 配置的响应速度和可扩展性，尤其是在处理高争用更新负载时。

在一些 OLTP 负载中，超过半数远程读取的目的是通过 UNDO 块来满足读取一致性。Exadata 独特地利用超高速 **RDMA** 从其他数据库实例**读取 UNDO 块**，从而进一步提高 OLTP 性能。

**独特的智能融合块传输**功能消除了重做日志写入延迟的影响，因此进一步提高了 RAC OLTP 配置的性能 — 特别是需要发送节点与接收节点之间传输热块时。当发送节点向重做日志发出 I/O 指令时，系统会立即传输该块，而无需等待重做日志 I/O 操作完成。Oracle 内部测试表明，对于通信密集型负载，智能块传输提高了吞吐量（大约 40%），缩短了响应时间（大约 33%）。

为了进一步加速 OLTP 负载，Exadata X8M 大容量和极速内存服务器添加了持久性内存提交加速器，让数据库能够通过 RDMA 直接对多个存储服务器的持久性内存缓冲区并行执行日志写入操作。该加速器部署为 **Exadata 智能 PMEM 日志**，可避免与操作系统的交互以及通过标准 I/O 路径传输数据的开销。这种单向事务可将日志写入性能提高多达 8 倍。

此外，Exadata **独特地**利用机器学习实现了 **Oracle Database 19c 的自动索引功能**。自动索引持续分析执行中的 SQL，并创建新的索引来提高性能。当底层数据模型或使用模式发生变化时，自动索引**不断学习并对数据库进行调优**。

Exadata 还**独特地**实现了**实时统计收集**，可在 DML 操作插入、更新或删除数据时实时收集统计信息。利用实时统计，SQL 优化器可在数据分布发生变化时动态调整执行计划。

### 通过压缩优化存储使用和 I/O

Exadata 存储服务器提供了一种独特的压缩功能，称为**混合列压缩 (HCC)**，可**显著减少大型数据库的存储占用**。

**“Exadata 是我们预订引擎的核心，没有 Exadata，我们就无法进行业务经营，无法售票。”**

Westjet  
首席技术专家  
James Callaghan

**“Exadata 为我们的数据仓库带来了 20 倍的压缩，这实在令人惊叹。”**

Morrisons,  
Plc. 商务智能和数据仓库主管  
Jonathan Walsh

混合列压缩技术是在数据库表中组织数据的一种创新性的方法，它结合使用行方法与列方法来存储数据。这种混合方法既可获得列存储的压缩优势，又可避免纯列格式的性能劣势。

利用混合列压缩，Exadata 可对 Oracle 数据库实现极高水平的数据压缩并减少 I/O，从而大幅降低成本并显著提高性能，这对于分析负载尤为有效。存储节省取决于具体的数据类型，通常在 5 至 20 倍之间。平均存储节省达到了 10 倍，处于行业领先水平。对于传统系统而言，较高的数据压缩会以性能的降低为代价，因为这会增加 CPU 的解压缩负担。而对于 Exadata 数据库云平台，它能够将解压缩分流至 Exadata 存储的处理器中，而且因实现了高度压缩而减少了 I/O，因此在多数情况下，使用混合列压缩可加快分析负载的运行速度。

混合列压缩支持两种模式。**仓库压缩**模式适用于读取密集型负载，可大幅节省存储，同时增强分析性能。**归档压缩**模式可提供极高的压缩率，适用于极少访问但仍须保持联机的数据。此外，这类数据现在可以无缝存储在 XT 存储服务器中，以便进一步降低成本。

在 OLTP 系统中，混合列压缩可用于压缩较旧且不太活跃的数据，而较新、活跃度较高以及经常需要更新的数据则可以采用高级行压缩技术进行压缩。Oracle Database 18c 及更高版本能够联机更改各个表分区所采用的压缩类型（即使表上使用了全局索引），以确保随着数据的老化和活跃度降低在不同压缩类型之间进行无缝分层。

针对数据分析，Exadata 智能闪存缓存实施了一种用于加快报告生成和分析查询的特殊算法，即 **Exadata 列闪存缓存**。列闪存缓存功能在 Exadata 闪存中实现了一种双格式架构，当频繁扫描的混合列压缩数据加载到闪存缓存时，它自动将这些数据转换为纯列格式。对闪存中的纯列数据的智能扫描速度更快，因为扫描只读取选定的列，从而减少了闪存 I/O 和存储服务器 CPU 占用。这在保持 OLTP 式单行查找的卓越性能的同时加快了报告和分析查询的速度。

### 适用于分析和混合负载的容错、超快的 Database In-Memory 云平台

Exadata 是运行 Oracle Database In-Memory 的理想平台。在 Exadata 上运行 Oracle Database In-Memory 不需要将所有数据都存储在内存中。数据可以存储于多个存储层中：最热数据存储在内存中，确保超高查询性能；活跃数据存储在闪存中，确保超高 I/O 吞吐量；活跃度较低或较早的数据则存储在成本极低的磁盘上。**单个查询可以完全透明地访问所有三层中的数据：内存、闪存和磁盘。**因此，与同类产品相比，Exadata 运行速度更快、容量更大且成本较低。

此外，Exadata 独创性地在**闪存缓存中实现了内存中列格式**。此特性扩展了 Exadata 列闪存缓存，可以在数据加载至闪存缓存中时自动将其转换为内存中列格式。智能扫描还可利用超高速单指令多数据 (SIMD) 向量指令，从而通过单一指令来处理多个列值。智能扫描结果将以 Oracle Database In-Memory 的格式传回数据库服务器，从而进一步降低了数据库服务器的 CPU 负载。这会将内存中列存储大小从数据库服务器中的 DRAM 容量无缝扩展至存储服务器中的闪存容量。Exadata X8M-8 全机架 HC 具有 360 TB 闪存缓存，可支持超大规模的内存中负载。不使用 Oracle Database In-Memory 的数据库仍然可以从 Exadata 列闪存缓存中受益，这种情况下没有向量处理优化。

**“我再也没有半夜接到系统停机的电话了。我们认为 Exadata 是永远可用的。”**

*Westjet*  
首席技术专家  
**James Callaghan**

**“自 2011 年初采用 Exadata 以来，我们的关键电子支付服务始终保持 100% 的正常运行时间。该服务每周能够可靠地处理数十亿欧元的转账业务，并且可实现亚秒级的在线查询响应。”**

*Martin McGeough*  
Vocalink 数据库技术架构师

Exadata 独创性地实现了针对 **Oracle Database In-Memory 的容错内存复制功能**。在通用集群配置中，当某服务器节点发生故障时，该节点上的内存中数据会丢失，需要数分钟时间才能在正常节点上重新填充内存中数据。在此期间，分析查询运行速度会降低数个数量级。这意味着一般的平台将无法达到业务 SLA 的要求。而在 Exadata 上，容错内存复制特性会在集群数据库服务器之间复制内存中数据的所有子集，从而避免这种速度下降的情形。当某个数据库服务器发生故障时，查询将透明地访问正常数据库服务器上的复制副本，从而继续执行而不会中断。

Exadata 独特地集成了**活动数据卫士 (ADG)**，允许客户在备用数据库上运行内存中分析，从而进一步提高备用系统的投资回报，同时提升可用性和整体性能。

### 兼顾企业级安全性与卓越性能

Exadata 数据库云平台具有极高的安全性。在透明数据加密 (TDE) 等 Oracle 数据库高安全性功能的基础之上，Exadata **独创性地将解密处理从数据库服务器软件移至 Exadata 存储服务器硬件**。Exadata 存储通过硬件解密和压缩为数据库带来了极高的安全性和性能。**加密发生于数据压缩之后，因此解密开销也因压缩得到相应的降低**。借助这两项技术，Exadata 能以极低的开销、极快的速度（每秒数百 GB 原始用户数据）查询完全加密和压缩的数据库。此外，TDE 提供了一个完备的密钥管理解决方案来确保所有数据处于安全的加密状态。

Exadata 是一个集成设计的整体交付系统，而不是一系列组件的简单集合。在传统的数据库部署中，客户自行负责所有的系统集成工作，包括确保各个独立软件和硬件组件的安全性以及确保整个产品体系的安全性。**在 Exadata 数据库云平台中，Oracle 交付全堆栈的安全性**。Exadata 虚拟机在操作系统层面提供一个额外的隔离层。此外，无论是物理部署还是虚拟部署，Exadata 系统均使用最精简的 Oracle Linux 发行版，确保只安装和启用 Oracle 数据库的运行所需要的 RPM。采用该方法，系统安全性远比默认的 Linux 安装更强，避免了许多安全漏洞。此外，Exadata 利用 Oracle Linux 的 ksplice 功能，在 OS 保持联机的情况下应用安全更新。

Exadata 在机器加电时就通过 Secure Boot 开始提供安全保护，确保系统 UEFI 固件只允许执行有加密签名的、系统认为可靠的开机载入程序。每次重新启动服务器时，系统都会对执行的每个组件进行验证。这种做法可以防止恶意软件在启动链中隐藏嵌入代码。

此外，Exadata X8M 中使用的磁盘和闪存技术支持存储数据加密功能。使用存储数据加密，磁盘和闪存存储设备对进入设备的所有用户数据进行加密。当客户重用或停用 Exadata 时，Exadata 的安全擦除特性将利用该功能，通过改变加密用户数据所使用的加密密钥来即时擦除存储设备上存在的所有用户数据。利用安全擦除特性，由于前一个加密密钥已被删除，因此无需担心因超额供应或扇区备用而留在存储设备上的隐藏数据。

全球数百家行业领先银行、电信公司和政府组织已对 Exadata 的安全性进行了测定和评估。Oracle 将所有这些安全性评估结果都整合到了 Exadata 标准配置中。因此，Oracle 安全专家和世界各地数以百计的行业安全专家双方通过审查确保了 Exadata 的高度安全性。

**“将 350 台数据库服务器和存储系统整合到 Oracle Exadata 后，我们获得了一个高性能、可靠、可伸缩的移动计费平台，借助该平台，我们的账单数据计算速度提高了 10 倍，维护成本减少了一半。”**

**Tomoki Shimamura**  
计费系统部门高级经理  
NTT DoCoMo, Inc.

**“我们将信托投资销售系统的 20 台原有数据库服务器整合到四个 Oracle Exadata 数据库云平台中后，为客户提供信息的速度提高了 136 倍，增强了自身的竞争优势，并且能以更低的成本支持未来 10 年的事务增长。”**

**Tomoshiro Takemoto**  
高级董事总经理  
Nomura Research Institute Ltd.  
云计算服务部

## 任务关键的高可用性

Exadata 数据库云平台经过精心设计，旨在提供极高水平的可用性。**该平台针对所有故障类型提供全方位防护**，从磁盘、服务器和网络之类的简单故障，到复杂的站点故障和人为错误，可防范所有故障。每个 Exadata 数据库云平台都采用**完全冗余的硬件配置**，包括冗余网络、冗余配电单元 (PDU)、冗余电源以及冗余数据库服务器和存储服务器。Oracle RAC 可防范数据库服务器故障。Oracle ASM 通过数据镜像防范磁盘或存储服务器故障。Oracle RMAN 可以极其快速高效地将数据备份到磁盘或磁带。Oracle 闪回技术支持回退数据库、表甚至行级的用户错误。使用 Oracle 数据卫士 (DG)，用户可以在高可用性架构 (MAA) 配置下部署第二个 Oracle Exadata 数据库云平台，在远程站点透明地维护数据库的实时副本，针对主数据库故障和站点灾难提供全面保护。

分析公司 IDC 认为，采用 MAA 配置的 Exadata 系统的**可用性达到了 99.999% 以上**，该系统与 HP Integrity NonStop 和 IBM z Systems 一起被归为 IDC AL4 容错细分市场类别。<sup>1</sup>

Exadata 通过许多独特的方法来确保各种不同故障情况下的高可用性，这也体现了 Exadata 的软硬件深度集成的设计原则。**即时检测计算和存储服务器故障**就是这样一种**独有功能**。在非 Exadata 平台上，检测服务器故障需要停工等待很长的超时，从而导致应用长时间中断。Exadata X8M 实现了基于 RDMA 的独有的**亚秒级节点死机检测**，几乎可完全消除应用中断的情况。

磁盘和闪存设备有时会因内部故障扇区恢复、内部固件重启或耗损均衡而出现 I/O 操作延迟很长的情况。这些延迟很长的 I/O 操作可能会导致任务关键型 OLTP 数据库出现停滞。借助 Exadata **独有的 I/O 延迟限制**功能，当读取 I/O 延迟大大超过预期时，Oracle Exadata 系统软件可自动将读取 I/O 操作重定向到数据的 ASM 镜像副本。同样，它可自动将高延迟的写入 I/O 操作重定向到正常运行的闪存设备，从而消除写入异常。Exadata 系统软件利用机器学习技术来预测濒临故障的组件，并采取主动措施稳妥地停用此类组件。如果磁盘确实发生故障，ASM 对磁盘上保存的数据执行重新平衡操作。Exadata 支持在线更换磁盘、风扇、电源和 PCIe 闪存卡，从而避免停机。Exadata 系统软件进一步改进了重新平衡，当在存储服务器之间移动数据时保留闪存缓存填充信息和存储索引以保持一致的应用性能。在极少数情况下，当网络子系统内发生延迟尖峰时，Exadata 将数据库服务器发出的 I/O 重定向到另一台存储服务器。

Exadata 自动监视 CPU、内存、I/O、文件系统和网络。这种自动化将机器学习技术与我们通过数千次重大实际部署积累下来的深厚经验结合起来。例如，Exadata 可以检测出因不当使用系统资源而影响数据库性能的情况，识别出这是由哪个进程造成的，进而发出警报，对此用户无需进行任何预先设置。

**“在 Oracle 白金服务的支持下，我们的运营效率提高了 40%，因为我们几乎不需要专用的资源来执行更新，监视系统以及解决事件。”**

**Roland Schiller**  
Swiss Re 架构师

**“我们发现系统性能提高了 4 倍，总存储拥有成本降低了 30%，而这些还是在我们的数据量激增的情况下实现的。”**

**Holger Haun**  
董事总经理  
IDS GmbH – 分析与报告服务部

---

<sup>1</sup> Worldwide Fault-Tolerant Servers Market Shares, 2014: *Vendors Are Hearing the Customer — More Bold Moves Needed to Grow the Segment*, IDC, Peter Rutten, Lloyd Cohen, 2015 年 10 月

由于具备行业领先的可用性，领先企业纷纷部署 Exadata 数据库云平台来支持其至关重要的应用，包括银行间资金转账、在线证券交易、实时呼叫追踪以及基于 Web 的零售。Exadata 的任务关键型可用性功能不仅适用于 OLTP 负载，也适用于数据仓库和分析负载。

### 部署数据库即服务的理想平台

Exadata 数据库云平台可托管很多数据库，支持大规模数据库整合或复杂的数据库即服务私有云。所有多数据库环境都需要处理多样化、复杂和不可预测的混合负载，包括 OLTP、分析和批处理等可采用顺序和随机访问模式的操作。Exadata 可运行任何类型或任意组合的数据库负载，并且可提供行业领先的可扩展性和性能，这使其成为理想的整合平台 — 无论是整合多数据库负载、还是利用 Oracle Database 12c、18c 和 Oracle Database 19c 中的 Oracle 多租户整合可插拔数据库。

多数据库环境面临一个固有的风险，即一个数据库会占用过多资源，因而会影响其他数据库的服务质量。Exadata 数据库云平台对从应用到数据库的 CPU、网络和存储提供独特的端到端优先级排序。用户可以在物理数据库、可插拔数据库、连接、应用、用户甚至作业层面指定优先级和资源限制，确保各个整合数据库或 SQL 操作均可获得必要的资源和达到目标响应时间要求。

Exadata 采用了独特的数据库和 I/O 资源管理。为数据库层面的操作指定的细粒度优先级会自动传递至 Exadata 存储服务器并应用于各个 I/O 操作，确保 CPU 操作和 I/O 操作均以数据库操作优先级为准。当在一个 Exadata 机架上部署多个数据库时（整合私有云通常都是这种做法），同样可以采用这些资源管理原则。

作为 X8M 系列产品中的新特性，Exadata 还利用 RDMA over Converged Ethernet (RoCE) 协议确保报告、批处理和备份等网络密集型负载不会影响对延迟敏感的交互式负载。RAC 缓存融合通信和日志文件写入等对延迟敏感的网络操作将在融合以太网结构中的高优先级网络通道上传输。对延迟不敏感的流量则依托于各自的网络交换缓冲区在其他通道上传输。

得益于独有的整合和数据库即服务功能，Exadata 成为了仅有的能够在单个 Oracle 多租户容器数据库中支持多达 4000 个可插拔数据库的平台。

### 借助 Exadata 快照和容器快速部署开发和测试数据库

用户可以直接在 Exadata 上为测试和开发环境快速创建可节省空间的数据库快照。Exadata 数据库快照功能与 Oracle 多租户相集成，提供了一个极其简单的接口来创建新的可插拔数据库 (PDB) 快照。

快照始自生产数据库（即 PDB）的一个共享只读副本，其中的敏感信息已被清理。从这个共享副本可以创建读写快照的层级结构。当发生任何更改时，每一个快照都会将发生更改的块写入一个稀疏磁盘组。由于多个用户可以从同一个基础数据库副本创建独立快照，因此，多个测试和开发环境可以共享空间，但同时每一个任务保持独立的数据库。

**“Exadata 让我们每年节省了至少 50 万美元的运营成本。”**

*Westjet*  
首席技术专家  
**James Callaghan**

所有 Exadata 专有特性（如智能扫描、资源管理和智能闪存缓存）都可以在通过 Exadata 快照创建的数据库实例上无缝运行，因此只需占用一小部分宝贵的存储资源即可提供精确的测试和开发环境。Exadata 上的快照备份也能节省空间，因为只用备份更改的信息。

此外，Exadata 支持 Docker 容器，可帮助客户轻松、快速地供应用于测试和开发的数据库，推动敏捷开发。客户还可以使用 Docker 容器来托管轻型应用和代理。

### 全面的系统管理

Oracle Enterprise Manager 可以全面管理 Exadata 数据库云平台，提供从监视和报告到主动生命周期管理的全面功能。具体包括：

- 统一监视 — Oracle Enterprise Manager 13c 对所有硬件和软件组件（如数据库服务器、存储服务器和网络交换机）提供独有的统一视图，监视在这些组件上运行的操作及其资源使用情况。DBA 可以从数据库监视屏幕下钻至 Exadata 存储层，快速确定任何性能瓶颈的根源。
- 无人值守监视 — Enterprise Manager 中的无人值守监视功能针对 Exadata 进行了优化，预先定义了相关量度和阈值，因此管理员可在出现问题时及时收到通知并处理这些异常。此外，还可自动检测硬件故障并记录维护服务请求，从而缩短了解决问题的时间。
- Exachk 工具 — 该工具与 Enterprise Manager 强大的合规性框架相集成，让系统管理员能以自动化手段评估集成系统，以便发现已知的配置问题并遵循最佳实践。管理员可利用一致性检查功能检查各机架间或一个机架中的各个数据库服务器间的配置偏差。
- MS 进程 — Exadata 的内置 Management Server (MS) 进程持续监视硬件和软件组件的运行状况，并在检测到故障组件时，向管理员和 Oracle 支持部门发送警报。

### 极高的服务级别

Oracle 为 Exadata 产品系列提供了全面的支持服务，包括：24x7 硬件服务、系统监视、软件安装和配置以及其他标准服务和定制服务。

此外还有专为 Oracle 集成系统提供的特别有用的 **Oracle 白金服务**。白金服务提供故障监视服务，具有更快的响应速度，并且可以在需要时将问题快速上报至开发部门。根据白金服务协议，Oracle 支持工程师将远程执行软件维护和修补。白金服务为集成系统内所有软硬件（包括 Oracle 数据库）提供比以往所享有的更高的支持级别。对于 Exadata 客户来说，白金服务无需另外付费。

### IT 敏捷性

Exadata 为数据库的运行提供了一个完备的系统，包含存储、服务器和内部网络。传统数据库系统的管理工作通常由多个管理团队分担执行，如数据库团队、存储团队和系统管理团队，各自负责不同的组件。相比之下，**Exadata 系统的管理工作通常仅需一个统一的数据库云平台管理 (DMA) 团队即可完成**。数据库云平台管理员全面掌控 Exadata 数据库云平台中的所有资源，包括存储资源。数据库云平台管理员可以实施新的数据库部署和进行配置更改，不再需要分管不同组件的多个管理团队协同作业，这些团队通常都有繁重的工作并且它们对工作轻重缓急的考量也各不相同。这样，数据库云平台管理员可以将工作重点放在与应用和业务相关的增强任务上，而不是与各组件团队沟通协调，或者调优和修复低级别的配置问题。

## 显著降低成本

由于 Exadata 数据库云平台可提供超强性能、大存储容量和独有的压缩功能，因此，原本需要超大型传统硬件系统的负载现在可以在小得多的 Exadata 系统上运行。与传统系统相比，部署于 Exadata 系统上的应用对硬件的需求通常要低 2-4 倍。

Exadata 可为大型数据集提供巨大的 RAM、闪存和磁盘空间。一个全机架 Exadata 上的原始磁盘存储可超过 3 PB，而原始闪存存储可高达 920 TB。此外，混合列压缩通常可将存储和内存容量提升 10 倍。通过在磁盘、闪存和内存层之间智能地移动活跃数据，Exadata 不但极大地提高了性能，而且还将成本降至极低。

Exadata 可将支持多种负载的众多数据库整合到单一云平台中，这是一项**独有的能力**。高端 OLTP、分析、批处理、报告和备份操作都能以绝佳的性能同时在一个或多个数据库中运行。**Exadata 具备超强性能和超大容量，支持用户将大量数据库和负载整合到一个 Exadata 平台中。**将数据库整合到 Exadata 中不仅可以降低系统硬件成本、软件成本，还能大幅降低持续运营成本。

统一的 Exadata 数据库云平台配置有助于显著节省成本。**Exadata 不仅实现了技术标准，而且还实现了集成、测试、安全性、增强、调优和支持的标准化。**与传统系统相比，客户部署 Exadata 系统要快得多，所需工作量也少得多。低级的调优、集成和维护任务得以减少甚至完全消除。每家 Exadata 用户采用的配置都有数千家其他用户和 Oracle 内部团队在使用，由于众多用户采用同样的配置，用户遇到问题的可能性大为减少，有了问题也能轻松快捷地解决，从而能够降低运营成本和停机成本。

## “按需扩容”软件许可模式

X8M-8 数据库服务器搭载 8 个 24 核 x86 处理器（共 192 个内核），具有很大的计算容量。按需扩容特性允许在硬件安装期间禁用每台数据库服务器的一些内核，而保留启用至少 56 个内核。当负载增加，用户需要更多内核时，可使用按需扩容模式重新启用内核，按一次 2 个内核的方式获得软件许可。这种“按需购买”的软件许可方式是 Exadata 帮助用户根据业务增长控制成本的另一种途径。

## EXADATA 创造的业务效益

除了超强性能、高可用性、高安全性以及跨本地部署和云部署的部署灵活性等运营优势外，Exadata 还能直接创造业务效益。

由于大大缩短了系统配置、调优和测试所需的时间，**Exadata 可加快新业务应用的市场投放速度。**此外，部署时间也从数月缩短至数日，并大幅降低了上线后遇到意外系统级问题的风险。在部署新应用时，不可预测的应用使用模式往往会造成性能问题。Exadata 具有极高的 I/O、网络 and 计算吞吐量，可轻松承受不可预测的峰值负载，不会减慢任务关键型负载的响应速度。Exadata 总体上加快了应用部署速度并降低了风险，从而让企业加快了创新步伐。

**Exadata 可加快新业务应用的市场投放速度。**此外，部署时间也从数月缩短至数日，并大幅降低了上线后遇到意外系统级问题的风险。在部署新应用时，不可预测的应用使用模式往往会造成性能问题。Exadata 具有极高的 I/O、网络 and 计算吞吐量，可轻松承受不可预测的峰值负载，不会减慢任务关键型负载的响应速度。Exadata 总体上加快了应用部署速度并降低了风险，从而让企业加快了创新步伐。

Exadata 具备超强的性能以及庞大的内存和闪存容量，可大幅加快用户响应速度，从而提高员工效率和客户满意度。**用户将有更多的时间来进行有益的工作，而不是将时间浪费在等待系统响应上。**

Exadata 的超强性能不仅可以提高业务效率，还可以**帮助业务用户做出更明智的决策、发现增长机遇和降低成本。**用户可以实时分析数据、探索各种可能性以及通过快速反复找到更优秀的解决方案。

Exadata 可助力实现：

- 实时业务数据分析
- 更快的财务结算
- 更完善的规划和预算
- 更有效、更快速的预测

## 结论

Exadata 提供了一个完全集成的数据库平台，采用新的硬件技术和**独有的**软件来确保实现超强性能、高可用性和高安全性。此外，Exadata 还有助于降低成本、简化管理和增强可支持性，从而能够大大提高业务敏捷性和效率。凭借诸多优势，Exadata 当之无愧地成为了运行 Oracle 数据库的新的全球标准 — 无论是本地部署还是 Oracle 云部署。

## Exadata 服务器硬件<sup>1、2</sup>

服务器类型	CPU	内存	磁盘	闪存	网络
数据库服务器	8 个 24 核 Intel Xeon 8268 处理器 (2.9GHz)	3 TB (默认) 到 6 TB (最大)	无	2 个 6.4 TB PCIe NVMe 闪存卡	8 个 10/25 Gb 以太网端口 (客户端) 8 个 1/10 Gb 铜缆以太网端口 (1 个用于主机管理) 8 个 100 Gb QSFP28 RoCE 网络端口 1 个 ILOM 以太网端口
HC 存储服务器	2 个 16 核 Intel Xeon 5218 处理器 (2.3GHz)	192 GB 1.5 TB 持久性内存	12 个 14 TB 7200 转 RPM 磁盘	4 个 6.4 TB NVMe PCIe3.0 闪存卡	2 个 100 Gb QSFP28 RoCE 网络端口 1 个 1/10 Gb 铜缆以太网端口 (管理) 1 个 ILOM 以太网端口
EF 存储服务器		192 GB 1.5 TB 持久性内存	无	8 个 6.4 TB NVMe PCIe3.0 闪存卡	
XT 存储服务器	1 个 16 核 Intel Xeon 5218 处理器 (2.3GHz)	96 GB	12 个 14 TB 7200 RPM 磁盘	无	

<sup>1</sup> 所有服务器均配有冗余的可热插拔风扇和电源

<sup>2</sup> 该表仅包含可单独购买的服务器。HC = 大容量，EF = 极速闪存，XT = 扩展。

## Exadata 机架配置<sup>1、2</sup>

机架规格	数据库服务器和内核	存储服务器和内核	大容量存储服务器容量 (原始)	极速闪存存储服务器容量 (原始)
半机架 <sup>3</sup>	2 台服务器, 384 个内核	3 台服务器, 96 个内核用于 SQL 分流	504 TB 磁盘, 76.8 TB 闪存, 4.5 TB 持久性内存	153.5 TB 闪存, 4.5 TB 持久性内存
全机架 <sup>4</sup>	2 台服务器, 384 个内核	14 台服务器, 448 个内核用于 SQL 分流	2352 TB 磁盘, 358.4 TB 闪存, 21 TB 持久性内存	716.8 TB 闪存, 21 TB 持久性内存
+数据库服务器	额外增加一台服务器 <sup>5</sup> , 每个机架最多 576 个内核	不适用	不适用	不适用
+存储服务器	不适用	每个机架最多 14 台服务器 <sup>5</sup> , 448 个内核	每个机架最高 2352 TB 磁盘, 354.8 TB 闪存, 27.6 TB 持久性内存	每个机架最高 921.6 TB 闪存, 27.6 TB 持久性内存

<sup>1</sup> 每个机架高度为 42 RU (机架单元), 配有 2 个冗余配电单元 (PDU)、2 个 36 端口 100 Gb/秒 RoCE 交换机和 1 个用于管理的 48 端口管理以太网交换机。随附的备件包内含:

- 1 个 6.4 TB NVMe PCI 闪存卡和 1 个 14 TB 大容量磁盘, 或者
- 1 个 6.4 TB NVMe PCI 闪存卡

<sup>2</sup> 弹性配置允许在半机架中添加数据库服务器或存储服务器, 以便达到恰合应用需要的计算与存储比例。一个全机架弹性配置最多支持 16 台服务器和 42 RU (机架单元)。数据库服务器 = 5 RU, 存储服务器 = 2 RU

<sup>3</sup> 一个弹性配置最多支持 3 台数据库服务器。一个弹性配置最多支持 14 台存储服务器。

<sup>4</sup> 添加了全机架配置, 作为之前以静态规模实现弹性配置的示例。

## 其他弹性扩展方案

多机架连接	通过 RoCE 网络结构连接达 18 个任意组合的 Exadata 数据库云平台机架或 Exadata 存储扩展机架。通过外部 RoCE 交换机创建更大规模的配置所连接的机架必须包含 Exadata RoCE 硬件。
-------	--

## Exadata 容量和性能指标：各个服务器

服务器类型	最高 SQL 闪存带宽 <sup>2</sup>	最高 SQL 读取 IOPS <sup>1,3</sup>	最高 SQL 写入 IOPS <sup>4</sup>	持久性内存容量 (原始) <sup>5</sup>	PCI 闪存容量 (原始) <sup>5</sup>	磁盘数据容量 (原始) <sup>5</sup>
HC 存储服务器 <sup>1</sup>	25 GB/秒	1500000	470000	1.5 TB	25.6 TB	168 TB
EF 存储服务器 <sup>1</sup>	40 GB/秒	1500000	470000	1.5 TB	51.2 TB	不适用
XT 存储服务器 <sup>1</sup>	不适用	不适用	不适用	不适用	不适用	168 TB

<sup>1</sup> HC = 大容量, EF = 极速闪存, XT = 扩展。PMEM = 持久性内存。实际系统性能随应用而不同。

<sup>2</sup> 该带宽是在不采用数据库压缩的情况下运行 SQL 时得到的峰值物理扫描带宽。在采用数据库压缩的情况下, 有效数据带宽将高于此值。

<sup>3</sup> 基于运行 SQL 的 8K I/O 请求得出。请注意, I/O 大小对闪存 IOPS 的影响很大。其他产品供应商报出的 IOPS 可能基于与数据库无关的、较小的 I/O 大小。

<sup>4</sup> 基于运行 SQL 的 8K I/O 请求得出。执行 ASM 镜像之后在存储服务器上测得的闪存写入 I/O, 这通常会发出多个存储 I/O 来保持冗余性。

<sup>5</sup> 物理容量以标准磁盘驱动器术语来计量, 1 GB = 10 亿字节。

## Exadata 典型机架配置：闪存容量和性能指标 (HC 和 EF)

闪存指标		最高 SQL 闪存带宽 <sup>2</sup>	最高 SQL PMEM 读取 IOPS <sup>1,3</sup>	最高 SQL 闪存写入 IOPS <sup>4</sup>	PCI 闪存容量 (原始) <sup>5</sup>
全机架	HC <sup>1</sup>	350 GB/秒	6000000	5000000	358.4 TB
	EF <sup>1</sup>	560 GB/秒	6000000	5000000	716.8 TB
半机架	HC	75 GB/秒	4500000	1410000	76.8 TB
	EF	120 GB/秒	4500000	1410000	153.6 TB

<sup>1</sup> EF = 极速闪存; HC = 大容量; PMEM = 持久性内存

<sup>2</sup> 该带宽是在不采用数据库压缩的情况下运行 SQL 时得到的峰值物理扫描带宽。在采用数据库压缩的情况下, 有效数据带宽将高于此值。

<sup>3</sup> 基于运行 SQL 的 8K IO 请求得出。请注意, IO 大小对闪存 IOPS 的影响很大。其他供应商报出的 IOPS 可能基于与数据库无关的、较小的 IO 大小。

<sup>4</sup> 基于运行 SQL 的 8K IO 请求得出。执行 ASM 镜像之后在存储服务器上测得的闪存写入 I/O, 这通常会发出多个存储 IO 来保持冗余性。

<sup>5</sup> 物理容量以标准磁盘驱动器术语来计量, 1 GB = 10 亿字节。测量可用容量时使用的是常规二进制空间术语, 1 TB = 1024

\* 1024 \* 1024 \* 1024 字节。

## Exadata 典型机架配置：磁盘容量和性能指标 (HC)

磁盘指标	最高 SQL 磁盘带宽 <sup>1</sup>	最高 SQL 磁盘 IOPS <sup>2</sup>	数据容量 (原始) <sup>3</sup>
全机架	25 GB/秒	36000	2352 TB
半机架	5.4 GB/秒	7800	504 TB

<sup>1</sup>该带宽是在不采用数据库压缩的情况下运行 SQL 时得到的峰值物理扫描带宽。在采用数据库压缩的情况下，有效数据带宽将高于此值。

<sup>2</sup>基于运行 SQL 的 8K IO 请求得出。请注意，IO 大小对闪存 IOPS 的影响很大。其他供应商报出的 IOPS 可能基于与数据库无关的、较小的 IO 大小。

<sup>3</sup>原始容量由标准磁盘驱动器术语计算得出，1 GB = 10 亿字节。测量可用容量时使用的是常规二进制空间术语，1 TB = 1024

\* 1024 \* 1024 \* 1024 字节。

## Exadata 典型机架配置：综合指标 (HC 和 EF)

综合指标		数据容量 (可用) 常规冗余 <sup>1</sup>	数据容量 (可用) 高冗余 <sup>1</sup>	最大数据加载速度 <sup>2</sup>
全机架	HC	953 TB	699 TB	35 TB/小时
	EF	282 TB	206 TB	35 TB/小时
半机架	HC	191 TB	150 TB	7.5 TB/小时
	EF	56 TB	44 TB	7.5 TB/小时

<sup>1</sup>可用容量以常规二进制空间术语来计量，即 1 TB = 1024 \* 1024 \* 1024 \* 1024 字节。该容量是在考虑实现 ASM 冗余性（可从1个驱动器故障中恢复）所需的空间之后用于创建数据库的实际可用空间。常规冗余计算反映了 Grid Infrastructure 12.2.0.1 或更高版本的使用。

<sup>2</sup>加载速率通常受数据库服务器 CPU 而非 I/O 的限制。根据加载方法、索引、数据类型、压缩和分区不同，速率也会有所变化。

## Exadata 数据库云平台组件环境规格

指标	EXADATA 数据库服务器 X8M 8 + 网络结构	EXADATA 存储服务器 X8M 2 大容量 (HC) + 网络结构	EXADATA 存储服务器 X8M 2 极速闪存 (EF) + 网络结构	EXADATA 存储服务器 X8M 2 扩展 (XT) + 网络结构	EXADATA 八分之一机架存储服务器 X8M 2 大容量 (HC) + 网络结构
高度	8.63 英寸 (219.25 毫米)	3.42 英寸 (86.9 毫米)			
宽度	17.5 英寸 (445 毫米)	17.52 英寸 (445.0 毫米)			
深度	32.8 英寸 (833 毫米)	29.88 英寸 (759.0 毫米)			
噪音 (工作)	8.7 B	8.2 B	8.2 B	8.2 B	8.2 B
重量	198.8 磅 (90.2 千克)	76.7 磅 (34.8 千克)	60.6 磅 (27.5 千克)	66.7 磅 (30.2 千克)	67.5 磅 (30.6 千克)
最大功耗	3.5 kW (3.6 kVA)	0.8 kW (0.8 kVA)	0.8 kW (0.8 kVA)	0.5 kW (0.5 kVA)	0.6 kW (0.6 kVA)
常规功耗 <sup>1</sup>	2.5 kW (2.5 kVA)	0.5 kW (0.6 kVA)	0.6 kW (0.6 kVA)	0.3 kW (0.3 kVA)	0.4 kW (0.4 kVA)
最大负荷下的散热率	11953 BTU/小时	2631 BTU/小时	2730 BTU/小时	1570 BTU/小时	1947 BTU/小时
	12610 kJ/小时	2775 kJ/小时	2880 kJ/小时	1656 kJ/小时	2054 kJ/小时
常规负荷下的散热率	8367 BTU/小时	1842 BTU/小时	1911 BTU/小时	1099 BTU/小时	1363 BTU/小时
	8827 kJ/小时	1943 kJ/小时	2016 kJ/小时	1159 kJ/小时	1438 kJ/小时
最大负荷下的气流 <sup>2</sup>	553 CFM	122 CFM	126 CFM	73 CFM	90 CFM
常规负荷下的气流 <sup>2</sup>	387 CFM	85 CFM	88 CFM	51 CFM	63 CFM

工作温度/湿度: 5 °C 至 32 °C (41 °F 至 89.6 °F) , 相对湿度 10% 至 90%, 无冷凝 工作海拔: 最高 3048 米, 900 米以上每上升 300 米最高环境温度下降 1 °C

<sup>1</sup> 常规功耗随应用负载而不同。

<sup>2</sup> 气流方向必须从前往后。

## Exadata 数据库云平台环境规格

指标	全机架	半机架
高度	78.74 英寸 (2000 毫米)	
宽度	23.66 英寸 (601 毫米)	
深度	47.13 英寸 (1197 毫米)	
噪音 (工作)	9.6 B	9.3 B
采用大容量磁盘的环境		
重量	2049.5 磅 (929.6 千克)	1243.8 磅 (564.2 千克)
最大功耗	18.7 kW (19.0 kVA)	10.2 kW (10.4 kVA)
常规功耗 <sup>1</sup>	13.1 kW (13.3 kVA)	7.1 kW (7.3 kVA)
最大负荷下的散热率	63643 BTU/小时	34705 BTU/小时
	67144 kJ/小时	36614 kJ/小时
常规负荷下的散热率	44550 BTU/小时	24293 BTU/小时
	47001 kJ/小时	25630 kJ/小时
最大负荷下的气流 <sup>2</sup>	2946 CFM	1607 CFM
常规负荷下的气流 <sup>2</sup>	2063 CFM	1125 CFM
采用极速闪存驱动器环境		
重量	1824.1 磅 (827.4 千克)	1195.5 磅 (542.3 千克)
最大功耗	19.1 kW (19.4 kVA)	10.3 kW (10.5 kVA)
常规功耗 <sup>1</sup>	13.3 kW (13.6 kVA)	7.2 kW (7.3 kVA)
最大负荷下的散热率	65029 BTU/小时	35002 BTU/小时
	68605 kJ/小时	36927 kJ/小时
常规负荷下的散热率	45520 BTU/小时	24501 BTU/小时
	48024 kJ/小时	25849 kJ/小时
最大负荷下的气流 <sup>2</sup>	3011 CFM	1620 CFM
常规负荷下的气流 <sup>2</sup>	2107 CFM	1134 CFM

工作温度/湿度: 5 °C 至 32 °C (41 °F 至 89.6 °F), 相对湿度 10% 至 90%, 无冷凝 工作海拔: 最高 3048 米, 900 米以上每上升 300 米最高环境温度下降 1 °C

<sup>1</sup> 常规功耗随应用负载而不同。

<sup>2</sup> 气流方向必须从前往后。

## Exadata 数据库云平台 X8M-8 相关法规和认证

法规 <sup>1, 2, 3</sup>	产品安全:	UL/CSA 60950-1、EN 60950-1、IEC 60950-1 CB
	EMC	
	辐射:	FCC CFR 47 第 15 部分、ICES-003、EN55032
	抗干扰性:	EN55024
认证 <sup>2, 3</sup>	北美 (NRTL)、欧盟 (EU)、International CB Scheme、HSE Exemption (印度)、BSMI (中国台湾)、CCC (中华人民共和国)、EAC (EAEU 包括俄罗斯)、RCM (澳大利亚)、VCCI (日本)、墨西哥、KC (韩国)	
欧盟指令 <sup>3</sup>	2014/35/EU 低电压指令、2014/30/EU EMC 指令、2011/65/EU RoHS 指令、2012/19/EU	

<sup>1</sup>本文中引用的所有标准和认证都是截止到编写本产品介绍时的新的官方版本。如需了解更多详细信息，请联系您的销售代表。

<sup>2</sup>其他国家和地区的法规/认证也可能适用。

<sup>3</sup>有些情况下，仅在机柜级别的系统达到了法规和认证合规性（如适用）。

## Exadata 数据库云平台支持服务

- 硬件保修：提供 1 年保修。如正常营业时间（周一至周五，上午 8 点至下午 5 点）内接到维修请求，则 4 小时内通过 Web/电话进行回复，2 个营业日内到达现场调换部件
- Oracle 标准系统支持服务包括 Oracle Linux 支持服务以及 24x7 的 2 小时内现场硬件维修响应（维修地点需处于维修中心附近）
- Oracle 标准操作系统服务
- Oracle 客户数据与设备保留服务
- 系统安装服务
- 软件配置服务
- Oracle 白金服务
- 关键业务系统服务
- Oracle Exadata 启动包
- 系统升级支持服务，包括硬件安装和软件配置
- Oracle 自动服务请求 (ASR)

## 主要特性和功能

### Exadata 和数据库软件特性 — 分析

- 自动并行化处理数据扫描并将数据扫描分流至存储（独有特性）
- 在存储中基于“where”子句筛选行（独有特性）
- 在存储中基于所选列筛选行（独有特性）
- 对 JSON 和 XML 分析查询进行存储分流（独有特性）
- 在存储中基于与其他表的联接筛选行（独有特性）
- 混合列压缩（独有特性）
- 存储索引数据忽略（独有特性）
- 按用户、查询、服务、数据库等管理 I/O 资源（独有特性）
- 在闪存缓存中自动转换为列格式（独有特性）
- 对表扫描进行智能闪存缓存（独有特性）
- 对索引快速全扫描进行存储分流（独有特性）
- 对加密数据扫描进行存储分流并确保 FIPS 合规性（独有特性）
- 针对 LOB 和 CLOB 的存储分流（独有特性）
- 针对最小值/最大值运算的存储分流（独有特性）
- 针对数据挖掘的存储分流（独有特性）
- 如果存储 CPU 繁忙，则反向分流到数据库服务器
- 在闪存缓存中自动列化数据（独有特性）
- 当数据加载至闪存缓存时，自动将其转换为内存中格式（独有特性）

#### Exadata 和数据库软件特性 — OLTP

- 数据库感知式 PCI 闪存 (独有特性)
- Exadata 智能闪存缓存 (独有特性)
- Exadata 智能闪存日志 (独有特性)
- 智能写回闪存缓存 (独有特性)
- 按数据库、用户或负载确定 I/O 优先级以确保服务质量 (独有特性)
- Exafusion Direct-to-Wire 协议 (独有特性)
- 数据库智能网络资源管理 (独有特性)
- Exachk 全体系验证 (独有特性)
- 全体系安全性扫描 (独有特性)
- 数据库范围的安全性 (独有特性)
- 通过单元到单元重新平衡保留闪存缓存和存储索引 (独有特性)
- 全体系安全擦除 (独有特性)
- 即时创建数据文件 (独有特性)
- 智能融合块传输 (独有特性)
- 控制每个数据库的闪存缓存大小 (独有特性)
- In-Memory OLTP 加速 (独有特性)
- Undo 块远程 RDMA 读取 (独有特性)
- 通过多租户选项支持超过 252 个可插拔数据库 (独有特性)

#### Exadata 和数据库软件特性 — 高可用性

- 即时检测节点或单元故障 (独有特性)
- 内存中容错 (独有特性)
- 亚秒级故障切换停滞磁盘或闪存上的 I/O (独有特性)
- 将备份分流到存储服务器 (独有特性)
- Exadata 数据验证 (扩展式 H.A.R.D.) (独有特性)
- 优先恢复关键数据库文件 (独有特性)
- 通过读取其他存储服务器来自动修复损坏的磁盘数据 (独有特性)
- 避免在预计存在故障的磁盘上执行读取 I/O 操作 (独有特性)
- 限制使用、关闭再打开暂时性能低下的驱动器 (独有特性)
- 当镜像存储服务器发生故障时提供关闭防护 (独有特性)
- 检测并禁用不可靠的网络链路 (独有特性)
- 重新平衡时保留存储索引 (独有特性)

#### 可管理性特性

- 嵌入式 Oracle Integrated Lights Out Manager (ILOM)
- Oracle Enterprise Manager Exadata 插件
- 活动 AWR 包含存储统计信息, 以支持端到端监视 (独有特性)
- 针对以太网连接的 IPv6 支持
- 按需扩容
- 单元软件透明重启
- 闪存和磁盘生命周期管理警报
- 自动磁盘清理和修复
- 自动创建 VLAN
- Oracle Exadata Deployment Assistant
- 单独的管理交换机和连接
- 从远程服务器执行 Exacli 命令行管理
- 通过 Cellcli 命令行管理存储服务器
- DCLI 分布式命令行自动化工具

#### Oracle 数据库软件 (需单独购买) :

- **对于数据库服务器:** Oracle Database 11g R2 企业版、Oracle Database 12c 企业版 R1 和 R2、Oracle Database 18c 企业版 R1 以及 Oracle Database 19c。Oracle 数据库选项, 如 Oracle 真正应用集群 (RAC)、Oracle 分区、Oracle 多租户、Oracle 活动数据卫士 (ADG)。关于所支持的特性, 请参阅具体版本的相应文档。
- **对于存储服务器:** Oracle Exadata 系统软件。允许将许可从一个系统转移至另一个系统, 或转移至新系统。

**Oracle 软件（随机附带）：**

- **对于数据库服务器：**采用 Unbreakable Enterprise Kernel 5 的 Oracle Linux 7 Update 7。零丢失复制数据报协议 (ZDP) RoCEv2 协议，用于 Exadata 存储服务器和 Oracle 数据库之间的通信。该协议基于可靠数据报套接字 (RDS) OpenFabrics 企业发行版 (OFED)

## 关注我们

请致电 400-699-8888 或访问 [oracle.com/cn](http://oracle.com/cn)。

中国地区的用户请访问 [oracle.com/cn/corporate/contact/index.html](http://oracle.com/cn/corporate/contact/index.html)，查找您当地 Oracle 办事处的电话号码。

 [blogs.oracle.com/oracl](http://blogs.oracle.com/oracl)

 [facebook.com/oracl](https://facebook.com/oracl)

 [twitter.com/oracl](https://twitter.com/oracl)

## Integrated Cloud Applications & Platform Services

版权所有 © 2019, 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。本文档不保证没有错误，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。其他名称可能分别是其所有者的商标。

Intel、Intel Optane 和 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。  
0919