

Vertica 高级分析平台

新一代大数据分析

产品介绍

处理当今海量的数据

在现代数据基础设施中，数据来自四面八方，例如 CRM 和 ERP 业务系统、用于收集机器生成数据的传感器、推文及其他社交媒体数据、Web 日志和数据流、天然气、电网和移动网络数据。由于这些数据来源庞杂，因此公司转而采用成本较低的不同存储位置来存储和管理这些海量数据，这样在对这些数据执行高级分析时，便会增加复杂性并形成孤岛。

为有效应对规模庞大、速度飞快、种类各异的数据，市场上涌现出一批功能新颖、极具创新性的大数据分析平台，旨在克服规模和复杂性空前的数据分析挑战，但其中有些平台苦于无法以统一方式处理所有数据，而另一些平台则仅提供了非常有限的基础设施选项。只有那些没有底层基础设施限制、能够涵盖所有数据的高级分析平台，才能满足绝大部分数据驱动型企业的需求。从数据变现到客户挽留，再从合规到流量优化，零售、医疗保健、通信、能源等诸多领域的企业在采用大数据分析平台后，正逐步改变各个行业的发展动向。

大数据分析平台有哪些主要的技术要求？

那么，您希望数据分析解决方案提供哪些功能特性来克服当下及未来的数据挑战呢？这里列出了以下几点：

以统一的方式分析海量数据：在分析数据时您可能希望不受规模限制，同时还需要在恰当的时机将数据存储于恰当的位置。今天的数据规模可能是 TB 或 PB 规模。但明天，您可能就要考虑如何处理 EB 规模的数据了。

满足业务预期：用户不希望被动等待结果。您的解决方案必须具有强大的可扩展性，才能满足服务级别协议 (SLA) 和业务需求，同时获得最佳的总体拥有成本 (TCO) 和快速的投资回报 (ROI)。

兼容主流工具：如果您的大数据分析依赖提取、转换、加载 (ETL) 工具或基于 SQL 的可视化技术，您的分析平台就必须提供稳定而强大的 SQL 工具，还要获得与您所有工具（而不仅仅是来自主要供应商的工具或仅限于单一基础设施平台的工具）配合使用所需的认证。

VERTICA 高级分析平台 – 无限制，不妥协

Vertica 高级分析平台的设计思路由堪称传奇的数据库专家 Michael Stonebraker 提出，而且从编写第一行代码开始便确定了打造一款大数据分析专用平台的目标。为什么要这么做呢？原因很简单，数据仓储、未经证实的开源解决方案和习惯性做法，都会给技术带来限制，致使企业不得不忍痛做出妥协。Vertica 高级分析平台在设计时便明确将速度、可扩展性、简便性和开放性确定为核心要求，而且使用了通过分布式压缩列架构来处理分析工作负载的构建方法。Vertica 高级分析平台具有速度极快（查询运行速度提高 50 到 1000 倍）、EB 规模（每台服务器多存储 10 到 30 倍的数据）、开放性和简便性（可使用任何商业智能 [BI]/ETL 工具、Hadoop 等）等特点，而且成本比传统数据仓储解决方案低得多，而上市速度也快于未经证实的开源解决方案。

让 VERTICA 变得如此强大的技术

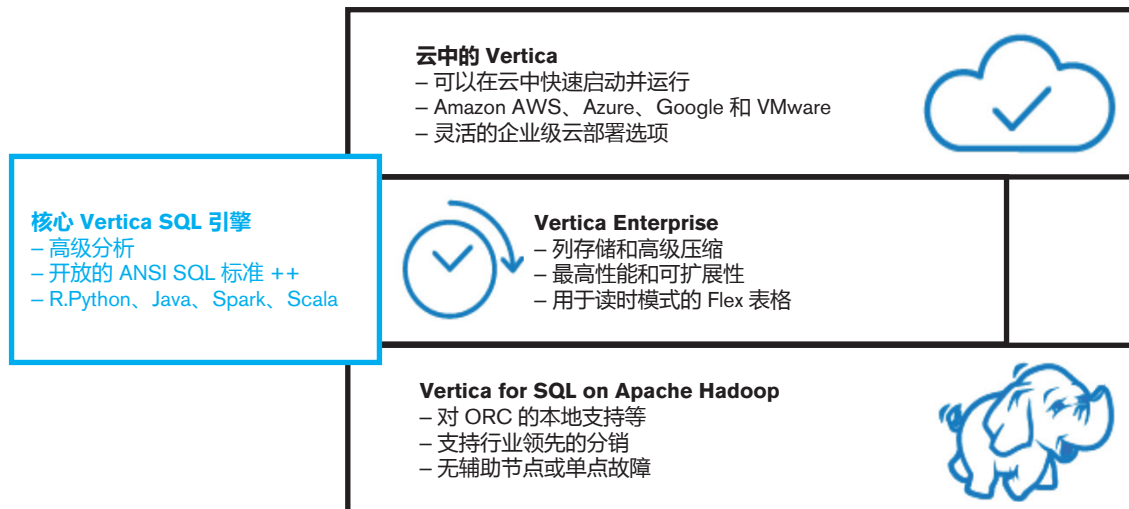
Vertica 是为了应对大数据分析挑战而全新构建的解决方案。凭借其大规模并行处理系统，它能够处理 EB 规模的数据，这一点已经在业内一些要求极为苛刻的使用环境中得到了印证。由于该方案采用列存储，而且对数据进行了大幅压缩，因而可以提供超快的大数据分析。与早些年已经过时的行存储技术相比，查询时间可从数小时缩短至数分钟，甚至从数分钟缩短至数秒。最后，Vertica 提供了十分先进的基于 SQL 的分析技术，包括图形分析、三角计数、蒙特卡罗模拟和地理空间等。这是一款功能全面的分析数据库系统。所有这些都应用于直接加载到 Vertica 的“热”数据，可用于要求极为苛刻的使用环境中，以及外部位置（如 HDFS 和 Amazon S3）中的数据，进而在无需移动数据的情况下，形成统一分析视图。

Vertica 的数据库内机器学习功能使用户可以充分利用大数据，同时简化并加快其预测分析过程，从而制定出更周全的决策，更有效地开展竞争，并缩短获取见解所需的时间。Vertica 支持从数据准备到部署的整个机器学习过程，并能够在 Vertica 群集中部署模型，这是将 Vertica 作为分析引擎解决方案的关键要求。

Vertica 的每个版本都通过了与可视化工具和 ETL 工具配合使用所需的认证和测试。它支持常用的 SQL 和 Java 数据库连接 (JDBC)/开放式数据库连接 (ODBC)。用户多年来对这些技术的投资和培训因此得以继续发挥作用，原因在于所有常用的 SQL 编程工具和语言都能与它无缝协作。领先的 BI 和可视化工具（例如 Tableau、MicroStrategy 等），乃至所有常用的 ETL 工具（例如 Informatica、Talend、Pentaho 等），都可与它紧密集成。

Vertica 高级分析平台的核心部分，是一个专门用来处理目前分析工作负载的列式关系数据库。与很久以前为支持小规模数据而设计的商业和开源行存储不同，Vertica 高级分析平台可为客户提供：

- 全面而又先进的基于 SQL 的分析功能，可执行强大的 SQL 分析
- 群集化的大数据存储，用于提供卓越的查询和分析性能
- 更出色的压缩功能，与同类数据分析解决方案相比，所需的硬件和存储空间更少
- 灵活性和可扩展能力，可随工作负载的增加轻松扩容
- 更高的负载吞吐量，以及更好的并行查询能力



- 通过 R 和 Python 及数据库内机器学习算法实现预测性分析
- 减少了数据库管理员 (DBA) 对日常管理和调整的干预

Vertica 提供高度可扩展性，适用于大规模的大数据分析。它采用了内存与磁盘平衡的分布式压缩列架构，这一独特设计使得它在运行当前数据分析工作负载时，速度较之传统技术，有指数级提升。

最广泛的部署和使用模式

Vertica 具有内部部署、Hadoop 上和云中等多种部署和使用模式，是统一的单一分析引擎，而且经实践证明，其大数据分析能力在速度和规模方面都无与伦比，永远不受底层基础设施的影响。

- Vertica 可以内部部署在您自己的数据中心、私有云或公有云中，包括 Amazon、Azure、Google 或 VMware® 云。操作简单，只需要使用 Vertica Enterprise 许可证，并在云中部署相同的统一引擎即可。如果需要更大的容量，但又没有时间在企业内部部署硬件，这将是一个颇具吸引力的方案。
- 相同的统一 Vertica 引擎可以本地部署在 Apache Hadoop® 上。Vertica SQL on Apache Hadoop® 在企业机构和机构首选的 Hadoop 分销版中以本地方式运行，可加快数据探索和 SQL 分析速度。

云中 – Vertica 软件经过优化和预配置，可在 Amazon、Azure、Google 和 VMware 云上运行。Vertica 为用户提供强大的敏捷性和可扩展性，可实现快速部署、自配置，还可与各种 BI 和 ETL 软件工具相集成。Vertica 具有很强的灵活性，可随着业务增长由小到大逐渐扩容，从而方便您将数据仓储过渡到云、过渡到内部部署，再无缝过渡回去。如此水平的敏捷性，自然不必做出妥协。

内部部署 – Vertica 高级分析平台是一个“完全无共享”的分布式数据库，可在经济高效的商用服务器群集上运行，而且只需向群集中添加新服务器即可扩展其性能。传统数据库需要含有大量 CPU 和 SAN 的“大型机”服务器，相比之下，Vertica 的网格架构大幅降低了硬件和扩展成本（可降低 70% 到 90%）。

群集化架构还有助于在群集中的节点上并行处理查询和加载，提供更高的吞吐量，从而提升性能。

在 Hadoop 上 – 与 Hadoop 配合使用时，Vertica for SQL on Apache Hadoop 会直接安装在 Hadoop 群集中，从而为组织提供一组强大的数据分析功能，且作用比前两种平台都要强大得多。因为它在查询时不依赖辅助节点，所以不会出现单点故障。它甚至可以读取 ORC、Parquet、Avro 和其他本地 Hadoop 文件格式并将其写入 Parquet。通过在 Hadoop 群集中安装 Vertica SQL 引擎，您可以充分挖掘全部 SQL on Hadoop 高级功能，在无需修改的情况下完成完整的 TPC-DS 查询，而且可在任意 Hadoop 分销版上运行。

将大数据变现并让大数据与每个人都息息相关

Vertica 高级分析平台使企业和机构能够透过浮华的表象，从大数据中提取出真正的价值。以下示例介绍了各组织如何通过 Vertica 来发挥自身数据这种颇具战略性的资产的价值：

- **Intuit** – 处理数十亿笔事务，为几百万 TurboTax 纳税申报用户提供高度个性化且快速的退税服务。
- **保护国际基金会** – 通过近乎实时地对 8600 万份样地和物种记录进行对比分析，帮助科学家评估气候、人类和土地使用的影响。
- **Cerner** – 通过分析临床医生在急诊室 (EMR) 的效率来改善为患者提供的医疗服务。截至目前为止，临床医生已利用败血症预警模型挽救了 500 名患者的生命。
- **Anritsu** – 将他们原有的 EDW 替换为 Vertica，作为新的嵌入式分析数据库，将当初只是理论的预测性分析解决方案付诸实施，获得了 351% 的投资回报，仅 4 个月就收回了投资。
- **Guess** – 通过移动设备提供每日店铺基本报告，从而精确跟踪销售情况，改善商品调拨和配送，并深入掌握顾客购买行为。

快来尝试一下，将您的想法变成现实。Vertica 下一代高性能 SQL 分析引擎以内部部署、云中和 Hadoop 上这三种集成产品的形式提供，可满足您的不同需求。您的需求是独一无二的，您的分析数据库也应该如此。立即试用 Vertica，网址：www.vertica.com/try

有关详细信息，请访问

www.vertica.com



扫描二维码
关注 Micro Focus 微信公众号

官方网站：<https://www.microfocus.com/zh-cn/>
售前电话：021-80383010
邮箱：china.sales@microfocus.com

有关其他信息，请访问：www.vertica.com

© 2017 Micro Focus. 保留所有权利。Micro Focus 和 Micro Focus 徽标是 Micro Focus 或其子公司或附属公司在英国、美国和其他国家/地区的商标或注册商标。所有其他标志均归其各自所有者所有。
