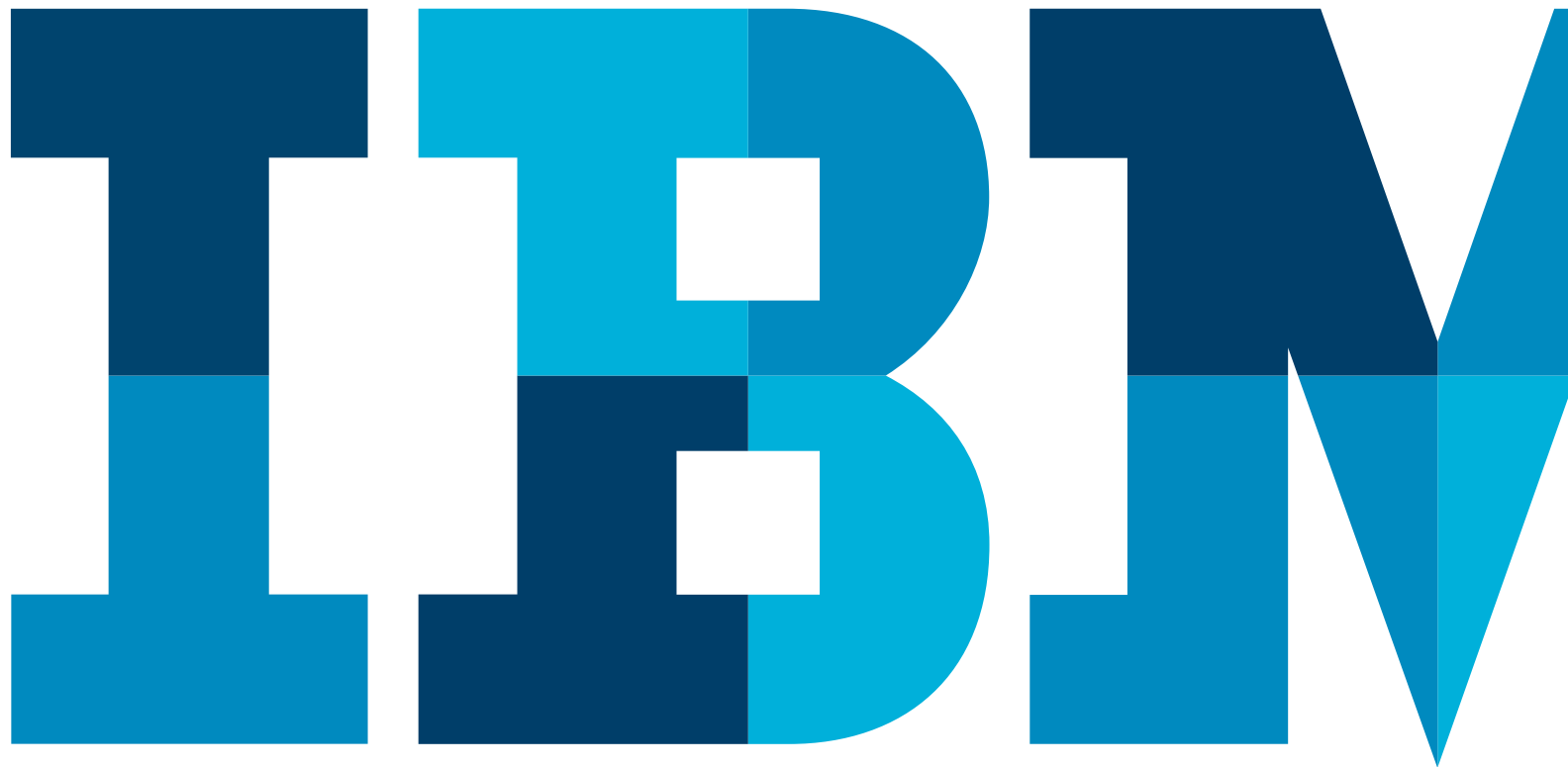


# Oracle Exadata 和 IBM Netezza 数据仓库设备比较

由 IBM 产品管理和产品市场营销副总裁 Phil Francisco 和 IBM 产品  
市场营销部资深总监 Mike Kearney 共同编撰



## 目录

简介	2
联机事务处理 (OLTP) 和数据仓储	3
查询性能	5
操作简易性	10
价值	14
结论	16

## 简介

IBM Netezza 数据仓库设备关注于查询和分析大数据的技术。IBM Netezza 数据仓库设备打破了市场的格局。为了以更低的运营和拥有成本对数据加以利用，许多客户从 Oracle 转为使用我们的数据仓库。Oracle 现已在市场上推出 Exadata，该机器似乎可以实现 IBM Netezza 数据仓库设备的所有功能，而且还可以处理联机事务。坦率的说，此次对 Exadata 和 IBM Netezza 作为数据仓库平台的检验是从 IBM 角度编写的，不过为确保可信性，我们采纳了 Bloor Research 的研究总监 Philip Howard 和 Monash Research 的总裁 Curt Monash 的建议。

为进行创新，要求我们以不同方式思考和操作，使用新方法来解决。IBM Netezza 数据仓库设备为我们客户的仓库查询提供卓越性能。IBM Netezza 数据仓库设备为客户提供简易性；任何具有基本 SQL 和 Linux 知识的人员都能胜任必须的管理任务以保证在动态变动的工作负载下提供始终如一的服务水准。IBM Netezza 数据仓库设备的性能及简易性降低了拥有和运行其数据仓库的成本。更重要的是，我们的客户可以通过部署分析应用程序来创造新的业务价值，这是他们以前可望而不可及的事情。

“Netezza 是 Exadata 灵感的一部分。Teradata 也是 Exadata 灵感的一部分。我们要感谢他们迫使我们开拓硬件业务。”

-Larry Ellison, 2010 年 1 月

“Netezza 是 Exadata 灵感的一部分。Teradata 也是 Exadata 灵感的一部分”，Larry Ellison 在 2010 年 1 月 27 日如是承认。“我们要感谢他们迫使我们开拓硬件业务。”<sup>1</sup> 尽管带着 Larry Ellison 惯有的潇洒，但在其评论中存在一个要点：Oracle 只对最佳产品才会予以关注。Exadata 代表 Oracle 的战略方向；调整其 OLTP 数据库管理系统，使其与来自 Sun 的大规模并行存储系统相结合。Oracle 推出了 Exadata V2，并且承诺在处理联机事务和分析查询方面实现极致的性能。因此，Oracle Exadata V2 是一种用于管理混合工作负载的通用平台。Oracle 数据库是为 OLTP 而设计的。但是与 OLTP 相比，数据仓储和分析对于软件和

<sup>1</sup> 请参阅 [http://oracle.com.edgesuite.net/ivt/4000/8104/9238/12652/lobby\\_external\\_flash\\_clean\\_480x360/default.htm](http://oracle.com.edgesuite.net/ivt/4000/8104/9238/12652/lobby_external_flash_clean_480x360/default.htm)

硬件的需求大不相同。简言之，数据仓储的一些工作负载在专为分析构建的系统上执行时性能更佳而且更为经济有效。我们还是要好好审视一番 Exadata 在数据仓储领域的性能，尤其是在简易性和价值方面。

本电子书开篇回顾了数据仓库中处理联机事务与处理查询和分析之间的差异。然后，从查询性能、操作简易性和价值方面对 Exadata 和 IBM Netezza 数据仓库设备进行了讨论。

我们仅仅希望读者如同我们的客户和合作伙伴一样，撇开对数据库管理系统工作方式的固有看法，对于新的思维方式保持开放性的态度，并且准备好事半功倍地实现更好的结果。

要说明一点：IBM Netezza 数据仓库设备团队无法直接访问 Exadata 机器。幸运的是，我们收到诸多对两种技术均已进行评估并选定 IBM Netezza 数据仓库设备的组织的详细反馈。鉴于 Oracle 的规模及其对 Exadata 的重视，有关 Exadata 的公开可获取的信息出乎意料的稀缺。Oracle 所引用的用例基本不提供可供讨论的信息，这本身引起不少行业关注者（例如，Information Week）的担心。<sup>2</sup>本文档中共享的信息本着开放性的精神提供。由于我们的错误产生的任何不准确的地方并非意在误导。

## 联机事务处理 (OLTP) 和数据仓储

OLTP 系统执行许多简短事务。每个事务的范围较小，限于一条或少数记录，并且由于这些事务的高度可预测性数据经常是缓存的。尽管 OLTP 系统处理大量数据查询，但其重点在于对当前数据集进行写操作

(UPDATE、INSERT 和 DELETE)。这些系统通常特定于业务流程或功能，例如管理支票帐户的当前余额。其数据通常按第三范式 (3NF) 构造。OLTP 系统的事务类型比较稳定，其数据需求也比较简单，因此索引之类的辅助数据结构可以先在磁盘上找到记录，然后将其传输到内存进行处理。

相比之下，数据仓库系统的典型特征表现为针对当前和历史数据集的明显繁重的数据库读 (SELECT) 操作。OLTP 操作只访问少数记录，而数据仓库查询可能会扫描包含数十亿行的表格，并将其记录与其他多个表中的行连接。此外，数据仓库中的查询通常在本质上不可预测，致使难以利用高速缓存和索引策略。用于在仓库中构造数据的选项比较多，范围从 3NF 到诸如星型和雪花模式之类的多维模型。每个系统中存储在典型仓库中的数据都构造为反映特定业务流程的需求。在将数据装入仓库之前，会对其进行清理、去重和集成。

<sup>2</sup> 请参阅 [http://www.informationweek.com/news/business\\_intelligence/warehouses/showArticle.jhtml?articleID=225702836&cid=RSSfeed\\_IWK\\_News](http://www.informationweek.com/news/business_intelligence/warehouses/showArticle.jhtml?articleID=225702836&cid=RSSfeed_IWK_News)

本电子书将数据仓库划分为第一或第二代。尽管此分类并不是最严格的定义，但从一代到二代的进化反映出我们的客户是怎样在数据中实现越来越大价值的。

第一代数据仓库通常在夜间装入数据。它们通过缓慢演变的基于 SQL 的报告和仪表板的稳定主体，向其业务部门提供信息。由于这些简单的数据仓库有点类似 OLTP 系统（其工作负载和数据需求明确而稳定），因此组织往往会采用与 OLTP 相同的数据库管理产品。随着产品进入实践阶段，数据库管理员会分析每份报告的数据需求，并构建索引以加快数据检索。在仓库中的数据量超过事务系统中通常管理的数据量之前，OLTP 技术和方法的演变看似成功。

在本世纪内，企业和公共部门机构每年 30%-50% 的数据增长率为正常情况。OLTP 领域的技术和实践已经证明越来越不适用于数据仓储；使用索引作为数据检索辅助便是一个例证。当数据库系统处理作业以装入数据时，它还要忙于更新多个索引。鉴于数据量巨大，此过程变得非常缓慢，导致装入作业超过其分配的处理时间范围。尽管长时间工作，但是技术团队仍无法满足业务部门所希望达成的服务级别。当业务单位等待报告和数据可用时，生产力就受到影响。

---

**OLTP 领域的技术和实践已经证明越来越不适用于数据仓储...**

---

组织正在提出新的关于如何利用数据的需求与愿望；本电子书将此发展称为第二代数据仓库。这种新一代数据仓库可以轻松管理大规模数据集并充当企业存储器。在查询时，它们会重新调用前几年记录的事件；这些久远的记忆提高了预测分析应用程序的准确性。细水长流式的数据供应正在取代夜间批量装入，从而减少事件记录及其分析之间的延迟时间。除用于填充报告和仪表板的简单 SQL 外，仓库还会处理高级分析的线性回归、朴素贝叶斯及其他数学算法。注意到五家店铺内的高利润产品销售突然激增，促使某家零售商了解发生的情况及原因。此零售商通过了解此情况，制定策略，在全部 150 家店铺都开展类似的促销活动。支撑该数据仓库的计算系统必须能够管理这些突如其来的需求激增，而不会扰乱定期报告和仪表盘。企业用户需要有选择利用数据的时间和方式的自由度。那些需要管理员不断调优的传统技术已无法满足用户对于即时分析的需求。

## 查询性能

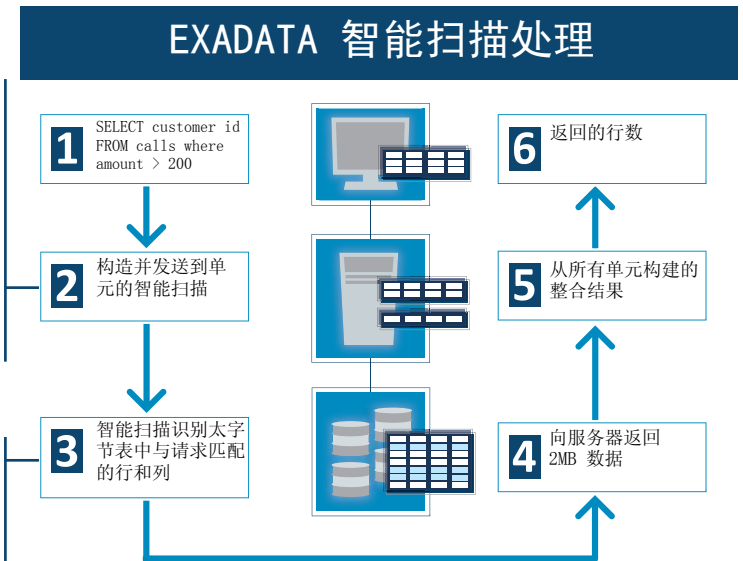
### Oracle Exadata 的查询性能

在并购 Sun 之时，Oracle 得出了 IBM Netezza 数据仓库设备团队在十年前便已得出的结论：当所有组成部件（软件和硬件）都根据其目标进行优化时，数据仓库系统便可实现最高效率。Exadata 通过由快速网络连接的两个子系统创建而成：一个智能存储系统，通过 InfiniBand 与包含 Real Application Clusters (RAC) 的 Oracle Database 11g V2 进行通信。单个机架系统在大规模并行处理 (MPP) 网格中包含由 14 台存储服务器（称为 Exadata 单元）组成的存储层，配以作为 8 个对称多处理节点的共享磁盘集群运行的 Oracle RAC 数据库。

在并购 Sun 之，Oracle 得出了 IBM Netezza 数据仓库设备团队在十年前便已得出的结论：当所有组成部件（软件和硬件）都根据其目标进行优化时，数据仓库系统便可实现最高效率。

**智能扫描限制**  
智能扫描并不全面；Exadata 的 MPP 存储层无法处理：索引组织表 - Oracle 推荐用于文本、图像、音频和空间数据；包含活动事务 (INSERT、UPDATE 和 DELETE) 的数据块；SQL 跨多个表连接或跨两个表复杂连接；Oracle 的 511 个数据库函数中的 192 个；用户定义的函数；或不同的聚集（在简单报告中较为普遍）。

**未充分使用的 MPP**  
Exadata 的设计未充分利用 MPP 体系结构。数据管理未完全集成到存储层中，意味着对其 MPP 网格中的硬件基本没要求。



Oracle 将 Exadata 的存储层标榜为智能存储层，因为它能够首先处理 SQL 投影、限制和连接过滤，<sup>3</sup>然后再将生成的数据集放在网络上，由 Oracle RAC 执行下游处理。此技术被称为智能扫描。不过，智能扫描并不全面；存储层不会处理所有限制。Oracle 的在线论坛<sup>4</sup>列出了许多操作，包括扫描索引组织表或集群表，都未能从智能扫描中获益。除此之外，《Troubleshooting Oracle Performance》一书的作者 Christian Antognini 还撰写了一篇博客，建议不要将智能扫描用于 TIMESTAMP 数据类型。<sup>5</sup>Oracle 建议将数据仓库中的事实表实现为索引表，以有效执行星型查询。<sup>6</sup> Exadata 的存储层不会处理对索引组织表的限制，但必须将所有记录传递到下游 Oracle 数据库。Exadata 的将完整记录从存储层传递到数据库层的方法对于 OLTP 非常有效率，因为各事务必须仅检索少数几行。不过，要求扫描长（几亿或几十亿行）而宽（几百列）的事实表的统计分析将会生成一大波数据，致使数据在网络中低效地移动。如果 Exadata 能够在其 MPP 存储层中处理所有 SQL 谓词（WHERE 子句），那么可以实现更理想的性能和效率。

Exadata 存储服务器无法相互通信；所有通信都会通过 InfiniBand 网络强制传递到 Oracle RAC，然后通过此网络返回存储层。该体系结构对于联机事务处理是有利的；其中每个事务（范围涵盖一条或少数几条记录）可通过将小数据集从存储层移至数据库层来得到满足。分析查询（如“查找华盛顿州、俄勒冈州和加利福尼亚州上个月出售的包含产品 X 和产品 Y 且总价超过 35 美元的所有购物篮”）必须检索大得多的数据集，所有数据集都必须从存储层移至数据库层。这种低效的大量数据移动会对查询性能产生不利影响。

Exadata 的存储层还表现出其他缺点。Exadata 单元无法处理不同的聚集，这甚至在简单报告中也是常事；它们无法处理分析应用程序中所使用的复杂连接或分析功能。由于无法在其存储层中解析这些典型的数据仓储查询，因此 Exadata 必须通过其内部网络将超大数据集推送至 Oracle RAC。该体系结构缺陷引起了对于 Exadata 在用于必须运行复杂分析查询的第二代数据仓库的稳定性的质疑。

Oracle 希望使用其 40 Gb/秒交换机 InfiniBand 超越 IBM Netezza 数据仓库设备；事实上，正是由于系统的不平衡和低效，Exadata 才需要这种昂贵的网络。Exadata 存储服务器执行的工作量极少，造成在网络上向下游移动以由 Oracle RAC 进行处理的数据超出所需，这就要求执行过多的工作。

Exadata 在其数据库层运行具有 Real Application Clusters 的 Oracle 11g V2，将其作为集群式共享磁盘体系结构。将该体系结构用于数据仓库平台让人们不得不担心，对于共享资源的争夺会限制数据库可处理的数据量及其可并发运行的查询数量。时间和客户体验会证明这一担忧是有道理的。

3 A Technical Overview of the Sun Oracle Exadata Storage Server and Database Machine - Oracle 白皮书，2009 年 10 月。

4 <http://forums.oracle.com/forums/thread.jspa?threadID=1036774&start=0>  
完整列表为：索引组织表或集群表扫描；索引范围扫描；压缩索引访问；反向键索引访问；安全企业搜索

5 请参阅 Christian Antognini 的博客，地址为  
<http://antognini.ch/2010/05/exadata-storage-server-and-the-query-optimizer-%E2%80%93-part-2/>

6 [http://www.oracle.com/technology/products/oracle9i/datasheets/iots/iot\\_ds.html](http://www.oracle.com/technology/products/oracle9i/datasheets/iots/iot_ds.html)

Exadata 存储层中的每个磁盘都由运行 Oracle RAC 的网格中的所有节点共享。这种共用存储器会造成某个页面由一个节点读取时正由另一个节点更新的风险。为解决此问题，Oracle 强制在节点之间进行协调。每个节点会检查其同级的磁盘活动以防止冲突。Oracle 技术人员将此活动称为块 ping。计算周期在每个节点检查其同级的磁盘活动时被无端的消耗，或者它们在一个节点处于空闲状态以等待另一个节点完成操作时被丢失，这些都造成了浪费。在专门设计用于数据仓储的体系结构中，这些周期本应该用来处理查询、挖掘数据和运行分析。

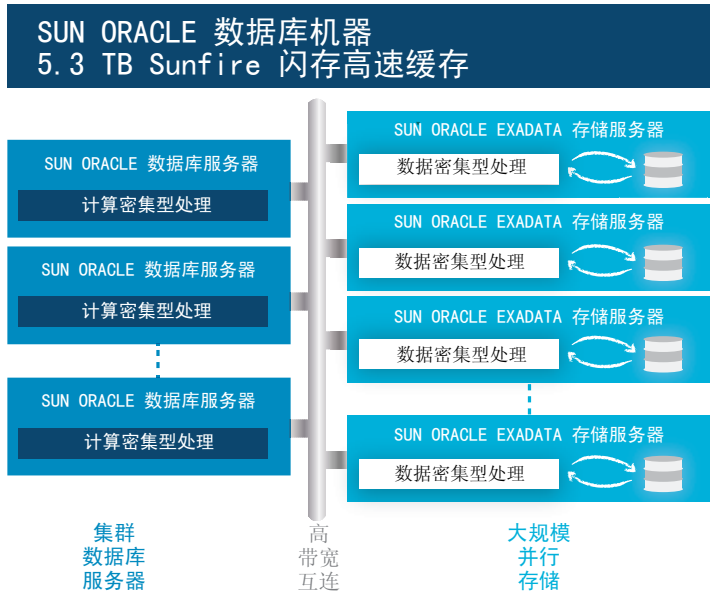
除简单查询外，Exadata 必须将大的数据集从其存储层移至其数据库层，由此引起对其作为现代数据仓库平台的稳定性的质疑。

**共享磁盘体系结构**

Exadata 的数据库层将具有 RAC 的 Oracle11g V2 作为集群式共享磁盘体系结构运行，从而限制数据库处理的数据量及其并行运行的查询数量。

**高网络流量**

Exadata 存储服务器通过 InfiniBand 网络相互通信，以传递到 Oracle RAC 并跨该网络回传到存储层。仅会向数据库服务器返回未压缩的数据，从而显著增加网络流量。



**性能瓶颈**

Exadata 的存储层不会处理对索引组织表的限制。所有此类记录都会装入到数据库服务器中进行处理。

**管理开销**

管理员通过分区、文件、表空间和块/扩展数据块大小来设计和定义数据分发。

**分析限制**

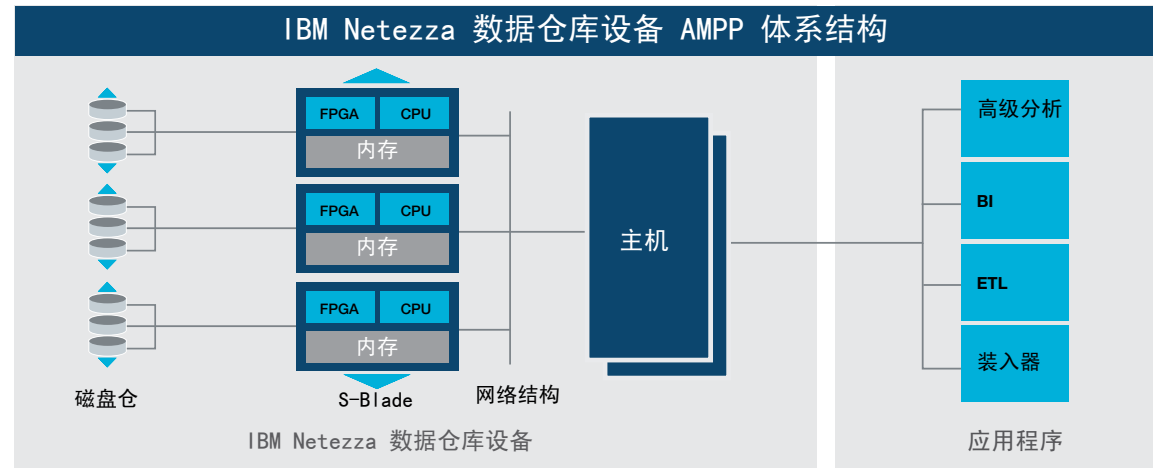
Exadata 单元不会处理独特的聚集（在简单报告中较为普遍）、复杂连接或分析函数（用于分析应用程序中）。



通过将现有数据库技术与新的“智能”存储层结合，Exadata 消除了 Oracle 在与传统存储产品配合使用时所遭受的磁盘吞吐量瓶颈。Exadata 可能看似为寻求将多个 OLTP 系统整合到单个平台的 CIO 提供了有趣的机会。但是，当进一步研究时就会发现，Exadata 中包含的存储量和存储软件很明显对于绝大多数 OLTP 环境都不适用。除简单查询外，Exadata 必须将大的数据集从其存储层移至其数据库层，由此引起对其作为现代数据仓库平台的稳定性的质疑。

### IBM Netezza 数据仓库设备的查询性能

IBM Netezza 数据仓库设备完全作为数据仓储平台进行设计。IBM Netezza 采用“非对称大规模并行处理”（AMPP）体系结构。“对称多处理”主机<sup>7</sup>面对由“大规模并行处理”节点组成的网格。IBM Netezza 数据仓库设备利用此 MPP 网格来处理急剧增加的仓储和分析数据。



<sup>7</sup> IBM Netezza 数据仓库设备具有两台 SMP 主机以实现冗余，但任何时候仅一台主机处于活动状态。



IBM Netezza 数据仓库设备网格中的节点称为 S-Blade（片段刀片），这是包含多核中央处理单元（CPU）的独立服务器。每个 CPU 与一个多引擎现场可编程门阵列（FPGA）和千兆字节的随机存取存储器结合使用。由于 CPU 具有自己的内存，因此它们唯一地专注于数据分析，绝不会转移精力跟踪共享内存数据库系统中其他节点上发生的块 ping 或集群冻结活动。

FPGA 是一种半导体芯片，配备有大量可编程以实现几乎所有逻辑功能的内部电路，在管理流式处理任务方面特别有效。FPGA 位于 IBM Netezza 数据仓库设备外部，在此类应用中用作数字信号处理、医学影像成形和语音识别。IBM 数据仓库设备工程团队已在我们设备的 FPGA 中构建软件机器，以在数据到达 CPU 之前加快对其进行处理。在每个 Exadata 机架内，Oracle 专用 14 台八路存储服务器，而所完成的数据处理量还少于 IBM Netezza 数据仓库设备利用我们刀片服务器内嵌入的 48 个 FPGA 所实现的处理。每个 FPGA 只是 1 平方英寸大小的半导体芯片，凭借其出色的效率完成工作，功耗和发热量都极低。

IBM Netezza MPP 网格中的节点间通信在运行基于 IP 的定制协议的网络结构上发生，从而可以充分跨区带宽，并可以消除拥塞，即使在发生持久、突发性的网络

---

**IBM 未采用存在明显缺陷的旧系统，而是使用新的更智能的存储层加以平衡；IBM Netezza 数据仓库设备设计为优化的数据仓储平台。**

---

流量的情况下，也是如此。该网络可优化为扩展至超过 1000 个节点，同时允许每个节点启动对所有其他节点同时进行的大量数据传输。这些传输为处理数据仓储和高级分析之类的任务带来了极大的效率。正如 SQL 语句得益于在 IBM Netezza 数据仓库设备 MPP 体系结构

中进行处理一样，高级分析核心的计算复杂的算法同样如此。前几代的技术将应用程序处理和数据库处理物理分离，由于大的数据集要在仓库和分析处理平台之间来回移动，从而造成效率低下并产生约束。IBM Netezza 数据仓库设备在其 MPP 网格中进行繁重的高级分析计算，在每个在物理上接近数据的 CPU 中运行算法，从而使数据移动成为多余，并推动性能的提高。得益于在 IBM Netezza 数据仓库设备 MPP 网格的许多节点上运行，这些算法不会受到伸缩性较差的集群系统的约束。

IBM 不象竞争对手那样使用新的智能存储层来平衡存在明显缺陷的旧系统；IBM Netezza 数据仓库设备本身就是专为数据仓储而设计的最优化平台。IBM Netezza 数据仓库设备能够实现出色的性能，从而使程序员、管理员和用户的工作更为轻松。

## 操作简易性

来自金融服务行业的 IBM Netezza 客户使用“精益”方法分析管理其 Oracle 数据仓库所需的资源支出。他们了解到，在构建和管理索引、聚集、具体化视图和数据集市方面，其 IT 团队 90% 以上的工作都浪费掉或花在非增值的处理上。

### Oracle Exadata 的操作简易性

必须在数据仓库中装入数据，然后才能运行查询。Exadata 的存储层为 MPP 网络。当所有节点都以同等机会参与当前的计算任务时，MPP 系统能够实现性能和规模效应。数据必须均匀分布，对于每个查询，每个节点上的相关数据量必须尽可能相同。要将数据均匀分布到 Exadata 的存储服务器网络上，要求管理员在设计、管理及维护复杂分区、文件、表空间、索引、表和块/扩展数据块大小方面经过培训并具有丰富经验。“如果系统不过分依赖复杂分区来实现出色性能，那么情况可能更为理想。”<sup>8</sup>

来自金融服务行业的 IBM 客户使用“精益”<sup>9</sup> 方法分析管理其 Oracle 数据仓库所需的资源支出。他们了解到，在构建和管理索引、聚集、具体化视图和数据集市方面，其 IT 团队 90% 以上的工作都浪费掉或花在不产生价值的处理上。这种浪费转化为不必要的硬件和软件许可成本、浪费太字节的存储空间、延长的开发和数据装入周期、数据长期不可用、数据闲置、负载和查询执行性能低下以及过多的管理成本。

Exadata 几乎未对 Oracle 数据仓库管理进行简化。管理员必须管理多个服务器层，每一层都具有要维护的操作系统映像、固件、文件系统和软件。Oracle 建议 DBA 管理 11g (Exadata 中的数据库版本) 所花费的时间比其花费在较旧的 10g 部署上的时间应减少 26%。如果这在实践中得以确认，并且 Exadata 将客户浪费在无价值管理方面的时间减少四分之一，那么表明 Oracle 已朝着正确的方向进了一步。IBM Netezza 的设备并非旨在浪费客户的任何时间。“DBA 团队只需备份环境和管理设备的高级别安全性模型，仅此而已。他们无需执行其他任何任务（例如，在与 IBM 开展业务时，索引的概念对他们而言非常陌生）。”<sup>10</sup>

<sup>8</sup> Curt Monash, 地址为 <http://www.dbms2.com/2009/09/21/notes-on-the-oracle-database-11g-release-2-white-paper/>

<sup>9</sup> “精益”一词源自制造业，是指使用六西格玛的工具和方法，分析浪费的资源支出，确定无法为产品和服务增添价值的活动以加以消除。

<sup>10</sup> 客户将 Oracle 用于 OLTP，将 Netezza 用于数据仓储，引用自 LinkedIn Exadata Vs Netezza 论坛，地址为 [http://www.linkedin.com/groupAnswers?viewQuestionAndAnswers=&gid=2602952&discussionID=11385070&sid=1275353329699&trk=ug\\_qa\\_q&goback=.ana\\_2602952\\_1275353329699\\_3\\_1](http://www.linkedin.com/groupAnswers?viewQuestionAndAnswers=&gid=2602952&discussionID=11385070&sid=1275353329699&trk=ug_qa_q&goback=.ana_2602952_1275353329699_3_1)

业务用户不仅要求快速完成查询，还希望能够体验一致的性能；如果报告昨天花五秒钟便完成，而今天花了三分钟，他们就可能会创建要求IT 帮助热线人员响应的凭单。仓库的性能不可避免地取决于变化且动态的工作负载需求。将会装入通过批处理作业或零星供应从 OLTP 系统到达的数据，在业务人员不可见的后台运行的诸如备份和复原之类的管理任务会持续更新。与此同时，计算密集型应用程序（如预测哪些索赔或贸易可能是欺诈或违法行为的应用程序）会在数据仓库基础架构上突然产生繁重的负载。为企业实现一致性能对数据库有两个要求：查询性能一致和有效的工作负载管理。这样可简化将可用计算能力分配给所有需要服务的作业的操作，通常根据和业务部门商定的优先级进行分配。

Oracle 的工作负载管理理念是为管理员提供多个调优参数。Oracle 的参数相互之间的依赖度高，在 Exadata 中，部分参数对于网格中的所有处理器必须设置为相同值。这种复杂性迫使管理员实验性地更改参数

设置，根据意外的数据仓库需求进行调优。通过增加和减少负载，为具有不同应用程序和数据需求的大群用户实现和维护一致性能，是一项复杂任务，要求具备丰富的 Oracle 经验和仓库管理员专业知识。

在事务量可预测且稳定的 OLTP 系统中，企业不受此复杂性的影响。数据库管理员在应用程序开发阶段有充分机会对每项操作的数据需求进行分析，并且有时间设计、测试和调优数据库。而数据仓库则不同。外部世界中的事件可能会产生以前所未有的方式分析数据的需求。即时信息需求使管理员没有时间来分析每项查询并优化数据检索。由于数据仓库的公式化本质，因此无法立即处理请求，从而可能会错失行动的商机。

---

“我们对其 [IBM Netezza 数据仓库设备] 进行概念证明的方式是，IBM 发来一个盒子，我们要做的就是将其放入数据中心并接入网络。在 24 小时内，设备已启动并正常运行。我并没有夸大其词，就是这么轻松。”

---

### IBM Netezza 数据仓库设备的操作简易性

IBM Netezza 的客户愿意证明我们的设备易于安装和使用。“我们对其 [IBM Netezza] 进行概念证明的方式是，IBM 发来一个盒子，我们要做的就是将其放入数据中心并接入网络”，他说道。“在 24 小时内，设备安装调试完毕并正常运行。我并没有夸大其词，就是这么轻松。”<sup>11</sup>

这段评论来自一家已在使用 Oracle 数据库和 RAC 软件，在业界领先的社交网络企业的技术副总裁。

<sup>11</sup> [http://www.computerworld.com/s/article/9126930/EHarmony\\_finds\\_data\\_warehouse\\_match\\_with\\_Netezza?source=rss\\_news](http://www.computerworld.com/s/article/9126930/EHarmony_finds_data_warehouse_match_with_Netezza?source=rss_news)

---

#### 针对简单方法的说明

- 无集群互连 (GES 和 GCS) 监控/调优
  - 无特定于 RAC 的知识/调优 (具备 RAC 经验的 DBA 并不多见)
  - 无数据库空间/表空间大小缩放和配置
  - 无重做/物理日志大小缩放和配置
  - 无日记记录/逻辑日志大小缩放和配置
  - 无页面/块大小缩放和表配置
  - 无扩展数据块大小缩放和表配置
  - 无临时空间分配和监控
  - 无操作系统内核建议集成
  - 无操作系统建议的补丁级别维护
  - 无用于配置主机/网络/存储器的 JAD 会话
  - 无查询 (例如 first\_rows) 和优化器 (例如 optimizer\_index\_cost\_adj) 提示
  - 无统计信息包 (统计信息、高速缓存命中、等待事件监控)
  - 无内存调优 (SGA 和块缓冲区等)
  - 无索引规划/创建/维护
  - 简单分区战略: 散列或轮询
- 

缩短实现生产力所需的时间是一个良好开端; IBM 的理念是在数据仓储的所有阶段都实现简易性。客户面临的第一个任务是装入数据。IBM Netezza 数据仓库设备可以自动执行数据分发。根据概念证明项目的经验, 客户使用自动分发功能将数据装入 IBM Netezza 数据仓库设备, 运行查询, 并将结果与高度调优的 Oracle 环境进行比较。除最简单的查询外, IBM Netezza 数据仓库设备仅凭自动分发便足以在性能方面胜过 Oracle。客户稍后可以分析所有查询, 以确定可通过按照不同键重新分发数据来加速的查询。IBM Netezza 数据仓库设备使这项任务变得简单。

提交到 IBM Netezza 数据仓库设备的所有查询都会在其大规模并行网格中自动处理, 而无需数据管理员介入。查询和分析通过主机进入 IBM Netezza 数据仓库设备, 在该设备中, 优化器、编译器和调度程序将查询和分析分解为许多不同的部分或片段, 然后将这些指令分发给由处理节点或 S-Blade 组成的 MPP 网格, 所有这些节点针对其本地受管数据片同时处理其工作负载。

到达每个 IBM Netezza 数据仓库设备 S-Blade 的片段会启动将压缩数据从磁盘读入内存的操作。然后, FPGA 从内存缓冲区读取数据, 并利用其压缩引擎进行解压缩, 然后立即将磁盘中的每个块转换为 FPGA 中等效的 4 到 8 个数据块。IBM Netezza 数据仓库设备的设计是对所有数据仓库中最慢的组件 (即磁盘) 进行加速。接下来, 在 FPGA 中, 数据流入项目引擎, 该引擎根据要处理的 SQL 查询的 SELECT 语句中指定的参数来过滤列。只有执行 SELECT 子句的记录会进一步向下游传递到限制引擎, 该引擎根据 WHERE 子句中指定的限制, 阻止不需要处理查询的行通过门。可视性引擎以流式方法速度保持 ACID (原子性、一致性、隔离和耐久性) 一致性。所有这些工作 (不断去除不需要的行和列) 都是在仅一英寸见方的节能型 FPGA 中完成的。如果 IBM Netezza 数据仓库设备的设计确定不需要移动数据, 那么就不会移动。

FPGA 的预处理完成, 它将刚生成的经修剪的记录集重新流回 S-Blade 内存, 然后由该 CPU 与 MPP 网格内的所有其他 CPU 一起并行执行诸如排序、连接和聚集

之类的高级数据库操作。CPU 还可能会应用片段代码中嵌入的复杂算法，以进行高级分析处理。CPU 最后汇集来自整个数据流的所有中间结果，为片段生成结果，并且按片段代码的指示，将结果通过网络结构发送到其他 S-Blade 或主机。当 JOIN 所需的数据未在节点上并置时，IBM Netezza 数据仓库设备节点间网络结构会在数据库完成限制和投影后，在处理周期后期高效简单地重新分发数据。一些高度复杂的算法需要在节点之间进行通信以计算其答案。IBM Netezza 数据仓库设备的设计是利用消息传递界面传达临时结果和产生最终结果。

并且，由于原始压缩数据块仍在内存中，因此可通过 IBM Netezza 数据仓库设备表高速缓存，稍后查询类似数据，从而使这些数据块得到自动复用，这种自动机制无需 DBA 培训或参与。

---

在迁移至 IBM Netezza 数据仓库设备仅三个月之后，一位客户便讲述其团队完成的分析应用比他们前三年使用 Oracle 所完成的都要多。

---

由于 IBM Netezza 数据仓库设备对所有任务应用完全并行性，因此其工作负载管理系统在控制可供所有作业使用的设备计算资源量方面起着至关重要的作用。在 IBM 设备体系结构中，由一个软件组件控制所有系统资源：处理器、磁盘、内存和网络。这种精巧的设计是 IBM Netezza 数据仓库设备“工作负载管理系统”的基础。通过 IBM Netezza 数据仓库设备“工作负载管理系统”，管理员根据与业务部门商定的优先级，将计算资源分配给用户和组，并且针对不同社区保持一致的响应时间。

IBM Netezza 数据仓库设备消除了浪费时间的数据库调优工作。IBM Netezza 设备能够自行做出智能的决策，无需调优工作，所需的系统管理工作也极少。对各种动态且不断变化的工作负载保持一致的性能所需的管理任务非常少，对于任何具有 Linux 和 SQL 经验的用户而言都轻而易举。只需要管理员将 IBM Netezza 数据仓库设备资源分配给用户社区中的各个组，然后将控制权移交给“工作负载管理系统”即可。技术人员摆脱了没有尽头的数据库管理，可以与业务部门合作，致力于研究有关利用数据创造价值的新方法。在迁移至 IBM Netezza 数据仓库设备仅三个月之后，一位客户便讲述其团队完成的分析应用比他们前三年使用 Oracle 所完成的都要多。处理与数据管理位置接近的分析应用程序，利用与用于处理 SQL 相同的 MPP 平台，这代表了真正的机遇，可使组织显著增加从数据中产生的价值。

## 价值

### Exadata 的价值

如 IBM Netezza 数据仓库设备和 Oracle 的金融服务客户执行的浪费分析所突出显示的一样，使用 Oracle 进行数据仓储是一项劳动密集性工作。Oracle 宣称其最新版本的数据库管理系统可能会将此浪费减少 26%。<sup>12</sup> IBM Netezza 数据仓库设备客户证明这些低级别、技术要求严苛的管理任务完全不必要；由此无可辩驳地证明，操作 Oracle 数据库需要管理员将其绝大多数时间花费在维护和供应底层技术上，而 IBM Netezza 数据仓库设备客户则将时间花在利用其数据创造价值方面。

Exadata 的新存储层需要管理员进行调优和管理，这又增加了一层复杂性。因为 Exadata 非常新，很少有使用该技术的数据仓库已投入生产环境，所以预测其拥有成本还为时尚早。不过，客户应当可以预计到，通过 Exadata 实现一致的高性能将在数据库设计和管理方面产生非常高的成本。

尽管添加新的存储层可以消除与 Oracle 数据库有关的磁盘吞吐量瓶颈，但是 Exadata 的设计是更倾向于适应大规模并行处理，而非充分利用体系结构。Oracle 无法将数据管理完全集成到 Exadata 存储层，这意味

着对其 MPP 网格中的硬件要求过少。这大幅增加了采购 Exadata 的成本；客户为从未完全利用的硬件付款，并且为新增但功能有限的存储软件层付款。这些成本在仓库生命周期内会逐渐累积。客户为其数据中心内利用率不足的空间付款，如果将这些空间用于容纳更高效的计算机系统，那么能够得到更大的回报。

尽管成本会削弱价值，但是基本问题在于 Exadata 能否帮助客户创造价值。第一代数据仓库在使组织时刻了解刚刚发生的情况方面起着重要作用，而第二代数据仓库则通过高级分析以及本文中先前所讨论的其他功能释放了数据的更大潜力。与传统存储产品结合的 Oracle RAC 在此方面取得了有限的技术成功，至今仍须证明在此角色中的成功。Exadata 的存储层无法处理复杂连接、独特的聚集和分析函数。难以想象两种本身并不胜任高性能深入分析超大数据集的技术通过快速网络连接并放置在同一机架中时，可以实现此类高级功能。

### IBM Netezza 数据仓库设备的价值

IBM Netezza 数据仓库设备的工程师将数据管理和分析深入集成在大规模并行且无共享的网格中。我们根据此创新规划的一个成果是为客户提供简易性，这直接转变为相比使用诸如 Oracle 之类的传统数据库产品，显著降低拥有和操作数据仓库的成本。

数据仓库需求不只是处理简单的 SQL；为充分利用数据，要求仓库能够运行预测模型、调查图及其他分析应用程序。举例说明，金融服务公司了解到某个家庭最近进行了抵押贷款，而且先前购买了投资产品和贷款产品，并具有支票帐户，但从未购买保单，那么预计该家庭接下来最可能购买的是另一次抵押之后的投资产品，那么该公司可针对此客户创建有针对性的价值市场促销活动，并极有可能获得成功。<sup>13</sup>

<sup>12</sup><http://www.dbms2.com/2009/09/21/notes-on-the-oracle-database-11g-release-2-white-paper/>

<sup>13</sup>请参阅“Dynamic Bayesian Networks for acquisition pattern analysis: a financial-services cross-sell application”，由曼彻斯特商学院 Marketing Group 的 Anita Prinzie 和根特大学市场营销系的 Dirk Van den Poel 合著



评估系统		IBM Netezza	Oracle
项目		IBM Netezza 1000	Exadata V2 (SAS)
性能和体系结构	MPP	<ul style="list-style-type: none"> <li>真正的 MPP</li> <li>针对数据仓储和分析进行优化</li> </ul>	<ul style="list-style-type: none"> <li>混合 - 并行存储节点和 SMP 集群头节点</li> <li>通用化的体系结构</li> </ul>
	硬件体系结构	<ul style="list-style-type: none"> <li>完全处理 S-Blade (1 个 CPU 内核 + 1 个 FPGA 内核/1 个磁盘驱动器)</li> <li>主要用于用户/应用程序界面的 SMP 主机节点</li> <li>独立的刀片到刀片重新分发</li> </ul>	<ul style="list-style-type: none"> <li>智能存储 (1 个 CPU 内核/1.5 个磁盘驱动器)</li> <li>运行 Oracle 11g RAC 的 SMP 集群节点</li> <li>InfiniBand (Exadata 节点到 SMP 集群)</li> <li>所有数据重新分发中的头节点处理</li> </ul>
	数据流	<ul style="list-style-type: none"> <li>针对 S-Blade 的 FPGA 性能辅助 - 分解、谓词过滤、行级安全性实施</li> <li>&gt;95% 的工作在 S-Blade 上进行</li> </ul>	<ul style="list-style-type: none"> <li>Exadata 节点主要用于分解和谓词过滤</li> <li>大多数数据仓库和分析工作在 SMP 头节点上进行</li> </ul>
	数据库内分析	<ul style="list-style-type: none"> <li>完全实施的用于分析的 MPP 平台</li> <li>用户定义的函数、聚集和表</li> <li>语言支持: C/C++、Java、Python、R 和 Fortran</li> <li>范式支持: SQL、矩阵、网格和 Hadoop</li> <li>内置的一组超过 50 种关键分析 (完全并行化)</li> <li>集成开发环境: Eclipse 和 R GUI, 包含向导</li> </ul>	<ul style="list-style-type: none"> <li>分析处理仅限于头节点集群</li> <li>用户定义的函数和聚集</li> <li>语言支持: C/C++ 和 Java</li> <li>范式支持: SQL 和矩阵 (次)</li> <li>基本分析函数</li> </ul>
	规模	<ul style="list-style-type: none"> <li>线性性能和数据大小可伸缩性</li> <li>功能齐全的企业级工作负载管理及其他功能</li> </ul>	<ul style="list-style-type: none"> <li>非线性性能和数据大小缩放 - 性能和 I/O 瓶颈 (在/到头节点集群)</li> </ul>
简易性	设备系统管理和集成	<ul style="list-style-type: none"> <li>无需调优、无需建立索引、无分区</li> <li>旨在实现最佳性价比的平衡系统</li> </ul>	<ul style="list-style-type: none"> <li>严重依赖于性能调优</li> <li>性能取决于物理数据库设计技能, 包括索引和分区</li> </ul>



## 总结

此分析超出 SQL 的能力，需要一种称为动态贝叶斯网络 (Dynamic Bayesian Networks) 的技术。不过，该分析使用与 SQL 所处理相同的数据来创建报告和仪表盘，从而建议对仓库使用扩展角色。

IBM Netezza 数据仓库设备是从头开始创建的，用于处理 SQL 和高级分析应用。IBM Netezza 数据仓库设备使客户不必使用专有语言。客户可以将现有应用程序移植到 IBM Netezza 数据仓库设备，或者选择使用所选语言（包括 R、C/C++、Java、Python 和 Fortran）开发新的分析应用程序。客户可以利用内置的并行化、数据库中的算法库，包括数据分区、数据挖掘、预测分析、地理空间数据和矩阵代数。此外，客户还可选择将 Hadoop/MapReduce 作为一种高度可伸缩的摄取机制，以便在将由面向公众的 Web 应用程序和 Web 日志生成的大量数据集装入到 IBM Netezza 数据仓库设备中进行按需应变分析之前，对这些数据集进行预处理。

IBM Netezza 数据仓库设备作为在数据仓储领域替代 Oracle 的主要方法的形式出现。将数据仓库和集市从 Oracle 移至 IBM Netezza 数据仓库将会创造新的机会而非风险。大多数 IBM 客户都已走过这条道路，其中许多客户通过与具有良好成功记录的系统集成公司合作，成功实现了迁移。Exadata 是 Oracle OLTP 平台的演进，定位为针对 OLTP 和分析的通用平台。Oracle 的数据库管理系统是为 OLTP 而设计的，其数据量相比数据仓库而言较小。OLTP 系统的数据库活动在投入生产环境之前可进行评估；管理员有时间分析、测试和优化每个事务的数据检索。数据仓库必须立即处理业务部门获取数据所需进行的任何查询；但需要管理员调解的技术并不适合此项任务。将此技术用于除事务处理之外的其他作用会对用于管理和运行数据仓库的人员和流程施加巨大压力。

---

“这 [IBM Netezza 数据仓库设备] 是第一款具有与我们自己的路线图完全一致的长期产品路线图的数据库产品。我们将此称为按需应变数据库。”

—一家大型证券交易集团的首席数据官

---

Oracle 向客户说，Exadata 在体系结构上与 IBM Netezza 数据仓库设备类似但更优越，因为 IBM Netezza 数据仓库设备并非每种数据类型或 SQL 标准都支持，并且不支持数据挖掘或高并发性。IBM Netezza 数据仓库设备的客户对此有不同意见：“这 [IBM Netezza 数据仓库设备] 是第一款具有与我们自己的路线图完全一致的长期产品路线图的数据库产品。我们将此称为按需应变数据库”，<sup>15</sup>一家大型证券交易集团的首席数据官说道。

鉴于不同的工作负载特征，很少有客户尝试在同一基础结构上运行 OLTP 和数据仓库系统；为此，要做到同时运行，就需要不断调优和优化。技术人员处于一种两难境地：要么接受 OLTP 和数据仓储性能同时受损的局面，要么不断重新配置数据库，以徒劳地尝试满足不同工作负载的冲突需求。

如本电子书中先前所提及，演进至第二代数据仓库的组织在不同的平台上运行 OLTP 和仓库系统，各系统根据其自己的工作负载需求专门进行配置。

真正紧要的唯一数据仓库是您自己的数据仓库，也就是在您的数据中心内针对您的数据运行的您自己的应用程序。现场概念证明 (PoC) 为 IT 部门创造了充分研究技术的机会，从而了解到可以如何使用 IBM Netezza 数据仓库设备来帮助业务部门的同事从数据中获取更大价值。要充分利用此机会，就要求 PoC 由适用于其他项目的同一规程进行管理。

Curt Monash 在其博客 “Best practices for analytic DBMS POCs,” 中提出了明智的建议，<sup>16</sup> 包括让独立顾问参与进来，以指引项目产生成功的结果。对于希望了解其仓库如何以零成本和零风险在 IBM Netezza 数据仓库设备上运行的合格组织，IBM 可提供测试。要订购，请转至

[ibm.com/software/data/netezza/](http://ibm.com/software/data/netezza/)

IBM Netezza 数据仓库设备：使用是一种享受。

## 传播

请将本文档与您的朋友和同事分享。

## 反馈

您对本电子书中的观点和论证有何看法？让我们了解您赞同哪些内容、不同意哪些内容，或者可能希望就哪些内容进行进一步讨论。

- 与 IBM Netezza 数据仓库社区畅聊：  
[www.enzeecommunity.com/groups/](http://www.enzeecommunity.com/groups/)

## 联系我们

- 访问 IBM Netezza 博客：  
<http://thinking.netezza.com>
- 访问 IBM Web 站点：  
<http://www.ibm.com/software/data/netezza/>
- 访问 IBM Netezza 数据仓库社区 Web 站点：  
[www.enzeecommunity.com](http://www.enzeecommunity.com)

<sup>16</sup> [www.dbms2.com/2010/06/14/best-practices-analytic-database-poc/ffmore-2297](http://www.dbms2.com/2010/06/14/best-practices-analytic-database-poc/ffmore-2297)

## 关于作者

### Phil Francisco, IBM 产品管理和市场营销副总裁

Phil Francisco 将 20 多年的经验带入技术开发和全球技术市场营销中。身为 IBM 产品管理和市场营销副总裁，他促进新的业务和产品战略，指导产品组合和推动产品市场营销计划。在加入 IBM 之前，Francisco 是 PhotonEx 的市场营销副总裁，这是一家领先开发商，主要为电信网络供应商提供 40 Gb/秒光传输系统。在加入 PhotonEx 之前，Francisco 担任 Lucent Technologies 中 Optical Networking Group 的产品市场营销副总裁，在此他与全球最大的一些电信运营商合作，规划和实施光纤网络解决方案。Francisco 先生持有高级光纤网络体系结构方面的专利。他从宾夕法尼亚大学莫尔学院电子工程系以优异成绩获得电子工程学士学位和计算机科学学士学位。他从斯坦福大学获得电子工程硕士学位，并在杜克大学福库商学院修完了高级管理课程。请阅读 Phil 的博客：[http://thinking.netezza.com/blogs\\_by/phil-francisco](http://thinking.netezza.com/blogs_by/phil-francisco)

### Mike Kearney, IBM 产品市场营销资深总监

Mike 已在信息技术领域工作超过 25 年。在 IBM 中，其专职工作是向希望从数据中获得更大回报的组织介绍 Netezza 的设备性能、价值和简易性。Mike 先前在 Vignette、BMC Software 和 Oracle Corporation 工作过，并曾在金融服务、电信和能源行业类公司任职。Mike 拥有伦敦大学学士学位和考文垂大学硕士学位。

## 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：

<http://www-01.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/software/cn/data](http://ibm.com/software/cn/data)**

IBM、IBM 徽标、ibm.com 和 Netezza 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/software/cn/data/legal/copytrade.shtml](http://ibm.com/software/cn/data/legal/copytrade.shtml)**

Microsoft 和 SQL Server 是 Microsoft Corporation 在美国和/或其他国家或地区的商标。

Oracle 和 Exadata 是 Oracle 和/或其子公司的注册商标。

其他公司、产品和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