

Oracle 大数据一体机 X8-2

ORACLE® BIG DATA APPLIANCE



主要特性

- 高度可扩展的开放式基础架构，用于存储、分析和管理工作数据
- 行业领先的安全性、卓越的性能和包罗万象的大数据工具集汇聚一堂，通过大数据一体机轻松实现部署
- 灵活的配置和弹性硬件选择，既能减少占地面积，又能利用 Cloudera 5.x 或 6.x 对 Hadoop 逐步扩容
- Oracle Big Data SQL 让大数据和 Oracle 数据库实现了前所未有的集成
- 性能卓越的高级 AI 和 ML 计算平台
- 全面的 Cloudera 软件套件，内含 Apache Hadoop 和 Apache Spark 的 Cloudera 发行版

Oracle 大数据一体机是一个灵活、高性能、安全的平台，用于在 Hadoop、Kafka 和 Spark 上运行不同的负载。凭借 Oracle Big Data SQL，Oracle 大数据一体机将 Oracle 行业领先的 SQL 实现扩展至 Hadoop/NoSQL 和 Kafka 系统。Oracle 大数据一体机兼收并蓄地将 Hadoop 生态系统的新技术和强大的 Oracle SQL 功能整合到了一个预先配置的平台，因而具有独到的能力，支持用户快速开发新的大数据应用并实现与现有关系数据的紧密集成。

Oracle 大数据一体机 X8-2

Oracle 大数据一体机是一个用于 Hadoop 和 Spark 负载以及流数据处理的多用途、开放式集成系统。大数据一体机旨在运行多种工作负载 — 从 Hadoop 专有负载（Yarn、Spark、Hive 等）到使用 Oracle Big Data SQL 跨 Apache Kafka、Hadoop 和 NoSQL 数据库执行的综合性交互式 SQL 查询。大数据一体机是一个经过 Cloudera 认证的平台，支持 Cloudera 5.x 与 Cloudera 6.x。

大数据一体机提供了开放的创新环境，同时保持了紧密的集成和企业级支持。企业可通过部署外部软件来支持新功能，例如图形分析，自然语言处理和欺诈检测。非 Oracle 组件将由其各自的支持渠道而非 Oracle 来提供支持。

降低 TCO 并更快创造价值

大数据一体机采用独特的定价机制，与自行构建的 Hadoop、Spark 或 Kafka 系统相比，不仅初始部署成本更低，而且其三年和四年的总拥有成本也显著降低。大数据一体机包括了高可用的硬件资源（服务器、高速网络、配电单元和外围设备）、操作系统和 Java 支持，以及 Cloudera 发行版（包括 Apache Hadoop）的订阅费用。Oracle 为硬件和集成软件（包括 Cloudera CDH）提供统一的支持联络点。

企业不希望将宝贵的人力资源用在组装和调优 Hadoop 基础设施上，尤其是这些资源要用于交付高价值的业务解决方案时。大数据一体机为 Cloudera 企业版和 Oracle NoSQL Database 提供了一个经过预先配置、高度调优的环境。借助这个优化的环境，企业可将其资源集中用于开发令人惊叹的业务应用，同时降低了解决方案的风险。此外，经过预先调优的环境还避免了新应用因性能和生产问题而迟迟不能上线。

- Oracle Enterprise Manager 与 Cloudera Manager 相结合，简化了整个大数据一体机的管理
- 使用 Oracle R 实现的高级分析可以与 HDFS 中存储的数据直接交互
- 节点之间、机架之间，以及与 Oracle Exadata 及其他 Oracle 集成系统之间均采用 InfiniBand 连接

主要优势

- 开放、安全且经过优化的大数据解决方案
- 整个堆栈（操作系统、Java、Oracle NoSQL Database 和 Cloudera 堆栈）使用同样的命令工具，简化了操作、更新和补丁管理
- Oracle Big Data SQL 与 Oracle Exadata 的集成为大数据提供了全面的 SQL 解决方案
- 单台设备上集成了全面的大数据工具集
- 无风险安装、快速实现价值
- 通过 Oracle Enterprise Manager 在单一管理控制台上集成大数据一体机硬件和软件监控
- 整个大数据解决方案（包括硬件和软件）由单一供应商提供支持

Oracle 将继续努力优化软件体系，让客户能够通过后续软件版本获得越来越多的优势，而不必为调优 Hadoop 花费高额咨询费。

全面的安全性

数据安全性对于企业的所有数据解决方案至关重要。大数据一体机对 Hadoop 中的数据提供了开箱即用的强身份验证、授权和审计功能。

使用 Kerberos 提供强身份验证。这可以确保所有用户身份的真实性，避免将恶意服务添加到系统中。大数据一体机利用 Apache Sentry 对通过 Hive、Impala 和 Big Data SQL 等工具进行的 SQL 访问进行授权。

Oracle 大数据一体机配置了网络加密和静态数据加密功能。它通过 Cloudera Key Trustee 支持 HDFS 透明加密，支持静态数据加密的新创新。此实现为 HDFS 中的所有数据提供了严密的安全性。网络加密可阻止网络嗅探捕获受保护的数据，在大数据一体机上通过一个简单的复选框即可启用。

为了确保安全性和数据访问合规性，Oracle 大数据一体机提供 Cloudera Navigator 来跟踪对大数据一体机中数据的所有访问。

除保护 Hadoop 系统外，组织在查询 Hadoop、Kafka 和 NoSQL 数据时，还可通过 Oracle Big Data SQL 利用 Oracle 数据库的安全功能。安全的大数据一体机再加上 Oracle Big Data SQL，为组织带来了超越当今市场上任何其他大数据系统的全面的安全性。

简化运维操作

Oracle 大数据一体机简化了日常运维操作，它提供简单的单命令安装、更新和打补丁以及扩展工具 Mammoth，Mammoth 可对不断发展中的 Hadoop 体系进行快速部署更新（通常每季度一次），而不会造成长时间停机。Mammoth 还支持在不同 Hadoop 版本之间进行经过 Oracle 测试的无缝升级，并且可通过自动化服务管理来确保在 Hadoop 主节点与数据节点之间达到平衡。

大数据一体机由 Oracle 提供支持，从而为组织提供对其硬件、所有集成软件（包括所有 Cloudera 软件）及所安装的任何其他 Oracle 软件的统一的支持。

灵活的配置

Oracle 大数据一体机旨在随着您的数据和需求的增长而扩展。最初实施大数据解决方案时，用户可以采用 Oracle 大数据一体机的入门机架。入门机架配备了全机架所需的全套交换机和配电单元 (PDU)。利用入门机架和交换连接装置，大数据一体机可以使用 Oracle 大数据一体机 X8-2 大容量 (HC) 节点以及 InfiniBand 基础架构以逐次添加一台节点硬件的方式轻松高效地扩展至更大的配置。

相关产品

- Oracle Exadata
- Oracle Big Data SQL
- Oracle Big Data Spatial and Graph
- Oracle Big Data Connectors
- Oracle Big Data Discovery
- Oracle NoSQL Database
- Oracle Exalytics
- Oracle Business Intelligence 企业版
- Oracle GoldenGate for Big Data
- Oracle Data Integrator
- Oracle Stream Explorer
- Oracle Enterprise Manager

相关服务

- 高级客户服务
- 产品支持服务
- 咨询服务
- Oracle 大学培训

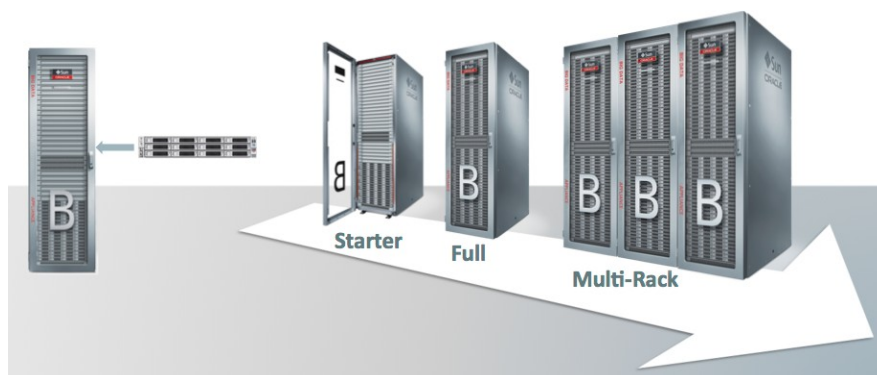


图 1. 模块化的硬件构建块

除了在一个机架内扩展系统之外，还可以通过集成的 InfiniBand 结构连接多个机架，形成更大规模的配置；支持以无阻塞方式连接多达 18 个机架，而无需任何外部交换机。如果增加外部 InfiniBand 交换机，还支持更大规模的无阻塞式配置；不增加交换机时，则支持更大规模的阻塞式网络配置。InfiniBand 的使用减少了对高带宽顶部机架交换基础设施的需要，因此显著降低了大规模配置的成本。

Oracle 大数据一体机采用多租户机制，可以将其配置为单个集群或一组集群。这使得客户能够灵活地部署开发、测试和生产集群。

将 Hadoop 作业连接到 Oracle 数据库

Oracle DataSource for Apache Hadoop 是 Oracle 大数据一体机的一项特性，可将 Oracle 数据库表转换为 Hadoop 和 Spark 数据源，从而可以利用最佳连接性从大数据一体机查询查询到 Oracle 数据库。DataSource for Apache Hadoop 支持使用 Hive SQL、Spark SQL，以及 Hadoop 和 Spark API 直接、一致地访问 Oracle 数据库中的数据，这些 API 支持 HCatalog、InputFormat、SerDes 和 StorageHandler（外部表）。对 Oracle 数据库的数据访问将通过安全连接（Kerberos、SSL、Oracle Wallet）并行执行。

Oracle Big Data SQL

Oracle Big Data SQL 是 Oracle 推出的一项数据虚拟化创新。它是一种支持在 Hadoop 上使用 SQL 的新架构，可将 Hadoop、Kafka 和 NoSQL 中的数据与 Oracle 数据库中的数据无缝整合。Big Data SQL 既可以用于 Oracle 大数据一体机，也可以用于其他支持的 Apache Hadoop 平台。使用 Oracle Big Data SQL，组织能够：

- 在同一个 SQL 查询中整合来自 Oracle 数据库、Hadoop/NoSQL 和 Kafka 的数据
- 查询和分析 Hadoop/NoSQL、Kafka 及其他系统中的数据
- 将大数据分析集成到现有的应用和架构中
- 将安全和访问策略由 Oracle 数据库延伸至 Hadoop 和 NoSQL 中的数据
- 利用智能扫描，尽可能提高对所有数据的查询性能

通过以下两项强大的特性，Oracle Big Data SQL 从根本上简化了大数据领域中的集成和操作：新扩展的外部表和 Hadoop 上的智能扫描功能。

Oracle 数据库用户可利用支持 Big Data SQL 的外部表类型来查看 Hadoop 和 NoSQL 中的数据。这些表一经定义，即可自动发现 Hive 元数据，包括数据位置和数据解析需求（例如 SerDes 和 StorageHandler）。这样，SQL 查询便可利用原生解析结构对数据进行访问，而数据格式维持不变。

智能扫描为 Oracle 所独有，是 Oracle Exadata 中的一项成熟的存储处理创新功能，如今，Oracle 将这项功能引入到了大数据一体机中。在数据处理过程中，很多性能损失往往缘于过多的数据移动。Hadoop 上的智能扫描不是将所有扫描数据都发送给计算资源，而是通过在存储层面运用以下技术来尽可能减少数据到计算节点的移动：

- 数据本地化扫描
 - 使用节点本地的原生运算符读取 Hadoop 数据
- 列投影
 - 仅将相关的列从数据源返回到数据库引擎
- 谓词评估和计算下推
 - 仅从数据源返回相关的行
 - 对查询利用底层存储格式（Parquet、ORC 等等）以实现高选择性
- 存储索引
 - 避免 Hadoop 大量 IO 扫描，提供查询加速仅从源返回相关行
- 复杂函数评估
 - 在数据源对 JSON 和 XML 类型应用 SQL 运算符
 - 在数据源评估模型评分和分析运算符
- 聚合
 - 聚合数据并将汇总结果返回数据库

智能扫描和存储索引可与 Hadoop 服务共存，而且不需要对 Hadoop 本身进行任何修改，因而符合 Oracle 大数据一体机环境的开放性原则。

Oracle Big Data Spatial and Graph

Oracle Big Data Spatial and Graph 在 Oracle 大数据一体机和其他受支持的 Apache Hadoop 和 NoSQL Database 平台上提供高级空间分析功能和图数据库。

属性图组件为用户提供了一个具有行业领先性的内存分析功能的可扩展的图数据库。它包含 35 个预制的图分析功能，用户可利用这些功能轻松发现关系、群体、影响者和其他图形模式。图数据库托管在 Apache HBase 或 Oracle NoSQL Database 上，支持 Python 和 Groovy 等流行的脚本语言，还支持开源 Tinkerpop 体系以及 Java API。

空间分析和图服务包括数据充实服务，用于基于位置和地名协调数据，还包括各种 2D、3D 和栅格算法，用于分析（举例来说）社交媒体或日志数据中人员之间和资产之间的位置关系。它可应用市、省/州、国家/地区分类方法，处理以及可视化地理空间地图数据和卫星图像。

Oracle Big Data Connectors

除了提供 Oracle Big Data SQL 和完备的 Cloudera 软件平台，大数据一体机还利用 Oracle Big Data Connectors 简化了数据集成和分析。使用 Big Data Connectors，可以从 Oracle Exadata 和 Oracle 数据库高速访问 Hadoop 中的数据，数据传输率高达 15 TB/小时。Big Data Connectors 还支持集成的、高度可扩展的分析，允许使用 Oracle R Distribution 以原生方式访问 Hadoop 数据和并行处理。此外，Oracle XQuery for Hadoop 促进了标准 XQuery 操作来处理 and 转换各种格式（JSON、XML、Avro 和其他格式）的文档，并跨 Hadoop 集群并行执行。

软件细节

大数据一体机 X8-2 — 随附软件

操作系统：

- Oracle Linux 6 或 Oracle Linux 7
-

集成软件:

Cloudera Enterprise 5 或 Cloudera Enterprise 6 — 数据中心版, 支持以下组件:

- 包含 Apache Hadoop 的 Cloudera 发行版 (CDH)
- Cloudera Impala
- Cloudera Search
- Apache HBase 和 Apache Accumulo
- Apache Spark
- Apache Kafka
- Cloudera Manager, 支持以下组件:
 - Cloudera Navigator
 - Cloudera Back-up and Disaster Recovery (BDR)

Oracle DataSource for Apache Hadoop

其他:

Oracle Java JDK 8

MySQL Database Enterprise Server — 高级版*

Oracle 大数据一体机 Enterprise Manager 插件

Oracle R Distribution

Oracle NoSQL Database 社区版 (CE)**

* 限制使用许可

** 不支持 Oracle NoSQL Database CE

大数据一体机 X8-2 — 可选 Oracle 软件

Oracle Big Data SQL

Oracle Big Data Connectors:

- Oracle SQL Connector for Hadoop
- Oracle Loader for Hadoop
- Oracle XQuery for Hadoop
- Oracle R Advanced Analytics for Hadoop
- Oracle Data Integrator

Oracle Data Integrator

Oracle GoldenGate for Big Data

Oracle NoSQL Database 企业版

Oracle Big Data Spatial and Graph

硬件细节及规格

大数据一体机 X8-2 — 硬件

全机架	入门机架	HC 节点 + InfiniBand 基础架构
18 个计算/存储节点	6 个计算/存储节点	1 个计算/存储节点

每个节点：

- 2 个 24 核 (2.4GHz) Intel® Xeon® P8260
- 8 个 32 GB DDR4-2666 MHz 内存，可扩展至 1.5TB
- 12 个 14 TB 7200 RPM 大容量 SAS 磁盘
- 2 个 240 GB M.2 SATA SSD 磁盘
- 2 个 QDR 40 Gb/秒 InfiniBand 端口
- 1 个双口 InfiniBand QDR CX3 (40 Gb/秒) PCIe HCA
- 1 个内置 RJ45 千兆以太网端口

2 个 32 端口 QDR InfiniBand 叶交换机 <ul style="list-style-type: none"> • 32 个 InfiniBand 端口 • 8 个 10 Gb 以太网端口 	使用入门机架中的叶交换机
1 个 36 端口 QDR InfiniBand 主干交换机 <ul style="list-style-type: none"> • 36 个 InfiniBand 端口 	使用入门机架中的主干交换机
其他硬件组件包括： <ul style="list-style-type: none"> • 以太网管理交换机 • 2 个冗余配电单元 (PDU) • 42U 架装式机柜 	使用入门机架中的管理交换机、PDU 和基本机架
备件套件包括： <ul style="list-style-type: none"> • 1 个 14 TB 大容量 SAS 磁盘 • InfiniBand 电缆 	使用入门机架中的备件套件

大数据一体机 X8-2 组件环境规格**X8-2 大容量节点 + InfiniBand 基础架构****物理尺寸**

高度	3.5 英寸 (87.6 毫米)
宽度	17.5 英寸 (445.0 毫米)
深度	29.0 英寸 (737.0 毫米)
重量	33.1 千克 (73.0 磅)
功耗	最大值：0.7 kW
	常规值 ¹ ：0.5 kW
散热	最大值：2481 BTU/小时
	常规值 ¹ ：1736 BTU/小时
气流 ²	最大值：115 CFM
	常规值 ¹ ：80 CFM

工作温度/湿度：温度为 5 °C 至 32 °C (41 °F 至 89.6 °F)，相对湿度为 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

¹常规功耗随应用负载而不同

²气流方向必须从前往后

大数据一体机 X8-2 — 环境规格

物理尺寸（全机架）		
高度	42U，78.66 英寸（1998 毫米）	
宽度	23.62 英寸（600 毫米）	
深度	47.24 英寸（1200 毫米）	
重量	机架	配送
入门机架	415 千克（915 磅）	546 千克（1203 磅）
全机架	836 千克（1843 磅）	979 千克（2158 磅）
功耗		
入门机架	最大值：5.3 KW	
	常规值 ¹ ：3.7 KW	
全机架	最大值：13.4 KW	
	常规值 ¹ ：9.4 KW	
散热		
入门机架	最大值：18087 BTU/小时	
	常规值 ¹ ：12661 BTU/小时	
全机架	最大值：45659 BTU/小时	
	常规值 ¹ ：31961 BTU/小时	
气流 ²		
入门机架	最大值：837 CFM	
	常规值 ¹ ：586 CFM	
全机架	最大值：2114 CFM	
	常规值 ¹ ：1480 CFM	

大数据一体机 X8-2 — 其他环境规格

工作温度/湿度：温度为 5 °C 至 32 °C（41 °F 至 89.6 °F），相对湿度为 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

大数据一体机 X8-2 — 法规和认证

法规³

安全性：UL/CSA 60950-1、EN 60950-1、IEC 60950-1 CB Scheme（不同国家/地区之间存在差异）
RFI/EMI：EN55022、EN61000-3-11、EN61000-3-12
抗扰性：EN 55024
辐射和抗扰性：EN300 386

认证³

北美地区 (NRTL)、欧盟 (EU)、国际 CB Scheme、BSMI（中国台湾）、C-Tick（澳大利亚）、CCC（中华人民共和国）、MSIP（韩国）、CU EAC（关税联盟）、VCCI（日本）

欧盟指令³

2006/95/EC 低电压指令、2004/108/EC EMC 指令、2011/65/EU RoHS 指令、2012/19/EU WEEE 指令

¹ 常规功耗随应用负载而不同

² 气流方向必须从前往后

³ 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。其他国家和地区的法规/认证也可能适用。有些情况下，在组件级别达到了法规和认证合规性（如适用）。

大数据一体机 X8-2 — 扩展配置

机架内	多机架
<p>扩展：</p> <p>每个入门机架可利用最多 12 个 HC 节点以及 InfiniBand 基础架构实现现场升级。</p> <p>每个 HC 节点以及 InfiniBand 基础架构包含的额外硬件：</p> <ul style="list-style-type: none"> • 1 个节点，采用直连式存储（如前所述） • 用于连接组件的 InfiniBand 和以太网电缆 <p>扩展配置支持多代硬件。</p>	<p>无需额外的 InfiniBand 交换机即可连接多达 18 个机架。</p> <p>机架备件套件中提供的 InfiniBand 电缆可用于连接 3 个机架。</p> <p>如需连接 4 个或更多机架，应增加 InfiniBand 光缆。</p>
<p>内存</p> <ul style="list-style-type: none"> • 可将任意一个节点或任意数量节点中的内存从每节点 256 GB 扩展至每节点 1.5TB 	

大数据一体机 X8-2 — 支持服务

<p>硬件保修：保修 1 年，正常营业时间（周一至周五早 8 点至晚 5 点）4 小时内通过 Web/电话进行回复，2 个工作日内到达现场/调换部件</p> <p>Oracle 标准系统支持服务：24x7 的 Oracle Linux 和集成软件支持服务，2 小时内现场硬件维修响应（需靠近维修中心）</p> <p>Oracle 标准操作系统支持服务</p> <p>其他可用服务（需要额外收费或安装激活）：</p> <p>Oracle 客户数据与设备保留服务</p> <p>系统安装服务</p> <p>软件配置服务</p> <p>系统扩展支持服务，包括硬件安装和软件配置</p> <p>季度现场补丁部署服务</p> <p>Oracle 自动服务请求 (ASR)</p>



联系我们

有关 Oracle 大数据一体机 X8-2 的更多信息，请访问 oracle.com/cn 或致电 400-699-8888 联系 Oracle 代表。

关注我们



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。Cloudera、Cloudera CDH、Cloudera Manager、Cloudera BDR 和 Cloudera Navigator 是 Cloudera, Inc. 的注册或未注册商标 1219

 | Oracle is committed to developing practices and products that help protect the environment