

Oracle9i中数据仓库的增强及其价值 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/142/2021_2022_Oracle9i_E4_B8_c102_142876.htm 数据仓库需要从各种不同的数据源取得各种不同的数据，并且把这些巨大数据量的数据转换成对于用户可用的数据，为企业的决策支持提供数据支持。这个过程常常被称为ETL（提取、转换、装载）。提取过程涉及把数据从不同的来源提取出来，比如，一些服务提供商需要从上百个网站提取数据，然后生成用户可用的数据。这个过程中，最最消耗时间的部分是转换和装载数据这两个步骤，在这些步骤中，要根据现有数据指定规则，然后按照这些规则来过滤数据，并且把合格的数据装载到数据仓库里边去。而这个步骤地最大困难就是要尽量不影响数据仓库和源数据库的性能，并且对于不同的数据量能够有一定的可伸缩性，并且在最短的时间内完成。当前绝大多数的ETL步骤都是通过第三方工具来进行的。这些工具能够在把数据转换并装入到数据仓库之前，对数据进行一些特定的处理。当数据转换完成以后，再用Oracle的并行插入和装载工具把这些数据插入到Oracle数据库。Oracle数据库的最主要的作用是管理这些数据行、索引和约束。有些ETL过程是串行进行的，因此需要使用更多的数据库资源来进行这些转换和装载过程。如果使用这类串行装载方法的话，首先利用一些第三方工具，数据先被提取出来，然后放到一个中间过渡区域里边，在使用PL/SQL或者java再在数据库里边进行转换，最后再把结果插入到数据库的表里边。这个复杂的过程导致了这种方法不可避免的弱点：可伸缩性差、在万一出现差错时难以控制。

Oracle9i 引进了新的"边装载边转换"的办法来取代那些过时的串行处理步骤：先转换然后装载或者先装载然后转换。在这种新方法里，数据库参与了数据转换和装载的过程，成为了ETL过程的一个有机组成部分。而另外有些原来是必须的步骤则没有继续存在地必要了，另一些则可以得到改进。

Oracle 9i提供以下功能来帮助这个转换、装载步骤更加快速而高效。Oracle Change Data Capture (OCDC) Framework 可以用来优化ETL过程中的数据提取这个步骤，建立一个可重复使用的执行步骤。OCDC能够捕获Oracle数据库中变化的数据。使用对称复制技术和Logminer技术，如果原始数据是来自Oracle数据库的话，那么Oracle能够很容易实现几个数据库的改变了的数据的同步化和异步化工作。对于非Oracle数据库，Oracle公司也提供了API，帮助第三方软件厂商开发工具，来实现非Oracle数据库和Oracle数据库的集成。External Tables 这个特性能够使一些外部数据源（比如一个普通文本文件）能够像存储在Oracle数据库普通表里边的数据一样被透明地存取。这个特性的唯一限制是：这张表是只读的，不能修改。对于SQL，PL/SQL和Java程序来讲，这些虚拟表不用首先被装载到数据库里边然后才能读取。这样的话，装载和转换步骤就被集成到一起了，不再需要在数据库里边存储中间数据。

Multi Table Insert 这是一个任何DBA和开发人员都会欣赏的新特性。在8i和以前的版本里面，想要把同样的数据插入到多张表里边，你不得不编写一个存储过程，或者执行多次SQL语句。在Oracle 9i里边，这些都可以不用了。一个SQL语句就可以解决这个问题。另外一个值得注意的新特性是Upsert语句，它用一个语句提供了Update和Insert的功能，根据条件分别

执行Insert或者Update语句。还有一个很有用的增强就是允许在一个Oracle数据库里边有多个Block的大小。这对于在不同Oracle数据库之间使用可移动表空间这个特性是非常关键的。从9i开始，可移动表空间的大小不一定需要是同样的块大小了。有了以上这些新特性以及其他一些本文没有提及的新特性之后，Oracle不仅仅适用于数据处理，也更加适用于数据仓库环境下的使用。100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com