

# InCompariso



## IBM Smart Analytics Systems 与 Oracle Exadata X2-2 的比较

Bloor Research 比较报告  
作者: Philip Howard  
出版日期: 2010 年 12 月

本文档由 IBM 出资制作。虽然本文档可能会利用来自包括 IBM 在内的各个供应商的公开可用的材料，但这未必反映了此类供应商关于本文档中所探讨的问题的立场。

---

当我们着手研究时，我们预计会发现 **IBM** 精于某些领域，而 **Oracle** 则擅长另一些领域。令我们惊讶的是，情况并非如此，**IBM Smart Analytics System** 几乎在我们测试过的所有领域中都优于 **Oracle Exadata**。

Philip Howard

## 执行摘要

本报告分为两部分：本节与结论部分，针对的是没有或只具备有限技术知识或兴趣的高级管理人员，本文的其余部分则针对更关注技术的人员。特别要指出的是，本节中的信息在本文其余部分有所重复，但我们将在后文提供更详细的内容。

本文及其姊妹篇（IBM pureScale Application System 与 Oracle Exadata X2-2 的比较）的基本主题是分别为数据仓储和联机事务处理 (OLTP) 方面提供对 IBM 和 Oracle 的产品综合比较。

Oracle 对这两组要求的观点是，一个单一的解决方案，即 Oracle Exadata，就可以理想地同时覆盖这两个领域；尽管在我们看来（我们认为 Oracle 不会有异议），这两个环境的要求有很大差异。相比之下，IBM 的态度是，在这些领域中，每个领域都需要有不同的侧重点，因此它提供针对 OLTP 环境的 IBM pureScale Application Systems 和针对数据仓储的 IBM Smart Analytics Systems。

在实践中，IBM 的方法并没有这么简单。在数据仓储方面，其实也有两种可能的方法：许可 DB2 和 InfoSphere Warehouse 或 IBM Smart Analytics System。Smart Analytics 产品包含了前者，但系统已配合选定的特定硬件平台以及所有可选模块（如 Cognos 商业智能模块）进行预建和预调优。实际上，许可 DB2 和 InfoSphere Warehouse 适合那些更喜欢 DIY 方法的用户，而 IBM Smart Analytics System 则适合那些想要一个随时可运行的完整系统的用户。类似的概念也适用于 IBM 的 OLTP 产品。

IBM 把其做法称为“工作负载优化的系统”。也就是说，这些产品，特别是软件包，已针对其特定环境进行设计和优化。这包括 System x（运行 Linux 或 Windows）、

System p（运行 AIX）或 System z（在大型机上运行 z/OS）。我们将在本文解决的问题是，在 Oracle 和 IBM 这两种方法中，那一种是最好的方法。当然，人们可以在理论上找到论据来支持 Oracle 或 IBM 的方法，但这样我们将一直

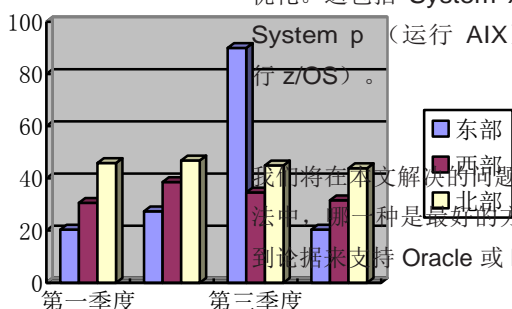
争论到猴年马月。从理论的角度来判断一个概念的好坏是毫无意义的；最重要的是，哪个方法在性能、可扩展性、行政和管理的简易度以及成本方面的表现最好。

首先要指出的是，最小配置的 Oracle Exadata 产品有 6TB 的可用数据。如果采用了压缩技术，这相当于至少 18Tb，并且实际上很可能会更多。相反，最小的 Smart Analytics System 的可用磁盘空间有 330Gb。因此，我们可以立即得出结论，Smart Analytics System 将适合于许多中小型组织以及部门，而 Oracle Exadata 对于这些用户将显得规模过大和过于昂贵。当然，这些用户可能会自己选择 Oracle Database 11g Release 2，但这不属于我们的考虑范围。

另一个区别是 IBM 升级 Smart Analytics System 的方式。不同于 IBM，Oracle 对 Exadata 有硬性规定：您可以选择一个 ¼ 机架，一个 ½ 机架或全机架的倍数。如果您想添加新的处理节点，您必须额外添加一台或多台存储服务器，并且除了涉及硬件成本之外，您还必须支付增加的许可费用来运行 Exadata 软件，其价格为每个磁盘 10,000 美元（而每台服务器有 12 个磁盘）加上 22% 的维护费用；而这一切只是因为您需要一些处理能力，即使您并不需要更多磁盘容量！

相反，IBM 提供了在用户喜欢的操作系统上运行的多种型号，您可以通过传统的方式对其进行升级。您可以选择提供固态硬盘的型号或不提供固态硬盘的型号，如果您的确选择使用这种技术的型号，那么您也可以升级这种技术。而另一个极端是，所有 Oracle Exadata 实施都使用 Flash 存储，即使它是 Oracle 在 OLTP 环境中的主要目标。

就迄今为止的表现而言，基于理论的比较（这两家公司都没有基于当前系统发布基准测试结果，虽然 IBM 目前拥有比例系数为 10TB 的 TPC-H 基准测试记录），与其说这个问题是在比较苹果和梨，还不如说是在比较水果盘：双方都有苹果，一个提供新西兰考克斯苹果，而另一个提供澳洲青苹；一个用香蕉和葡萄来配他们的苹果，另一个



## 执行摘要

则用油桃和奇异果。因此，我们要比较 Oracle (Exadata Storage Server 本机的功能加上存储索引及混合列式压缩、工作负载管理) 和 IBM (背驮式扫描、标记化压缩、多维集群、多维数据服务和工作负载管理)，再加上数据库的所有常见性能特征，以及厂商对 Flash 磁盘的不同使用方式。更不用说在大型机上运行的 Smart Analytics System 与在分布式平台上运行的系统所提供的设施亦有所不同。

我们的观点是，根据客户的特定要求，两个厂商可能各有性能上的优势。具体而言，我们预期 Oracle Exadata 在相对静态的数据方面表现得比较好（因为这种情况下可以得到更好的压缩），并且能更好地支持通常与分析应用程序、数据挖掘或统计分析相关的复杂查询类型。另一方面，我们更喜欢 IBM 的多维数据服务、压缩和工作负载管理。一般来说，要确定哪个供应商可以提供能满足您的工作负载和需求的最佳性能特征，概念证明将是最好的方法。

谈到定价本身，IBM 已经针对在 Oracle 环境中不需要任何软件许可的 Smart Analytics System 进行定价。也就是说，这些客户拥有至少数据库的无限许可协议 (ULA)。对于这些用户来说，一个小型 Smart Analytics System，拥有与 ¼ 机架的 Oracle Exadata 系统几乎相同的处理器性能和存储容量，成本与其对手相近，而中型和大型配置的成本则不如其竞争对手的 ½ 机架和全机架配置昂贵。当然，如果您没有 ULA，那么这些价格比较对于 IBM 是非常有利的。但是，请记住，标价一向都以变化无常著称，有时和现实并没有任何相近之处。

除了价格，两个系统之间的最大差异也许在于可管理性和灵活性。我们已经从所提供的配置和磁盘容量角度讨论过后者，说到可管理性，首先，Smart Analytics Systems 更容易安装，更容易扩展，并且在任何时候都不要应用程序调

优（这是指在添加一个节点到 Oracle 集群的情况下）。在某些环境中，Oracle Exadata 将优胜于 IBM，而在另一些环境中，情况则正好相反。然而，在我们看来，IBM Smart Analytics System 在所有其他方面都优胜于 Oracle Exadata：它更容易管理和调优，更容易安装，更灵活，花（至少名义上）更少的钱。

## 系统描述

在我们开始进行任何比较之前，我们需要清晰了解每个产品所提供的架构，每个产品中包含什么，不包含什么。

### IBM InfoSphere Warehouse

IBM InfoSphere Warehouse 可能被认为就是 DB2。然而，有一套工具围绕着 InfoSphere Warehouse 中的 DB2，以简化在 DB2 之上构建的数据仓库的创建和管理。Linux/UNIX/Windows (LUW) 版本包括 DB2 服务器，而 z/OS 版本则不包括，因为大部分感兴趣的客户已经拥有许可。包括在 InfoSphere Warehouse 中的是图形化和基于 Web 的工具，以支持物理数据模型的开发和执行、数据移动流 (SQW)，OLAP 分析（多维数据服务）和数据挖掘（仅 LUW 版本支持该功能）。

就本文而言，我们将不讨论 InfoSphere Warehouse 本身，因为 Smart Analytics System 包括了它的所有元素。

### IBM Smart Analytics System

IBM Smart Analytics System 和使用 InfoSphere Warehouse 的 DB2 之间的区别是，后者仅仅包括可以在任何要安装它的平台上运行的软件，并且不需要任何方式的预配置、预调优或预构建。实际上，这是一个自己动手的解决方案。另一方面，Smart Analytics System，恰恰是以一种预配置、预调优和预构建的方式在所有 IBM 操作系统环境（大型机 z/OS、AIX、Windows 和 Linux）中提供，配有可以从入门级系统就开始使用的各种不同选项。这种对各个组件的预集成和预调优，意味着更容易获得更好的性能，并且可以获得更快的价值实现时间。

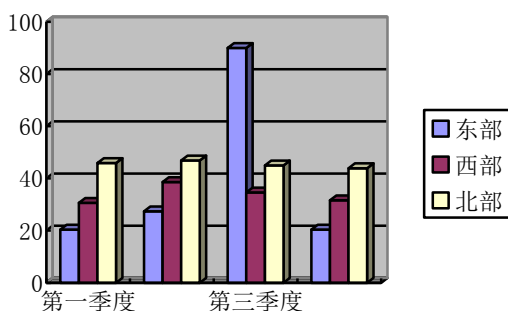
### Oracle Exadata X2-2

实际上有两个 Oracle Exadata Database Machine 产品：Exadata X2-2 和 Exadata X2-8。后者只提供全

机架系统，主要针对最大的 OLTP 和整合环境。该型号目前尚未提供，因此我们将重点放在 Exadata X2-2。其组成为 Oracle Database 11g Release 2、Oracle RAC（真正应用集群）Database 服务器网格、一个 InfiniBand 互联、Oracle Enterprise Linux 操作系统和使用高性能 (600Gb) 或高容量 (2Tb) 磁盘存储的 Exadata Storage Server Grid，后者提供更多容量，但性能较低。

系统运作的方式是，数据存储 Exadata Storage Server 网格中，在把结果传递给数据库本身之前，存储服务器作为一种预处理器使用 Oracle 称为智能扫描的技术，以优化的方式从磁盘访问数据。这大大降低了数据库必须处理的数据量，并且在数据仓库环境中特别有效。为了提高在 OLTP 环境中的性能，Oracle Exadata 还包括用于缓存热数据的 Flash 存储。

在一个 Exadata 环境中可以运行多个数据库，也可以在一个单一 RAC 节点上运行多个小型数据库，还可以有跨节点大型数据库。这意味着，OLTP 系统与数据仓库实施可以共享一个 Exadata 环境。在 pureScale 环境中不能像这样共享 Smart Analytics System。这是因为 Oracle 完全使用共享磁盘环境，而 IBM 在其 pureScale 实施中对 OLTP 使用共享磁盘，但对在 Linux、UNIX 和 Windows 上的数据仓库却使用了无共享的架构。唯一的例外是大型机 Smart Analytics 产品对 OLTP 和数据仓库/业务分析都提供了一个共享磁盘环境。这可以通过适用于 System z 平台的 9600 的逻辑分区选项很轻松地部署。至于不利的方面，不可否认它是一个小问题，您不能重复利用 Exadata Storage Servers。如果您决定在将来迁移到另一个厂商，那么您可以重新使用您的 RAC 服务器，也可以重新使用您的 IBM 服务器，但使用 Exadata Storage Servers 要实现同样的目的就不那么容易了，原因是它们的特殊功能设计。



## 系统的扩展

Oracle Exadata X2-2 的实施选项说明见

表 1:

	¼ 机架	½ 机架	全机架	2-8 机架
数据库服务器	2	4	8	16-64
Exadata 存储服务器	3	7	14	28-112

表 1: Oracle Exadata 2-2 配置选项

请注意，在升级时，只可以使用这些选项：一个 ¼ 机架可以被升级到一个 ½ 机架，并且一个 ½ 机架可以升级到一个全机架；例如，不允许 ¼ 机架，并且超过全机架时，您只能升级到全机架的整数倍。四分之一机架配置容纳 21Tb 原始磁盘容量，而一个全机架（使用高性能驱动器）包含几乎 100Tb 的原始磁盘容量。如果您使用高容量驱动器，一个全机架的容量是 336Tb。每台 Exadata Storage Server 还包括 4 块 Flash 卡，每块 Flash 卡容量为 96GB，这使全机架的容量扩展为大约 5Tb。请注意，如果不增加额外的磁盘，您就不能向上扩展：这意味着，比如说您有一个 CPU 瓶颈问题，您不能只是添加新的处理能力：您必须增加存储容量，即使您并不需要它。这在数据仓库环境中应该不会成为主要的问题。

当然，在实践中，实际的磁盘容量和可用磁盘容量是完全不同的概念。首先是实现弹性所需要的磁盘镜像，这使可用容量减少了一半，然后是日志、临时空间、索引等所需要的空间方面的考虑。Oracle 自己的估计是，在考虑镜像之前，有 55% 的磁盘容量真正可用于存储数据，这意味着，¼ 机架实际上提供约 6Tb 的可用空间，½ 机架约 14Tb，而全机架则有 28Tb 左右。当然，这些数字还没有考虑压缩。

IBM 提供了六种型号的 Smart Analytics System:

- 1050 – 基于 System x (Linux 或 Windows) 的入门级系统，备有 3 种规模选项，每个选项配有一个数据模块（见下文）、2 个插槽、最多 12 个核心，最大 48GB 内存、磁盘总容量分别为 0.9、4.2 或 7.2Tb、相应的用户容量分别为 0.333、1.65 和 3.3Tb。1050 的标准配置不包括压缩。
- 2050 – 基于 System x 的较大部门级模型，备有 3 种规模选项，每个选项配有一个数据模块、4 个插槽、最多 32 个核心、最大 1Tb 内存和最高 28.8Tb 的旋转磁盘容量（13.2Tb 用户空间）。不包括压缩。这个系统，连同 1050，通常会在不到 100 个用户的环境中部署。
- 5600 – 基于 System x 的中等规模配置，备有 4 种标准规模选项，配 2 到 8 个数据模块（但您的数据模块可以超过该数量），每个模块包含 6 个核心和 64GB 内存。旋转磁盘容量范围是 14.4 到 57.6 Tb，提供（压缩比是 2.5 倍）24 至 96Tb 的总用户容量。这个系统，连同使用 SSD 选项的 5600，通常会在多达几百个用户的环境中部署。

## 系统的扩展

5600 with SSD Option (使用 SSD 选项的 5600) – 基于 System x 的最大型系统, 备有 4 个标准规模选项, 配有 2 到 8 个数据模块, 每个模块 (再次说明, 这不是限制数量) 包含 12 个核心和 128Gb 内存。磁盘容量范围是 21.6 至 86.4Tb。此外, 带有 SSD 选项的 5600 包括 1.28 至 5.12Tb 固态硬盘。压缩后的总可用容量范围是 24 至 96Tb。

- 7700 – 适用于非大型机环境的最强大的 IBM 产品, 这是围绕运行 AIX 的 Power7 服务器所构建的。有六个基本选项, 分别配有 1、2、3、6、10 或 20 个数据模块, 不过同时还有更大型或中间的规模可供选择。每个模块包含 16 个核心和 128Gb 内存。旋转磁盘容量范围是 28.8 到 560Tb, 固态硬盘容量范围是 0.7 至 14 Tb。还可以选择添加更多固态硬盘容量, 最多可增加到 80.4Tb。考虑压缩后的总可用容量范围是 28.75 至 575Tb。据 IBM 称, 该系统将适用于 800 位或更多用户。

- 9600 – IBM 的大型机版本 Smart Analytics System 提供了与 7700 相近的性能, 但在磁盘容量方面, 其规模介于 5600S 和 7700 之间 (较接近后者)。该产品包括配有 DS8700 或 DS8800 磁盘驱动器的 zEnterprise 或 z10 服务器。一大优势是, 从 System z 加载事务数据时可以实现非常低的延迟。

9600 产品的另一个重要特点是, 能够将资源添加到一个现有的 System z, 以支持同一系统上的 9600 逻辑分区 (LPAR)。这降低了部署成本, 并避免在为独立工作负载安装独立的服务器时引入支持问题和费用。今天, Oracle 正成功地运行在 Linux on System z 上, 特别是对整合的工作负载, 但 Exadata 不能运行在 System z 上。IBM 的最新产品是专门针对大表扫描而设计的 IBM Smart Analytics Optimizer, 利用该产品可以扩展 9600。

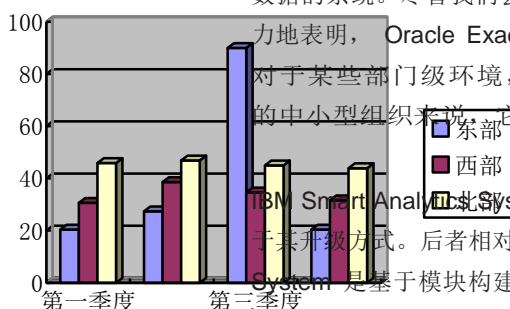
从这些细节中应该明显看出, 与使用 Oracle Exadata 不同, 使用 IBM, 您可以从很小的规模开始: 包含小于 1Tb 数据的系统。尽管我们会在适当时候讨论定价, 但这已有力地表明, Oracle Exadata 的规模不会缩小, 因此至少对于某些部门级环境, 以及对于没有大量数据分析的中小型组织来说, 它将不适合或过于昂贵。

IBM Smart Analytics System 与 Oracle 的不同之处还在于其升级方式。后者相对缺乏弹性, 但 Smart Analytics System 是基于模块构建的, 每个系统包含一个基础模

块、一个或多个数据模块、以及多种可选模块。这些可选模块包括管理员/用户模块、故障切换模块、管理模块和应用程序模块。这里的概念是, 如果需要更多磁盘容量, 那么许可一个额外的数据模块, 但如果需要支持更多用户, 那么许可一个额外的用户模块, 如此类推。换句话说, 只为所需要的东西付费。Oracle Exadata 的情况可能有所不同, 如果您需要更多的 CPU 能力, 您必须有更多的磁盘。

### 压缩

Oracle 使用两种不同类型的压缩: 其一, 被称为“高级压缩”, 专门为 OLTP 环境设计, 在本文的配套文件中将有更详细的讨论; 还有一种被称为“混合列式压缩”, Oracle 在数据仓库和归档环境中使用它。使用基于列的压缩的好处是, 每列都有其自己的数据类型, 因此, 可以针对特定的数据类型优化压缩算法。因此, 从理论上讲, 列式压缩应该是最优压缩, 虽然在实践中的某些算法比其它算法效果更好。然而, 正如其名称所示, Oracle 的压缩并不是纯粹的列式压缩, 而是混合的。这是因为 Oracle 数据库本身并不是列式的, 而是基于行的。这意味着为了压缩, 必须先转换成列, 而这又意味着, 加载时间减慢, 因为数据库处理行中的数据, 这也意味着必须解压缩数据并把数据转换成行格式才可以处理。因此, 虽然压缩可能会显著节省空间, 并导致 I/O 的减少, 但您将损失一些由此获得的性能优势。此外, 您必须牢记, Oracle 压缩的工作方式是, 它有一系列压缩单元。这种方式唯一的问题是, 当您更新一个压缩单元时, 整个单元都会被锁定。所以, 如果要插入一个新行, 那么这将是写入到压缩单元以外的新页。



## Scaling your system

相反，如果删除一行，那么在压缩存储中的相关空间将不能重用（除非您做 reorg），使压缩的效果随着时间的推移逐渐变差。此外，这将导致不稳定的查询处理时间，因为您正在访问的一些数据可能在原来的压缩单元中，而另一些数据则在新的页面（它们没有以任何方式排序，并因此将可能有较差的压缩比），还有另一个原因是在各个压缩单元中夹杂着空块。

IBM 使用一种完全不同的压缩技术：它查找重复的模式（这可能是在同一列中或跨列），然后把每一个这些字符串都替换为由全系统的等价符号串词典支持的符号（标记）。因此，如果您有 5000 个 ‘Milwaukee’ 实例，您在数据库中存储的只是一个 ‘Milwaukee’ 再加上 5001 个符号实例。同样的方法也应用于索引压缩，并且与 Oracle 的方法进行对比是有用的：假设您正在压缩客户索引，则每次

‘Bloor’ 出现时，Oracle 将存储 ‘Bloor’ 加一个行 ID。因此，如果有 250 条 Bloor，就会有 250 个不同的 Bloor-行 ID 对。然而，IBM 只会存储 Bloor 一次，再加上由 250 个行 ID 所组成的一个字符串。因此，IBM 的方法对于索引压缩更有效。

Oracle 的数据压缩在其最佳性能时，会比 IBM 的方法产生更多空间节省。然而，正如我们已经指出，空间节省将会变差。在实践中，Oracle 宣称 3 至 10 倍的节省，而 IBM 则宣称最高 4 至 7 倍。以平均数计算，我们预期 Oracle 在数据方面做得更好。另一方面，Oracle 在压缩临时空间和索引方面处于劣势，所以整体来看，我们预期结果对它不利，或者说更有利于 IBM。不过，还有性能的问题，因为数据已被压缩，所以读取相同数据量时需要较少的 I/O，在这种情况下，由于在 Oracle 环境中所

需的行-列-行转换，我们会期望从 IBM 压缩获得的性能增益（IBM 估计自己在 40%）超过 Oracle Exadata 环境中所实现的性能增益。把这一点添加到节省空间的组合，我们更喜欢 IBM 的方法。



## 性能

Oracle 和 IBM 在过去 12 个月内都公布了 OLTP 处理的 TPC-C 基准测试结果，与之不同，针对数据仓库的基准测试 TPC-H 并没有可比较的结果。IBM 保持 10Tb 系统的当前性能记录，但该结果并不是使用当前版本的 DB2 得出的，并且 Oracle Exadata 没有可用的数据。然而，这不应该引起我们的过份担心，因为基准测试结果只是指示性的。因此，我们将直接继续讨论每个系统中对性能产生有利影响的特定元素。其中一些功能我们已经讨论过，特别是 IBM 在压缩性能方面的优势。

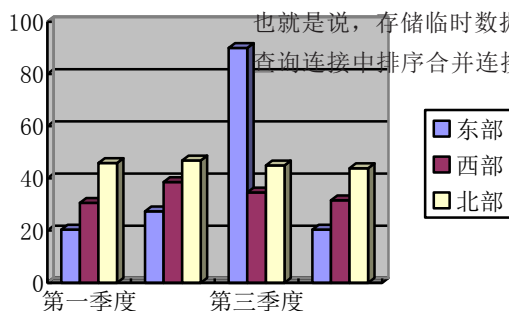
### Flash（固态）存储

正如前面所讨论的，Oracle Exadata 标准配置中配备了 Flash 磁盘，但 Flash 磁盘只是某些 IBM 系统的标准配置。但是，Oracle 把 Flash 的使用主要视为加快 OLTP 处理的工具，而 IBM 则更广泛地使用它。IBM pureScale Application System 标配不带固态硬盘，但您可以选择使用固态硬盘。IBM（配 Easy Tier）和 Oracle 在其使用 Flash 方面有两个主要差异。第一，Oracle 使用 PCIe Flash 卡，而不是固态硬盘。这样做的好处是，在 Flash 存储和处理器之间没有磁盘控制器，如果磁盘控制器尚未设计为以 Flash 速度运作，它就可能减慢环境的速度。Oracle 和 IBM 之间的其他差异是，这两家公司使用 Flash 的方式。Oracle 将其技术称为 Exadata Smart Flash Cache，实际把它用作一个读缓存。也就是说，它把热数据从存储复制到缓存。决定应在 Flash 缓存中保存什么数据是自动处理的，不过用户可以在数据库表中定义指令、索引或分类级别，以确保特定应用程序数据被保存在 Flash 中，但条件限于该软件足够智能，可以知道数据能否放入缓存。

另一方面，IBM 提供了两种方法来使用 Flash（固态硬盘）：

1. 作为 5600 和 7700 的标准配置，它用作临时空间。

也就是说，存储临时数据的地方，例如，在一个更大的查询连接中排序合并连接的结果。当然，多个查询可



能同时使用临时空间，从这些查询随机地访问临时空间比顺序地访问更有效，这就是为什么它会受益于存放在固态硬盘上，因为在数据仓库环境中的磁盘更适合顺序访问。

2 有了在 9600（在适当的时候，中档存储服务器）上可用的 IBM System Storage DS8700，它可以与 IBM 的 Easy Tier 技术配合使用。这里的概念是，有些（热）数据被保存在固态硬盘（SSD 阵列）和传统硬盘驱动器其余部分，在适当的时候，最热的数据向上迁移到 SSD 阵列或（在变冷时）向下迁移到硬盘，数据的重定位（每次可低至 1Gb）由磁盘系统的 Easy Tier 特性自动处理。Easy Tier 提供的功能还可以把单个逻辑卷从一个层动态地重定位到另一个层（例如，从较快的旋转磁盘重定位到较慢的旋转磁盘），并更改卷的 RAID 类型。

Easy Tier 显然比 Exadata Smart Flash Cache 拥有更广泛的功能，并且可更广泛地使用（它与应用程序无关，因此将同时适用于 OLTP 和数据仓库环境）。然而，它目前的可用性有限。对于大多数企业来说，将是在 Exadata Smart Flash Cache 与把固态硬盘用作临时空间之间作出选择。前者是 Oracle 针对 OLTP 环境，而临时空间则在仓库环境中尤为重要，鉴于本报告的上下文，我们必然预期 IBM 的方法将更有用。

### 数据库

在本报告之前，Bloor Research 已于 2003 年、2005 年和 2007 年定期进行 DB2 和 Oracle 数据库系统之间的性能比较。我们不打算讨论那些报告中已经完成的详细情况，否则我们将使这份报告变为两倍的长度！总体而言，每个产品都具有一些我们喜欢的特性，但这些大致平衡。这两家厂商在发布各自的新版本数据库系统时会超越对手，这种情况往往也同样是真实的。这些与数据仓库相关的要

## 性能

素，与 OLTP 不同，我们历来认为两大引擎或多或少可相媲美。例如，我们喜欢 Oracle 的索引，但选择 IBM 的多维数据服务（以支持 OLAP）以及其调优和管理功能。但是，Oracle 已经在减少其管理要求方面取得了重大进展，因此，它已经缩小这方面的差距，但我们仍然认为它落后于 IBM。

不过，在两个数据库内也有一些需要特别讨论的元素。

### 大表扫描

某些类型的查询需要读取整个表。如果这些表非常大，和/或如果复杂查询需要扫描多个表，这可能会对性能造成严重影响。IBM 和 Oracle 都有直接解决这个问题的技术，以及有助于缓解这一问题的辅助技术（在下一节讨论）。

这是作为 Oracle 数据库附加程序的 Exadata 专门设计要解决的问题。简言之，它把处理放在靠近磁盘的位置，以提供一个类似于 MPP 的、无共享的方法。在把结果传递给数据库本身进行最终处理之前，数据流出磁盘并在本地进行预处理。这导致了更快的表扫描并减少 Storage Server 和数据库之间的网络流量。Oracle 使用的实际技术被称为 Smart Scans。虽然在一般情况下，它会显著加快查询处理，特别（但不仅仅）是大表扫描，但它有一个弱点。即，Smart Scans 在一个比数据库更低的层次进行操作，如果对有问题的表进行一个写入操作，Smart Scans 将会被关闭。换句话说，它最适合于静态环境或条件，在静态情况下，定期更新数据是可以接受的。如果在实时或近实时模式下运行，那么将只需要为访问历史数据的那部分查询而保留 Smart Scans。

根据 Smart Analytics System 实施在分布式或大型机环境中，IBM 采用不同的方法：

- DB2 for LUW，当然，是一种无共享数据库，但它并没有像 Oracle Exadata 那样把处理放在靠近磁盘的位

置。从历史上看，它也遇到了大表扫描的性能问题。为了克服这个问题，它提供了一个称为背驮式扫描的工具。这里的概念是，一个查询可以

“背驮”另一个查询。假设查询 a 和 b 都需要扫描一个特别大的表。查询 a 开始。一分钟之后，查询 b 开始。在由查询 a 发起的扫描期间，两个查询都从该表中检索它们需要的数据。同时，查询 b 扫描的那部分表，查询 a 在第一分钟已经扫描过。最终结果有两种：第一，查询 b 将比本来需要的时间早一分钟完成，第二，所使用的 I/O 资源大大减少。

显然，IBM 的大表扫描方法与工作负载相关。如果您有很多用户和查询要访问相同的表，并且如果他们需要完全扫描，那么您使用背驮式扫描将会得到显著的效益。然而，需要全表扫描的复杂查询和其他查询的数量可能相当低，因而背驮式扫描的效益也相应较低。对分布式环境进行总结：Oracle 和 IBM 都没有一个理想的解决方案，任何一方都具有优于对方的可能，这取决于不同的情况和工作负载。

- 利用 DB2 for z/OS，IBM 最近推出了新的工具，通过 Smart Analytics Optimizer 与 zEnterprise BladeCenter Extension 的配合使用，处理大表扫描的问题。这个 Extension（扩展）结合了压缩的列式存储与多并行处理结构。在收到查询时，Smart Analytics Optimizer 将适当地把查询或其部件智能路由到正常的行存储或列式 Extension（后者在大表扫描时特别有用）。在许多方面，这都可以直接与 Oracle 在 Exadata 中所采用的方法相媲美。

### 存储索引和多维集群

采用某些数据库技术的目的是尝试提高总体性能并避免大表扫描的需要。其中最明显的是使用索引，但这些技术会

## 性能

增加磁盘的要求，减慢加载速度，并且当您添加更多索引时，环境会变得更加难以调优。第二种方法是使用分区。虽然 Oracle 和 IBM 对传统索引的支持能力相近，但他们所采用的分区方法一直都有很大差异，Oracle 致力于为您提供一个支持多种选择的灵活方法，而 IBM 则几乎完全专注于散列分区（只是在最近几年才增加了范围分区）。特别要指出的是，IBM 的散列分区实施非常好，但这必须针对 Oracle 的灵活性进行设置。

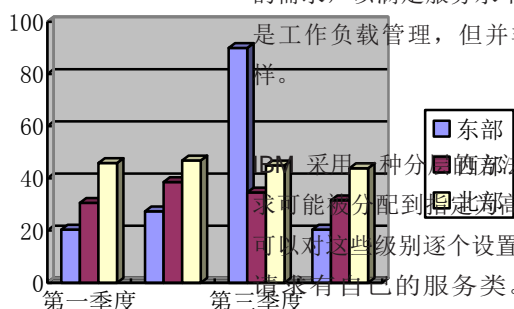
但是，有两个特性（Oracle Exadata 的存储索引和 IBM 的多维集群）与其竞争对手没有可比性。存储索引并非针对数据本身的索引，而是针对存储设备本身：实际上，它们会告诉您每个磁盘块中有（或没有）什么。因此，当您从磁盘中读取时，您就知道要查看哪些块，不需要查看哪些块，并因而节省了相当数量的 I/O，从而提高性能。另一方面，多维集群可以被认为是一种子分区类型。例如，一般情况下，您实施散列分区；然后每月实施范围分区；然后在每个月内，围绕相关的任何维度群集磁盘上的数据。实际上，可以根据集群的程度减少 I/O 并按比例提高性能。

这两种技术都将证明是有用的。但是，如果我们必须选择其中一种技术，我们会选择使用存储索引，因为它们更普遍，并且调优和管理更简单。

### 工作负载管理

典型的例子是，在企业级数据仓库中，有各种不同类型的查询需要同时运行。其中，有一些是低优先级的，另一些具有很高的优先级；有一些需要大量的 I/O，而另一些则需要非常少；有一些需要大量的 CPU 周期，而另一些则几乎不需要；有一些需要在几秒甚至在亚秒内作出反应，而另一些则很可能要花数个小时。问题是，如何平衡这些用户的需求，以满足服务水平协议，并让用户感到满意。答案是工作负载管理，但并非所有的工作负载管理系统都一样。

IBM 采用一种分区方法实现工作负载管理，用户请求可能被分配到指定高、中或低优先级的服务类，您可以对这些级别逐个设置阈值和可分配的资源量。系统请求有自己的服务类。另一方面，Oracle 使用基于



资源组的流程，应用程序被分配给特定节点。对于每个组，可以为某个应用程序设置最大 CPU 分配，设置并行度，等等；也可以把指令传递给 Oracle I/O Resource Manager，它用于保持磁盘（但不是 Flash）有效地执行（例如，防止低优先级任务淹没磁盘）。这种方法的问题是，它的粒度低于 IBM。您可以定义同样的事情，并获得几乎相当的结果，但您将必须定义（和维护）多个资源组，这将增加您的管理负担。

### 在数据仓库环境中的 OLTP

您将不能在与数据仓库相同的环境内正常运行 OLTP，不过，使用 Oracle Exadata，您可以把群集的一部分专用于 OLTP，而其余部分用于数据仓库，IBM 在大型机上也有类似的功能（但实现方式不同）。然而，这并不是说，没有类似于 OLTP 的应用程序在数据仓库内运行。一个典型的例子是主数据管理 (MDM)。如果您仔细考虑 MDM，它实质上是一个带有随机 I/O 而不是顺序 I/O 以及大量查找、更新和插入的 OLTP 式系统。出于这个原因，我们不提倡使用数据仓库来支持 MDM。然而，有一些公司确在他们的仓库中实施 MDM，而 Oracle 和 IBM 都支持此功能。即使没有 MDM，还是会出现具备类似特点的环境，特别是（近）实时的环境，这些环境往往与更传统的查询处理配合。例如，对冲基金可能要针对历史数据与实时股票变动数据运行复杂的查询。当然，这可以通过使仓库与前端 OLTP 系统或一个复杂的事件处理系统联合

## 性能

---

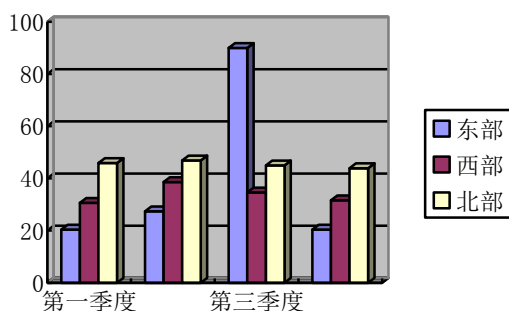
来完成，但有些用户可能更喜欢在一个系统内实现所有这些功能。在这种情况下，混合工作负载管理就显得尤为重要，Oracle 的 Smart Flash Cache 也可能很有意义。

就目前真正的混合环境而言，Oracle Exadata 和 IBM Smart Analytics System 9600 都代表了现实的选择。

## 管理升级

我们已经讨论了两家厂商所采用的一般升级方法，Oracle 提供基于机架的升级，而 IBM 采取一个更传统的姿态，允许根据需要添加额外的容量（内存、硬盘等）。因此，只要不会完全超出所选择的系统的容量，扩大 Smart Analytics System 的容量相对来说比较简单。

另一方面，添加一个新的节点到 Oracle RAC 实施中，意味着配置新节点、安装 CRS、安装 RAC 软件、添加 LISTENER 到新节点、添加数据库软件、手动添加一个 ASM 实例和手动添加一个数据库测试实例。此外，使用 Oracle RAC 时，应用程序需要具备群集感知能力，以优化与集群相关的性能效益，这意味着应用程序可能需要随环境规模扩大（或缩小）而进行调优。



## 行政和管理

DB2 的主要特性之一是它对 Oracle 环境的支持。您可以把 Oracle 架构直接导入 DB2 数据库，DB2 原生（不是模拟的）支持 Oracle 并发控制（但是 DB2 以不同的方式实现，以避免出现在 Oracle 环境中造成性能退化的锁定问题）、SQL、PL/SQL、包、内置包、OCI（Oracle 调用接口）、JDBC、在线架构变更和 SQL\*Plus 脚本，及其他功能。

这一切都意味着，针对 Oracle 数据库编写的绝大多数应用程序、存储过程和其他结构的运行将不变，甚至可能因为针对 DB2 数据库改进的锁定而获得更好的性能。据 IBM 称，它已经测试超过

750,000 行 PL/SQL，并且其平均兼容性已达到 98.43%。这确实令人印象深刻。

我们已经谈过，尽管在最近发布的版本中，Oracle 已经在减少管理要求方面取得了重大进展，我们仍然不认为其自主和自我调优能力可以与 IBM 相匹配。除此之外，也许最大的区别在于实施和高可用性。

### 实施

虽然我们没有关于安装 Oracle Exadata 系统的数据，我们却有 Oracle Database 11g Release 2 在 4 个节点的集群上实施的数据：根据 Winter Corporation 所进行的独立调查，安装这样一个系统涉及 208 个步骤。形成鲜明对比的是，IBM Smart Analytics System 在交付之前，已在 IBM 内完成构建、测试和软件安装。在实际运输过程中，它被拆开，然后工程师将在现场把它重新连接。从那时起，您就可以开始加载数据。换句话说，RAC 有 208 个步骤，而 Smart Analytics System 没有任何步骤，而 Exadata 的步骤很可能比 208 还多。在升级和补丁方面也有相似（虽然不是那么极端）的差异，IBM 所有软件组件的升级和补丁都只是一个单一安装。

### 高可用性

**Oracle 依赖基于软件的磁盘镜像：** 当有一个磁盘发生故障时，该故障由软件自动检测出来，并且它在其它正常的磁盘上重新创建和重新平衡新的镜像。此外，Oracle 提供了一个高冗余选项，让您设置数据的三个副本，这样即使双重故障也不会引起问题。

更普遍的是，Oracle 声称，它已不存在单点故障源。从技术上讲，这是事实。然而，在每台 Exadata Storage Server 中只有一个磁盘控制器，这意味着，如果有一个磁盘控制器出现故障，您必须切换到另一台 Exadata Storage Server。

Smart Analytics System 中的高可用性在不同的平台上亦有所不同。在 System x 架构上，您指定一个高可用性组，把一台备用服务器用作故障转移设备。鉴于 System x 服务器的成本相对较低，这是合理的。另一方面，在 7700 上，如果您愿意，您也可以采用这种方法，较为典型的方法是，把服务器放在有四个节点的组中，再加上一台备用服务器作为服务器组中其余服务器的故障转移服务器。

另外，9600 提供带有共享磁盘的 SMP 架构，让所有处理器都可以访问所有数据，以提供高度可用的环境。9600 提供了工作负载管理功能，让您可以在查询工作负载的同时运行实用程序（备份、reorg、更新等）。当然，并非所有数据仓库都要求高可用性，这就是为什么即使高可用性是标准产品的一部分，IBM 的故障转移模块仍然是可选项。在一般情况下，IBM 通过双适配器、双控制器、双电缆等确保整个服务器的冗余。从软件而不是硬件的角度来看，IBM 还提供了镜像的日志，这将进一步增强高可用性。

## 成本

Oracle 对 Exadata 采用非捆绑式定价结构，而 IBM 则对 Smart Analytics System 采用捆绑的方式。因此，在前者的情况下，您必须单独许可数据库、RAC、分区、高级压缩、调优和诊断包。表 2 说明了不同 Exadata 配置的目录价，包括这些额外组件以及第一年的维护和支持，针对 Smart Analytics Systems 的对比数字，附有相近服务器的性能特征和存储容量配置。请注意，表中的磁盘容量是指可用的未压缩容量，其根据是允许 RAID 格式化后有 55% 的容量这个 Oracle 的既定假设。如表中所示，IBM 在 5600 和 7700 型号之间的容量各有不同。Oracle 的价格不包括 OLAP、数据挖掘或文本挖掘，而 Smart Analytics 产品中包括了所有这些功能。

与 Oracle 的价格相比较，这些 IBM 价格看起来非常有吸引力，并且值得记住的是，即使在您已拥有无限许可协议 (ULA) 的情况下，IBM 的价格已被设置为可与 Oracle 竞争，或比 Oracle 便宜，所以当包括许可时，他们的价格明显低于 Oracle，这并不足为奇。其它型号中的情况也一样：例如，磁盘容量相当于一个 ¼ 机架系统的一台 2050，其目录价只是 \$164,394，但不包括压缩。

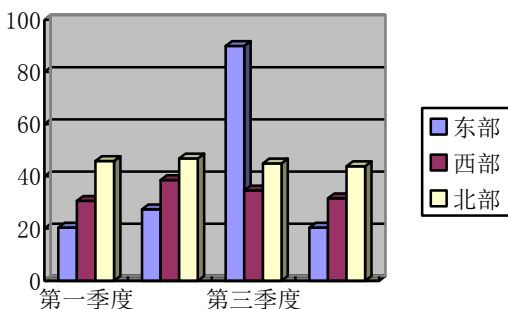
当然，必须说明的限制条件是，这些都只是目录价，实际可能会有可观的折扣。

另外，在考虑成本时必须牢记两点。第一，升级成本。使用 Oracle，您只能从 ¼ 机架升级到一个 ½ 机架、全机架，然后是添加更多机架。例如，您不能在存储服务器之外单独添加更多数据库服务器（反之亦然），而存储服务器许可成本是每个磁盘驱动器 \$10,000（加上 22% 的维护费用），如果您只需要更多计算能力，这是一个昂贵的选项。相反，升级 Smart Analytics System 则没有这些限制。然而，这样也的确有缺点，任何特定的系统将只能扩展到其最高容量，所以有可能会发生型号升级的情况，比如说从 2050 更换为 5600。

其次，如果我们的判断是正确的，DB2 比 Oracle 更容易管理，而 Smart Analytics System 环境比 Exadata 更容易管理，那么我们预期后者要求的管理比 IBM 所需的管理更多。这本身就代表了一种成本。

Oracle 系统	Oracle 标价	IBM 配置	IBM 5600	IBM 5600 with SSD	IBM 7700
¼ 机架 (6Tb)	\$2.32m	3 个数据模块 (10/11Tb)	\$1.15m	\$1.74m	\$2.68m
½ 机架 (14Tb)	\$4.73m	6 个数据模块 (20/22Tb)	\$1.56m	\$2.65m	\$3.73m
全机架 (28Tb)	\$9.30m	12 个数据模块 (30/33Tb)	\$2.14m	\$3.68m	\$4.69m

表 2: 同等配置的标价比较



## 结束语

---

当我们着手研究时，我们预计会发现 IBM 精于某些领域，而 Oracle 则擅长另一些领域。令我们惊讶的是，情况并非如此，IBM Smart Analytics System 几乎在我们测试过的所有领域中都优于 Oracle Exadata。唯一例外的是性能。为了实现预处理，把处理放在靠近磁盘的位置，并与 Oracle 存储索引的使用结合，这意味着将会有某些类型的查询，以及因此带来的某些类型的环境，在这些环境中，Oracle 的性能将优于分布式平台上的 IBM。然而，这些环境相对比较特殊，并且以复杂分析为重点。[注意：虽然这超出了本文的范围，但 IBM 最近收购了 Netezza。正是在这些领域中，我们预期 Netezza 将以明显的优势战胜 Oracle]。

在大型机上，情况可能有所不同：有了 Smart Analytics Optimizer 和 zEnterprise BladeCenter Extension，我们看不出有任何理由 Oracle 会 有任何性能优势。

此外，在其他类型的环境中，以及有混合工作负载的要求的环境中，我们没有看到这么明显的性能问题，IBM 将会在许多领域中优于 Oracle。换句话说，我们不能一概而论地对两家公司之间的比较作出一个有现实意义的判断：这将取决于您的具体情况。

在所有其他方面，从可扩展性到灵活性，一直到易用性和高可用性，再到成本（至少以目录价格计算），IBM 似乎都提供了明显的优势。

### 更多信息

关于本主题的更多信息可从以下网址获取  
<http://www.BloorResearch.com/update/2069>



## Bloor Research 概述

Bloor Research 是欧洲领先的 IT 研究、分析和咨询机构之一。我们阐明了如何通过有效地治理、管理和信息利用为企业的 IT 系统带来更大的敏捷性。我们已经通过独立、明智、表达清晰的交流内容以及关于 ICT 行业各个方面的出版物树立了“准确精辟”的美誉。我们相信，准确精辟的目标是：

- 说明技术及其相关业务价值，还有与其交互的其他系统和流程。
- 了解如何使新技术和创新技术配合现有的 ICT 投资。
- 观察整个市场，并阐述所有可用的解决方案，以及如何可以更有效地评估它们。
- 过滤“噪音”，使发现更多支持投资和实施的信息或新闻变得更容易。
- 确保我们所有的内容都可通过最适当的渠道提供。

成立于 1989 年，我们花了超过二十年时间，通过网上订阅、定制研究服务、活动和咨询项目向全球 IT 用户和供应商组织分发研究和分析。我们致力于把我们的知识变成您的商业价值。

## 关于作者

**Philip Howard**  
研究总监 - 数据



Philip 早在 1973 年就踏入计算机行业，已在包括 GEC Marconi、GPT、Philips Data Systems、Raytheon 和 NCR 在内的多个公司内担任过系统分析员、程序员和销售人员、以及营销和产品管理等多个岗位。

在别人名下工作了 25 年之后，Philip 于 1992 年成立了自己的公司，即现在的 P3ST (Wordsmiths) Ltd，他的第一个客户是 Bloor Research（当时是 ButlerBloor），Philip 以助理分析师的身份为该公司工作。他与 Bloor Research 之间的关系从那时开始一直持续至今， he 现在是研究总监。他的专长范围包括与数据和内容有关的任何领域，在这方面与他一起工作的还有五位分析师。除保持对全局的把握以外，Philip 本事的专长是数据库、数据管理、数据整合、数据质量、数据联合、主数据管理、数据治理和数据仓库。他在事件流/复杂事件处理方面也有兴趣。

Philip 除了代表 Bloor Research 编写大量报告之外，他也定期在 [www.IT-Director.com](http://www.IT-Director.com) 和 [www.IT-Analysis.com](http://www.IT-Analysis.com) 上发表文章，并且曾经代表 Cambridge Market Intelligence (CMI) 出任“Application Development News（应用程序开发新闻）”和“Operating System News（操作系统新闻）”的编辑。他还在各种杂志上发表文章，并通过 CMI 和 The Financial Times（金融时报）等公司出版了一系列报告。

在工作之外的时间，Philip 的主要休闲活动是运河船、滑雪、打桥牌（他是该项目的终身大师）和遛狗。

## 版权及免责声明

---

本文件版权所有 © 2010 Bloor Research。未经 Bloor Research 事先同意，不得以任何方式复制本出版物的任何部分。

出于本材料的性质，文中多处提及硬件和软件产品的名称。绝大多数情况下，这些产品名称都由制造该产品的公司声明为商标。Bloor Research 无意声明这些名称或商标属于自己。同样，所复制的公司徽标、图形或屏幕截图均经过所有者的准许，其版权属于该所有者。

虽然在编制本文件时已力求审慎，以确保信息的准确性，但出版人接受对任何错误或遗漏承担责任。



2nd Floor,  
145–157 St John Street  
LONDON,  
EC1V 4PY, United Kingdom

电话: : +44 (0)207 043 9750

传真: +44 (0)207 043 9748

网址: [www.BloorResearch.com](http://www.BloorResearch.com)

电子邮箱: [info@BloorResearch.com](mailto:info@BloorResearch.com)