



重点:

- **严格分析** - 解答先前认为过于复杂、需要过多数据或者过长时间来分析的问题
 - **数据探索** - 发现不断增长的数据量中的微妙模式，以解答互相关联的复杂业务问题
 - **按需应变的分析** - 快速响应动态业务状况，以选择最佳操作方式
 - **易于使用** - 使用现有分析技术和技能集（包括开箱即用的功能、常用的分析软件包和语言）来构建模型和记分数据
 - **高性能** - 数据库内的并行算法针对组织的大量数据集快速执行，并充分利用 IBM Netezza 的数据仓库非对称大规模并行处理（Asymmetric Massively Parallel Processing™, AMPP）体系结构
-

IBM Netezza Analytics

IBM Netezza 的嵌入式数据库内分析平台

现今，企业面临着数据量不断增长的挑战，同时快速响应复杂且彼此关联的业务难题的需求也越来越高。高级分析可帮助企业制定更加明智的决策，并保持竞争力。IBM Netezza Analytics 是嵌入式的专用高级分析平台，它随各种 IBM Netezza 设备一起交付，支持分析型企业满足并超出各种业务需求。

IBM Netezza Analytics 的高级技术将数据仓储和数据库内分析融入到可扩展的、高性能、大规模并行嵌入式分析平台中，此平台旨在快速处理千兆级数据量。这使大量用户可以提出先前无法在其他体系结构上仔细考虑的各种数据问题。IBM Netezza Analytics 旨在快速有效地为最复杂的业务问题提供更快更好的答案。

IBM Netezza Analytics 是 IBM Netezza 最强大的高级分析平台，可提供技术基础架构以支持企业部署数据库内分析。此分析平台将含有先进分析工具（例如，R 和 Hadoop™）的强大内建分析集与 IBM Netezza 的核心数据仓库设备相集成。IBM Netezza 是现代化数据仓库设备的引领者，客户遍及全球，这些客户已实现了将数据仓储与分析结合到单一的高性能集成系统中的价值。IBM Netezza Analytics 允许分析型企业通过新的业务模型实现显著的业务价值，并帮助各公司最大限度地提升收入，同时将成本降至最低。

IBM Netezza 可提供与众不同且便于使用的严格分析方法。传统的分析构建并部署于独立的分析服务器上。这延长了从模型先启到部署的端到端时间。同时，这需要将数据从某个数据仓库或其他数据源移动到分析服务器。该过程不仅耗时，而且效率低下，限制了使用的数据量，约束了分析建模的范围，并且阻碍了以迭代方式进行实验的能力。



有了 IBM Netezza Analytics, 分析模型可直接在数据所在位置(即数据仓库中)进行构建和部署。由此可显著缩短在整个企业内构建并部署分析模型所需的时间。通过缩短从模型先启到部署的时间, 各公司可以将更多决策与富于洞察力的按需应变的分析相结合, 从而制定出企业级基于事实的决策。

每个 IBM Netezza 数据仓库都随附有数据库内分析功能库, 可通过任何 SQL 兼容的界面进行访问。此外, 客户可以使用该平台的用户定义扩展来开发新的功能。它是最方便、可扩展性最高的平台, 支持多种工具、语言和框架。

使用 IBM Netezza Analytics, 您可以进行:

- 数据探索和发现
- 数据变换
- 模型构建
- 模型诊断
- 模型计分

IBM Netezza 数据仓库设备是强大的并行计算平台, IBM Netezza Analytics 充分利用该平台来交付高速且可扩展的分析处理。该设备使用非对称大规模并行处理 (AMPP) 体系结构的高速吞吐量, 以最大化数据库内分析处理的速度和效率。AMPP 体系结构是基于刀片的流式体系结构, 使用商品级刀片和存储器, 并结合 IBM Netezza 获得专利的数据过滤(使用现场可编程网关阵列 (FPGA)) 来交付针对大量数据的高速分析。IBM Netezza 已将所有分析活动整合到一个强大简单的设备中。

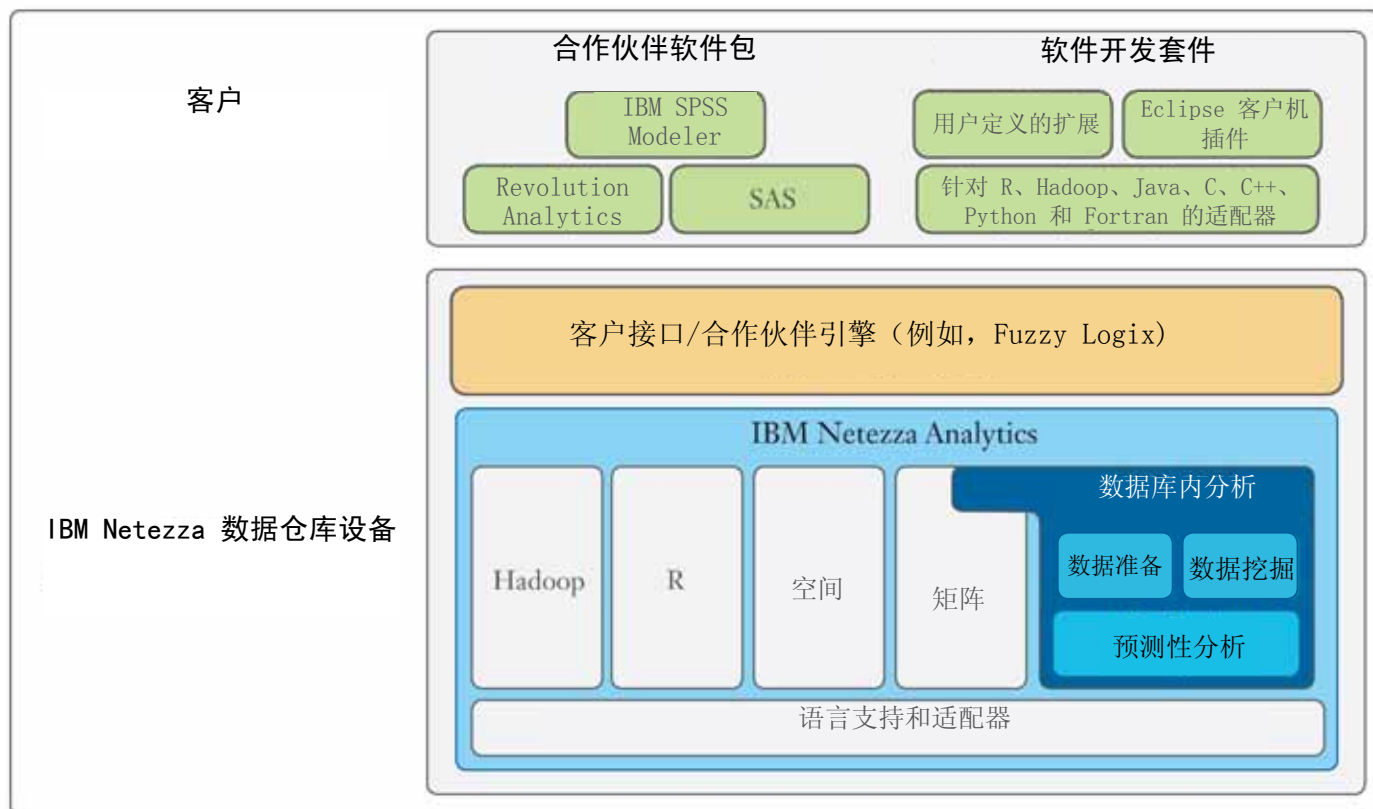
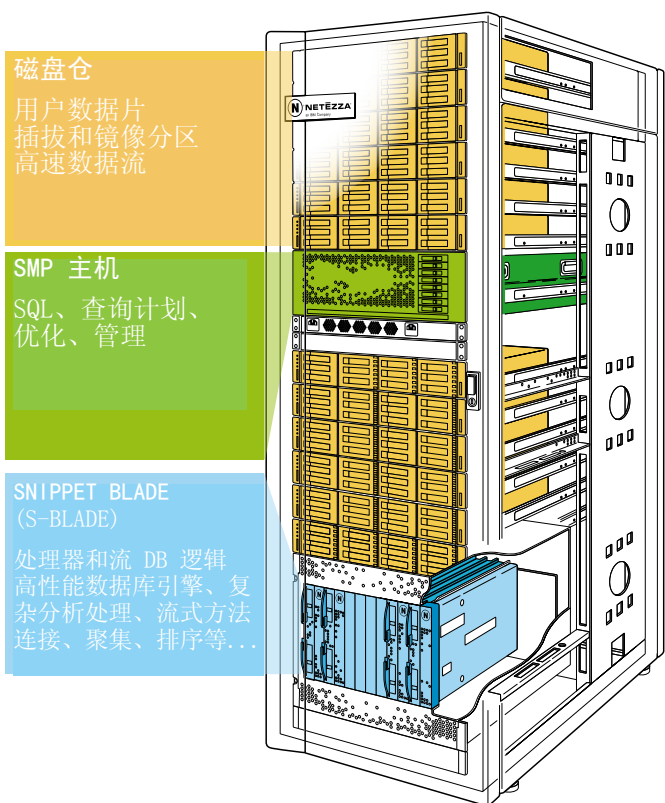


图 1: IBM Netezza Analytics 体系架构



IBM Netezza Analytics 专用于简化模型构建和部署，以满足分析型企业对大量复杂数据的最高性能需求。

严格分析

企业收集和跟踪的信息量之大前所未有，并且以更高效且有效的方式开展运营的压力与日俱增。分析数据、预测结果并搜寻方式以改善业务的能力推动着各公司充分利用高级分析。理清大量数据并将其转化为有意义的成果对于使用传统数据库技术的公司而言着实有些困难，甚至从技术角度而言不可能实现。这些系统仅仅为了维持用户和数据量的增长速度已捉襟见肘。

曾经看似不可能或者不切实际的分析现在可能通过 IBM Netezza Analytics 得以实现。利用 IBM Netezza 的简单设备方法，组织的所有数据都可用于生成更细致的结果集，从而帮助推动新的收入商机并获取竞争优势。通过使用 IBM Netezza 数据仓库设备上的高级分析，整个企业均可实现从财务团队到业务线、销售、IT 及主管办公室的价值。这可为整个企业提供更高的清晰度，以确保所有人员都在利用相同的数据并使用所有可用的数据。

使用 IBM Netezza Analytics，各组织不再需要在大数据量和严格分析之间做出抉择。

数据探索

利用 IBM Netezza Analytics，可以超越基于历史记录数据的传统商业智能 (BI) 和即席报告，并开辟从数据中挖掘价值的新方法。各组织现在能够构建并部署复杂的分析模型，从而更轻松有效地映射出现实世界的复杂性。他们可以连续实验、改善并调整分析模型，以发现趋势和找到降低业务风险、降低成本、增加收入的途径，同时制定基于事实的决策。

利用 IBM Netezza Analytics 来研究从数据中挖掘价值的新方法，同时执行并发的并行模型实验。现在，所有数据均可利用，而不仅仅是样本或聚集，从而可以提高准确性并指定更具针对性的决策。

电信业案例研究

在竞争白热化的电信业，参与者需要准确且最新的信息来管理其业务，利用新的商机并将顾客流失率保持在最低。

- 收入保障部门使用分析来识别收入漏洞，并填补收入链中的任何缺失。
- 产品营销和定价团队可主动计划并预测电话关税更改的效果，并计划如何对竞争对手的产品做出最有效的反应。
- 信贷服务部门使用分析来确定电话使用高峰，并在任何潜在的信贷问题出现前对其加以主动管理。

IBM Netezza 数据仓库设备可快速处理大数据量并处理数据库内分析，以帮助电信公司维持竞争力并精简流程。

医疗卫生业案例研究

某医疗卫生供应商对于预测糖尿病易感人群十分感兴趣。通过查看传统医疗卫生参数（例如体重和家族病史），并将更多属性（例如，财务背景）添加到模型中，该供应商发现个人财务状况确实影响着易感人群的发病风险。通过优化其分析模型，该医疗卫生供应商不仅可以确定某人是否会在生命期内某特定时间患上糖尿病，并且可预测该疾病的发病时间（一年后、三年后等）。

通过识别这些趋势，可以向这些易感患者提供最出色的预防性护理。该医疗卫生供应商利用 IBM Netezza Analytics 持续优化并创建模型，以了解其他趋势并加强患者护理。

按需应变的分析

各组织可通过使用 IBM Netezza 来预测并优化业务元素，从而超越竞争对手。过去，分析过程成本高昂且耗时，通常需要数周来从数据仓库中的数据开发预测模型。一旦开发完模型，仍需要数小时（有些情况下甚至数天）来执行所有数据，不仅如此，还可能需要在解决问题而添加昂贵的硬件。随着数据量的增长，问题日趋严重。

IBM Netezza 可为公司重大的预测计划提供快速实现价值的力量，从而对收入和利润的增长产生积极影响。通过 IBM Netezza 的加入，各组织可具备最准确的情报，以快速、自信地应对市场可能带来的任何商机或威胁。可根据需要快速部署并调整模型，并成倍地并行分发，同时还可利用 IBM Netezza 的并行数据库内技术。这样可以快速有效地处理最大数据量。

在各公司需要尽可能保持敏捷以应对不断变化的市场条件和需求的时代，可高速运行并分析 PB 级数据量的简单易用系统非常实用。

易于使用

IBM Netezza 的数据仓库设备易于使用，并可显著加速整个分析过程。编程界面和并行选项可以将大部分分析直接移到设备内，不论这些分析是使用工具（例如，IBM SPSS、SAS 或 R）执行还是以各种语言（例如，Java™、Python™ 或 Fortran）编写。此外，IBM Netezza 数据仓库设备随附有内建并行分析功能库，这是专为大数据量构建的，可快速启动并加速任何分析应用程序的开发和部署。

开发的简单性和易用性真正使 IBM Netezza 脱颖而出。该设备史无前例地将数百个处理核心的功能和扩展性囊括在一个体系结构中，从而使其成为并行分析的理想之选。IBM Netezza Analytics 将所有分析活动整合到强大的设备中，它并非带有复制数据所用多个系统的分段式分析基础架构。该设备便于部署，只需最低限度的持续管理，即可实现较低的总体拥有成本。

简化数据探索、计算、建模和计分过程，从而促进分析在整个公司范围内的成功采用。利用 IBM Netezza，业务用户可近乎实时地运行各自的分析，从而帮助由分析支持、数据驱动的决策遍布整个企业。

零售业案例研究

某基于行为的市场营销解决方案全球领袖企业向制造商和零售商提供了基于实时市场购物篮分析和历史采购模式来快速执行目标市场营销计划的能力，从而见证了将 IBM Netezza 引入业务的显著成果。例如，在某个依靠 POS 信息来向客户发放个性化优惠券的竞销活动中，使用 IBM Netezza 使优惠券兑现率增加了 30%。通过利用 IBM Netezza 的简便性，该公司减少了维护其数据仓库环境所需的 DBA 数量，从而显著提高了生产力，并且 IT 分析项目完成速度提升了 5 - 10 倍。

此外，由于消除了所有聚集表和索引，该公司主要的数据库存储空间自迁移至 IBM Netezza 后减少了近 80 TB。存储容量的减少同样减少了对应的数据中心占地面积。

数字媒体业案例研究

某家向客户提供详细分析业务的数字媒体公司可在适当的时间提供数据，以推动具有洞察力的决策制定，正确度量市场动态，并有效地构建零售商和制造商之间的桥梁。通过利用 IBM Netezza Analytics 和 IBM Netezza 数据仓库设备，该公司现在可为比以往更多的客户提供服务，并且这些客户可以运行定制定义的即席市场分析，但原先他们仅限于查看市场的静态视图。同时，可检索并分析的数据历史记录比以往更多，新的数据可近乎实时地供客户使用。分析不再受到制约，并且功能更强、灵活性更高。

IBM Netezza 帮助降低了 IT 成本，由此创建更有效的业务模型。利用其高级分析功能，该公司取得了超越其竞争对手的竞争优势。

专注于业务而不是过程。允许您的 IBM Netezza 数据仓库设备为您挑起重任。

高性能

IBM Netezza 已创建了极端灵活的分析平台，以提供千兆级的高性能。通过将分析引入数据，建模人员和定量分析团队可以直接在设备内部对数据进行操作，而无需将数据转移至其他位置并处理相关联的数据预处理和变换。

分析人员和建模人员可充分利用 IBM Netezza Analytics 的 AMPP 体系结构，在所有企业数据上询问最复杂的问题，从而摆脱基础架构的阻碍。实践人员可更快速地在不同模型间进行迭代，以实验并查找最佳匹配。

一旦模型开发完成，即可针对企业内所有相关数据无缝地执行。预测和记分可直接在数据驻留位置完成。用户可以近实时地获取预测分数结果，以帮助实施高级分析，并将其供给整个企业使用。

IBM Netezza 数据仓库设备使用现场可编程网关阵列 (FPGA)，它是由 IBM Netezza 专为高效处理大数据量而编程的。这些 FPGA 可在数据流出磁盘时即过滤掉不相关的数据。这样可消除 I/O 瓶颈并使下游组件 (例如 CPU、内存和网络) 免于处理不必要的数据，从而大幅提高系统性能。

在强大的多核 CPU 中执行严格的大量计算，其中，数据库基本和复杂分析在经过滤的数据流上执行。分析任务作为数据流上运行的独立过程在每个 S-Blade 间运行。IBM Netezza 的并行分析引擎可利用设备内所有计算核心的能力来提供强大的性能和可扩展性，从而进行严格分析。

通过利用 IBM Netezza 的设备来消除技术阻碍，从而使生活更简单。

利用 IBM Netezza Analytics，您将具备可管理大数据量上所有分析查询的设备，同时可利用 IBM Netezza 的数据仓库设备并行处理平台来提高业绩。IBM Netezza 可为您提供简单的设备，以执行严格的分析。

金融服务业案例研究

某金融机构需要计算某股票期权的风险值。IBM Netezza 平台能够在不到三分钟内对 1000 个标的股票的 20 万种状况执行 Monte Carlo 模拟 (25 亿次模拟)。通过利用数据库内分析方法，该金融机构得以在数据驻留位置使用数据，而不是为执行 Monte Carlo 模拟单独构建并行数据处理平台。

整合执行时间，同时消除在两个平台之间移动数据所需的等待时间，从而使该金融机构能够在评估投资策略的风险时包含更多变量，并更加频繁地执行此评估。

IBM Netezza Analytics 平台

核心模块:

IBM SPSS® In-Database Analytics	执行数据准备、数据挖掘、预测分析
R	运行并行的数据库内模型构建和记分
矩阵	利用并行的线性代数软件包
Hadoop	在数据库内执行 MapReduce™ 功能
空间	使用 Geospatial 数据类型和功能

软件开发套件 (SDK) 包括:

语言适配器	使用 R、Hadoop、Java、C、C++、Python 和 Fortran 开发
Eclipse 的插件	利用便于使用的、标准的集成开发环境来构建数据库内的分析
用户定义的扩展	创建定制的用户定义的函数 (UDF)、用户定义的聚集 (UDA)、用户定义的表函数 (UDTF) 和用户定义的分析过程 (UDAP)

第三方应用程序

数据集成	Ab Initio®、BusinessObjects™/SAP®、Composite® Software、DataFlux™ – 一家 SAS® 公司、Expressor® Software、GoldenGate® Software、Informatica®、IBM Information Server、Oracle® Sunopsis 和 WisdomForce™
数据分析	IBM SPSS、Revolution® Analytics、BusinessObjects/SAP、Kalido®、KXEN®、Quest® Software 和 SAS
BI/报表	IBM Cognos、IBM Unica®、Acutate™、BusinessObjects/SAP、Information Builders、MicroStrategy®、Oracle 和 QlikTech®
数据可视化	TIBCO® Spotfire® 和 BIS(2)™
数据库内分析	SAS 和 Fuzzy Logix™ / DB Lytix™
业务连续性/合规性	IBM DataMirror、IBM Tivoli Storage Manager、EMC® 和 Symantec™ Veritas™

分析功能

内建分析库随每个 IBM Netezza 数据仓库设备提供 (样本)

数据准备、验证和操作

存储过程

表和视图

数据挖掘

数据评估

分布

信息内容

数据变换

采样、培训/测试集

离散化

主成份组件

模型构建、解释和应用

项目集 (关联的基础)

集群

分类

回归

部署 (PMML 导出)

统计信息

ANOVA

关联、协方差、分散

分布、密度、概率

缺失的值

时机, 集中趋势

规范化, 标准化

分位数

重要性

矩阵函数

常规

矩阵代数、操作和描述

矩阵表操作

矩阵创建和变换

矩阵统计信息

空间函数

常规

几何, 单一

几何, 多重

关于 IBM 数据仓储和分析解决方案

IBM 提供最广泛且最全面的数据仓储、信息管理和业务分析软件、硬件和解决方案的组合, 以帮助最大程度发挥信息资产的价值, 并发现新的洞察力, 以更好更快地制定决策, 优化业务成果。所有 IBM 的数据仓储和分析解决方案均旨在简化并加速交付业务分析洞察力。

IBM 的产品服务组合所包含的数据仓库设备将数据库、服务器和存储器集成到易于管理的单一设备中, 该设备只需最低限度的设置和日常管理, 并提供更快且更一致的分析性能。IBM 还提供预先构建、预先集成且针对工作负载优化的数据仓储和分析平台, 以及数据仓库软件, 以实现运营智能化。这些产品通过针对大数据和新型的分析工作负载的额外支持得到了增强, 包括对大量传输中数据进行连续且快速的分析。

了解更多信息

要了解有关 IBM 数据仓储和分析解决方案的更多信息, 请联系 IBM 销售代表或者访问 :

ibm.com/software/data/infosphere/data-warehousing/



国际商业机器中国有限公司
北京市朝阳区工体北路甲 2 号
盈科中心 IBM 大厦 25 层
邮编:100027

IBM 首页位于:

ibm.com

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corporation 在美国和/或其他国家或地区的商标。如果这些名称和其他 IBM 已注册为商标的名称在本信息中首次出现时使用商标符号 (® 或 ™) 加以标记, 那么表明这些符号在本信息发布时已经是由 IBM 根据美国联邦法律注册或根据普通法注册的商标。这些商标也可能是在其他国家或地区的注册商标或普通法商标。以下 Web 站点上的“Copyright and trademark information”部分包含了 IBM 商标的最新列表:

ibm.com/legal/copytrade.shtml

Netezza 是 Netezza Corporation (一家 IBM 公司) 的注册商标。

其他产品、公司或服务名称可能是其他公司的商标或服务标记。

© Copyright IBM Corporation 2011
All Rights Reserved.



请回收再利用