

Oracle 数据仓库解决方案概述

Summarize of Oracle

Warehouse Solution

摘要: 本文介绍了用 Oracle 的产品来做数据仓库的设计与开发要完成数据仓库的设计建模、数据转换与集成、数据存储与管理、数据的分析和展示、数据仓库的维护和管理五个过程。

关键词: 数据仓库 设计建模 Oracle 元数据

1 引言

Oracle 数据仓库的实施是一个相当复杂的过程, 主要包括五个过程: 数据仓库的设计建模、数据转换与集成、数据存储与管理、数据的分析和展示和数据仓库的维护和管理。Oracle 提供了覆盖整个数据仓库建立周期的一套完整的产品包: Internet Developer Suite, 它包括数据仓库的建模、数据集成和转换、数据存储和管理、元数据管理和数据可视化分析等产品。以下是说明如何结合 Oracle 的产品来做数据仓库的设计与开发的。

2 数据仓库的设计、建模

数据仓库的设计工作对于决策支持系统起着至关重要的作用。它需要根据决策需求确定主题, 从数据源到数据提交, 对数据仓库的数据组织进行逻辑结构的设计, 还要按照业务用户最能理解的方式组织和提供信息。

在这个阶段, 我们使用了 Oracle Warehouse Builder。Oracle Warehouse Builder 是个高度优化的数据库工具, 广泛用于数据源的逆向工程、建模、数据仓库方案设计, 以适应每个业务需求。

通过对逻辑设计、物理设计和应用建模进行集成, Oracle Warehouse Builder 方便了数据仓库的开发和实现。

3 数据转换与集成

在进行数据仓库的建立时, 最大的挑战之一是如何将原始业务数据转化为一致的格式, 使之更好地为决策支持服务。这包括对已有数据的准确性 and 一致性进行检验、净化, 将数据进行转化、提取、转换、装载到数据集市或数据仓库以及对其进行定期更新和管理。Oracle Warehouse Builder 作为数据抽取工具, 从各种异构的数据源中抽取数据, 在数据抽取过程, 用户可以根据不同的抽取阶段, 灵活定制各种数据抽取流程, 并定时地将数据加载到数据仓库中。

Oracle Internet Developer Suite 是一个集成的软件产品套件, 用于建造和管理数据集市和分析应用。Oracle Internet Developer Suite 交付了一个开放的可伸缩的解决方案, 主要定位于数据集市完整的生命周期和分析应用开发及产品化的管理, 能够支持多种平台上快速变化的大量数据

作为数据来源, 进行复杂的转换处理以及支持高速的数据加载。其 metadata repository 能够协调并驱动一系列的核心功能, 包括抽取、转换、加载和管理等。

Oracle Warehouse Builder 的图形化用户接口帮助数据仓库管理人员很容易的设计复杂的 source-to-target 的映射, 然后可以由 Oracle 强大的服务器引擎来自动地执行。

4 数据存储与管理

数据仓库的存储可以选用多维数据库, 也可以选用关系型数据库或其它特殊的存储方式。数据的存储要保证数据的安全性、完整性、一致性, 同时还要具有复杂的分析查询的高效性。

我们选用了 Oracle 的数据仓库级产品 Oracle 9i Database Enterprise Edition。Oracle 9i Database Enterprise Edition 是一个关系型数据库, 为高性能决策支持和数据仓库的建立而进行了优化。IQ 中的关键技术是纵向数据存储(通过列而不是通过行来进行)、Partitioning Table、及 Partitioning Index 查询索引和数据压缩。

5 数据分析和展现

联机分析处理(OLAP)是一个分析处理技术, 它从企业的数据集中收集信息, 并运用数学运算和数据处理技术, 灵活、交互式地提供统计、趋势分析和预测报告。通过多种 OLAP 工具对数据仓库中的数据进行多维分析、汇总, 形成图表或报表的形式, 使决策者可以清晰、直观地看到分析结果, 这正是数据仓库系统所要达到的目的。OLAP 技术已经整合在 Oracle 9i Database Enterprise Edition 中。

数据仓库的开发应用主要有结构设计、数据集中组织和管理、数据的快速高效访问等。其中数据的访问一般都是由较为成熟的业务智能工具完成, 因此不同于 OLTP 系统, 数据仓库系统的
(文章未完第 22 页右下角)

前端开发编程量是比较小的，但是其维护工作的时间跨度要大，因为决策支持应用的随意性较强，不可能再象业务系统那样固定一个统一的操作模式。

Oracle 作为较早进入中国市场的业务智能提供商，其产品操作精简、功能丰富，并且有直观易懂的前端展现元数据管理部分，在这个解决方案中与 IQ 的高速查询效率相得益彰。

6 数据仓库的维护和管理

元数据是关于数据的数据，能够表示、定义数据的意义及系统各组成部件之间的关系的的数据，它包括关键字、属性、数据描述、物理数据结构、源数据结构、映射及转换规则、综合算法、代码、缺省值、安全要求及数据时限等。管理好元数据是管理数据仓库的关键。

Oracle 的 Warehouse Builder 通过对元数据仓库的集中管理，提供了数据仓库解决方案的保证技术。从设计和开发到实现到最终用户访问，由工具和数据库产生的对元数据的密集型集成和管理保证了真正企业级数据仓库的建立。Warehouse Builder 为数据仓库开发人员提供的数据库元数据管理工具，能够在数据仓库环境下进行数据采集、捕捉、存储、管理和发布逻辑的、物理的以及上下文相关的信息，而不用去管它的物理存储位置是在联合数据仓库上、分布式数据仓库上还是二者兼有。业务用户可以浏览根据其需求而生成的元数据对象，甚至可以使用发布和登记性能请求或选择附加性能。

7 结论

结合 Oracle 的最新系列产品 Oracle 9i Application Server 中的 Portal 技术以及 Internet Developer Suite 中的 Business Intelligence 的核心技术进行整个数据仓库系统的改进，强化查询数据的最终展现功能以及数据挖掘分析的功能。整个系统全面采用了 Oracle 的全线产品，发挥了其提出的统一的系统架构、集中的数据存储、整合的业务流程处理的完整的企业电子商务的解决方案。■