

Oracle Warehouse Builder 10g

升级至10g 版本的三个业务原因

*Oracle 白皮书*

2004 年 2 月



# 升级至 10g 版本的三个业务原因

## 执行概要

许多软件产品都在不断地进行更新和扩展，Oracle Warehouse Builder 10g 亦是如此。对于客户而言，在决定升级路径之前，首要之事是要了解新的版本可以带来哪些益处。

本白皮书将通过演示 Warehouse Builder 10g 版本中的新功能，为客户的升级提供指导性意见。其中将着重介绍客户通过使用 Warehouse Builder 10g 代替旧版本而可以获得业务利益。请注意，Warehouse Builder 9.2 版与 10g 版功能基本相同，只是 Warehouse Builder 10g 还受数据库 10g 版的支持并通过了其认证。除此之外，本白皮书中的讲述的 10g 的功能基本都与 9.2 相同。

Warehouse Builder 10g 在业务方面的主要增强功能有：

- 通过高级的嵌入式数据质量提升了数据质量
- 提高了调试和数据查看功能中提取、转换和加载过程的生产率
- 增强了元数据比较，从而可对元数据进行扩展而无需进行重新开发
- 增强了版本管理，允许针对开发人员和设计人员进行对象级版本控制。

对于所有希望提高其生产率并改善其数据质量的客户而言，新版本中针对这些业务问题做出的主要改进使得 Warehouse Builder 10g 成为了一款引人注目的产品。利用该产品，仓库系统的实施质量更高、速度更快。

## 引言

作为 Oracle 商务智能产品的基石，Oracle9i Warehouse Builder 解决了各种分散数据源与目标之间复杂的数据集成问题。Oracle9i Warehouse Builder 是一种为 Oracle 商务智能解决方案设计和管理数据与元数据集成的工具。

如今，许多企业都了解集中数据和元数据存储的必要性。集中存储使得公司能够快速做出更好的决策。

在部门级，数据仓库将划分为第 1 阶段或部门即席、实施，这表明企业已了解需求并起草了解决方案。多数第 1 阶段项目面临着增长问题。有一些成功地克服了困难，业务需求超过了新信息的提供。而另一些项目则受困其中，产生许多其无法解决的问题。

因此，许多首次尝试都是力图在部门级或数据仓库的第 2 阶段中开发出可靠、持续以及扩展的信息提供机制。

许多尝试进入数据仓库第 2 阶段的项目都面临着以下问题：

- 由于下列原因造成的实施生产率低下：
  - 开发人员效率低下以至于无法满足获取新扩展和信息的积极需求
  - 元数据更改功能的不足降低了交付速度
- 解决方案中的低信任度导致数据质量较低
- ETL、模式设计以及最终用户工具之间的元数据分散

为了从第1 阶段即席实施前进至第 2 阶段策略性实施，客户必须解决上述问题。

要前进至第 3 阶段 — 即，直接影响企业或策略性仓库关键所在的数据仓库实施，客户必须确保其数据精确且有价值。换言之，客户的数据有价值，利用这些数据可以降低成本，或提高企业的盈利能力。

Oracle Warehouse Builder 的 10g 版本增加了特定功能，可解决上述这些问题甚至涉及范围更为广泛的问题，使得客户能够从第1 阶段前进至第 2 和第 3 阶段。

## 提升实施生产率

要提升整体的生产率，首要考虑的是开发周期中最耗时的任务。毫无疑问，数据仓库项目中最为耗时的部分是提取、转换和加载 (ETL) 例程的开发。

若能够降低在项目的该阶段中投入的时间和精力，则整体生产率会得到最大限度的提高。第二个需要改善的领域是开发人员处理数据和元数据更改的能力。

### Warehouse Builder 10g 是如何缩短 ETL 过程耗费的时间的？

ETL 开发中难度较大的部分是创建将数据从数据源移至目标系统的转换。很多时候转换操作比较复杂，需要花费大量的时间进行开发。为了缩短开发时间，Warehouse Builder 引入了 Mapping Debugger，从而使得开发人员可以调试其逻辑 ETL 过程。

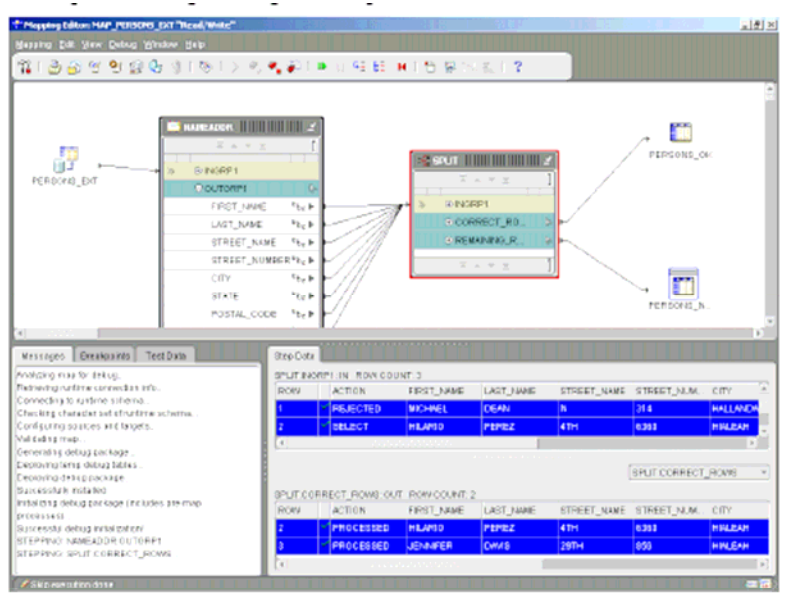


图 1 调试复杂的映射

开发人员已经可以在图形用户界面中创建一组转换，但是，借助调试器，他们还可以逐步演练过程并查看他们构建的个别数据行的效果。这一功能提升了 Warehouse Builder 中 ETL 设计的生产率。与 ETL 手工编码相比，映射调试器向前迈进了一大步。

热爱数据仓库的专业人士会说，市场上提供此种功能的工具比比皆是。这是事实，然而，这些工具中没有一个是能够提供 Warehouse Builder 所具有的那种细粒度调试。

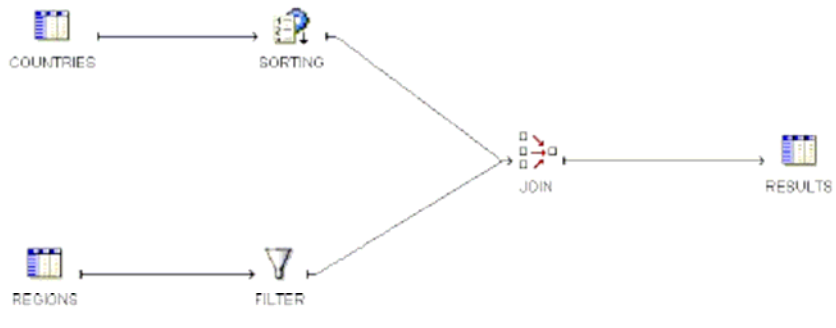


图 2 映射编辑器中的细粒度建模

与其他图形化 ETL 工具相比，Warehouse Builder 按照活动提供建模，而不是根据 SQL 语句。例如，在查询中，不会对一组源表上的联接和过滤操作进行建模，而个别的操作符允许开发人员查看有关其数据的单个以及组合操作效果。因为独特的设计，Warehouse Builder 可以查看 Warehouse Builder 中建模的查询，并提供强健的调试结果。

### Warehouse Builder 10g 如何协助管理更改？

任何仓库开发人员都知道，需要持续更改和更新数据仓库以反映新的或不断变化的业务要求。

为了解决这一需求，Warehouse Builder 从一开始就具备了支持复杂更改管理的功能，例如增量元数据部署和高级线性和影响分析。

随着 Warehouse Builder 10g 版本的推出，又新添了一个仓库元数据管理维度 — 检查多个元数据对象版本之间的差异并在更改前按照这些数据进行操作。

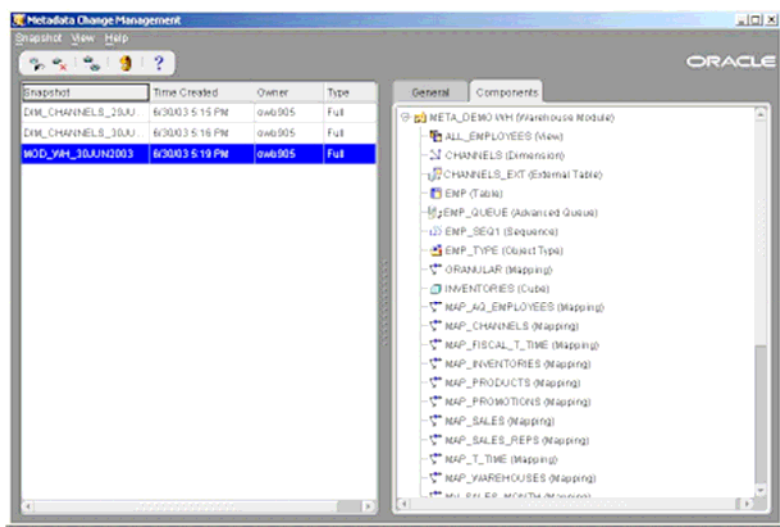


图 3 更改管理器

这样，Warehouse Builder 就杜绝了版本管理中的不确定因素，使得开发人员能够利用新的或现有的版本更新其元数据。

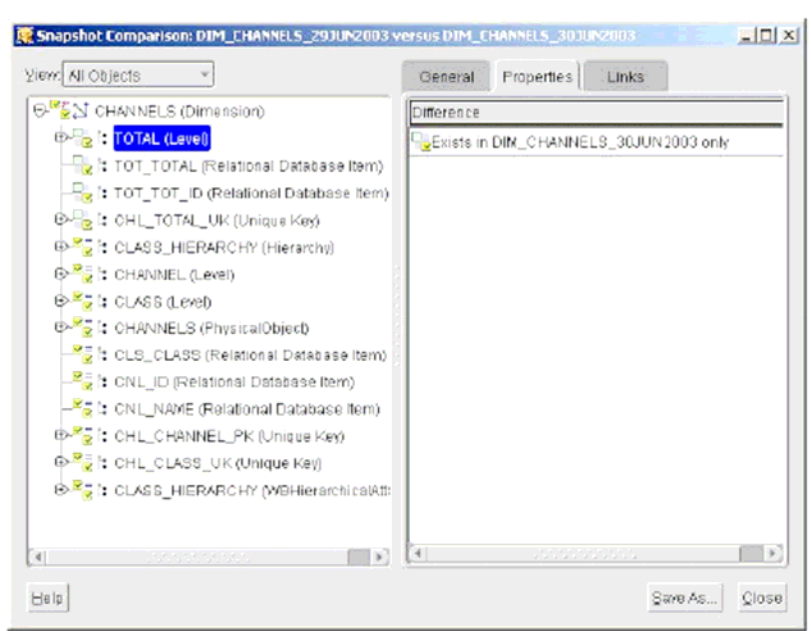


图 4 查看两个快照之间的差异

在原有广泛的更改功能之上新添了这一功能后，使得 Warehouse Builder 在更改管理方面独树一帜。

## 通过改善数据质量提升信任度

在 Warehouse Builder 的之前版本中，Oracle 引入了集成的高级名称和地址 (Name and Address) 清理功能。名称和地址清理可使重要数据 (客户数据) 质量更高。随着 10g 的推出，通过在名称和地址清理映射表中合并映射调试功能，用户可以获得更多的功能。

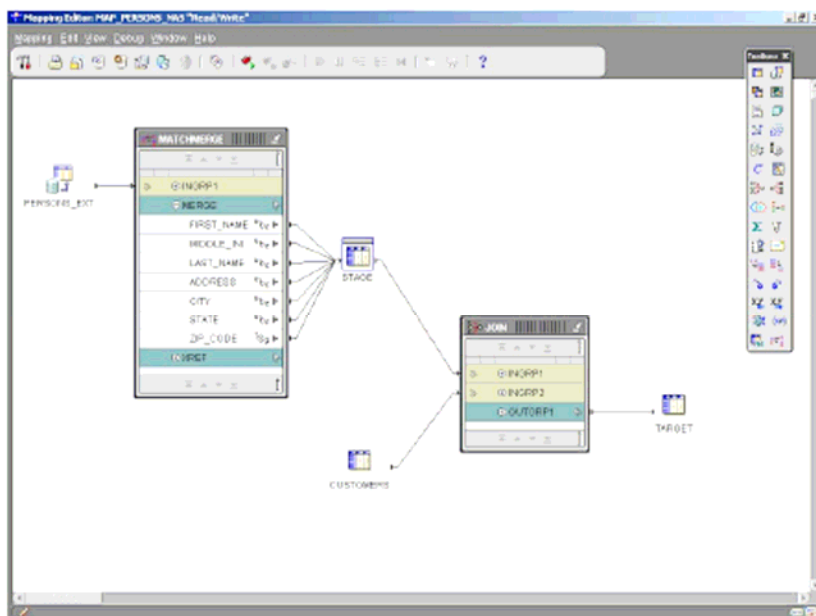


图 5 匹配与合并数据

然而，这就是客户数据的数据质量。如前所述，非常重要的一点是，数据质量的许多其他方面并不包含在客户数据内。例如，产品列表，通常也如客户数据一样很重要，但也不太清晰明了。

### Warehouse Builder 如何帮助在非客户数据中创建更高的信任度？

为了协助客户处理没有按照名称和地址排列的关键数据，Warehouse Builder 10g 提供了高级匹配与合并功能。

与名称和地址清理不同，顾名思义，匹配/合并使您可以在或多或少非结构化的数据中查找重复项。一旦标识出这些重复项，您就可以将其合并至单个实例的行中。这样，对于业务社区而言，您的数据将更为精确，聚合与预测更为准确、更有价值。

Warehouse Builder 中的匹配/合并实施允许您指定自己的匹配与合并规则，在同一用户界面中处理简单和高级用例。

## Warehouse Builder 的覆盖范围能遍布全球吗？

正如您对企业级工具的期望一样，Warehouse Builder 能够包含全世界的名称和地址。为此，Warehouse Builder 的选择与多数供应商背道而驰。

Warehouse Builder 允许您和您的当地数据供应商将特定的国家/地区信息插入到 Warehouse Builder 工具中，而不是将您限定在专有技术中。这种方式比较灵活，您可以选择最符合贵公司名称和地址要求的数据供应商<sup>1</sup>。

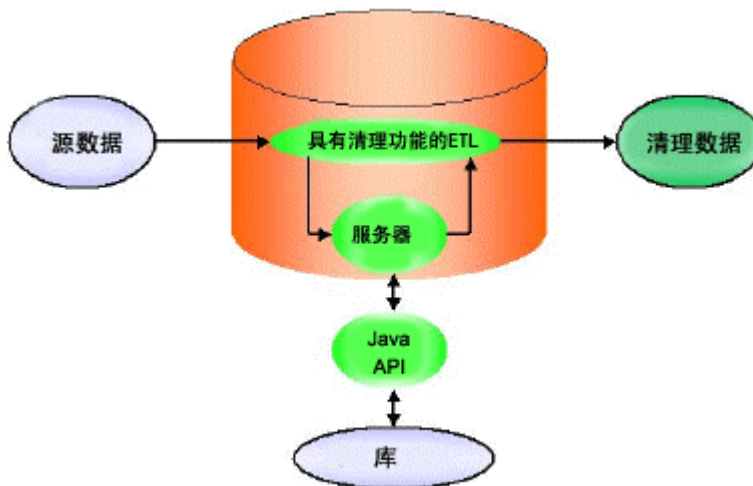


图 6 提高数据质量的 API 体系结构

这样也可提升您的开发人员的生产率，他们可以在熟悉的、用户友好的 Warehouse Builder 用户界面中进行开发，同时仍能获得专用的数据清理支持。

### 提升操作价值

一旦在数据和数据仓库中完成了上述步骤，就可开始前进至仓库开发的第 3 阶段了。请注意，在第 3 阶段，仓库对于日常业务而言不可或缺并具有战略意义。

### Warehouse Builder 如何帮助创建操作价值？

该功能中最新增加的一项是分类管理 (Householding)。一旦您控制了数据质量，就可以使用数据实现增值，例如降低业务成本。

分类管理是一种有效的方法，通过向个别记录添加含义来提升客户信息价值。分类管理就是将个别记录进行分组。换言之，将成员放置在应属的逻辑组中。

---

<sup>1</sup> 请访问我们的网站，查看当前的合作伙伴，或登录 OTN 技术论坛与我们联系，获取更多信息。

为什么这与您密切相关？简而言之，这样确保您只能将一份材料副本发送至一家（而不是将其发送至同一家的 3 个人），降低了发送至一定人数所需的金额。这是个显而易见的优点，然而，它也有不利的一面。有些时候，一个共享的地址可能只表示对方是室友关系，而实际上您希望发送多个副本来提高您的成功几率。

令人称赞的是，Warehouse Builder 在其 10g 版本中合并了这一高级的重要功能，利用了您在技术上的投资，同时又向您提供了增加仓库操作价值的方法。

## 结论

本白皮书中介绍的 Warehouse Builder 新功能只是管窥豹斑，但该产品会为面临数据和元数据集成问题的客户带来巨大的商业价值。

通过在数据仓库开发中运用 Warehouse Builder，您可以解决企业中许多常见问题。10g 版专门针对以下问题：

- 低下的实施生产率
- 解决方案中的低信任度导致数据质量不高
- ETL、模式设计以及最终用户工具之间的元数据分离

解决了上述这些问题，对于贵企业而言，您的项目将更为成功、效率更高，价值更大。



升级至版本 10g 的三个业务原因  
2004 年 2 月  
作者: Jean-Pierre Dijcks

Oracle Corporation  
全球总部  
500 Oracle Parkway  
Redwood Shores, CA 94065  
U. S. A.

全球咨询热线:  
电话: +1. 650. 506. 7000  
传真: +1. 650. 506. 7200  
[www.oracle.com](http://www.oracle.com)

Oracle 是甲骨文公司的注册商标。本文中提及的各种产品和服务的名称可能是甲骨文公司的商标。其他所有提及的产品和服务名称可能是各自所有者的商标。

版权所有 © 2004 Oracle 公司。保留所有权利。