

Oracle Data Integrator 技术白皮书：产品体系结构总揽

*Oracle*白皮书
2006年12月

ORACLE®

Oracle Data Integrator 体系结构

介绍

Oracle Data Integrator在多项组件上得到构建，且这些组件围绕同一集中式元数据库共同工作。上述组件--图形模块、运行时组件以及Web界面，通过与其他高级特性的结合，使Oracle Data Integrator成为一种代表当前最先进技术水平、非传统性（Legacy-free）的轻量级集成平台。本白皮书将对Oracle Data Integrator的体系结构予以详细描述。

体系结构总揽

Oracle Data Integrator体系结构的组织主要是围绕一种模块化存储库展开，通过完全在Java中编写的图形模块和执行代理（程序）、以客户服务器模式对该存储库进行访问。

Oracle Data Integrator体系结构的组织主要是围绕一种模块化存储库展开，通过完全在Java中编写的图形模块和执行代理（程序）、以客户服务器模式对该存储库进行访问。该体系结构还包括Web应用以及Metadata Navigator，它们可以使用户通过Web界面访问信息。

图形模块

四种图形模块分别是Designer、Operator、Topology Manager、以及Security Manager。这些模块可以被安装在所有支持Java Virtual Machine 1.5（J2SE），包括Windows、Linux、HP-UX、Solaris、AIX、以及Mac OS在内的图形平台上。

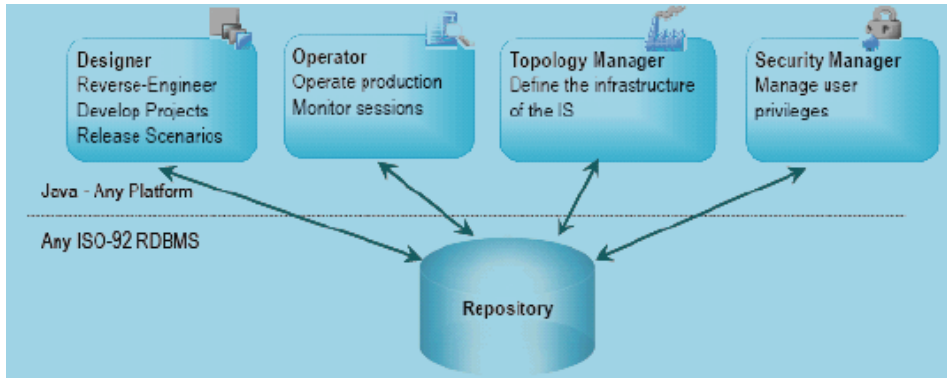


图1：连接存储库的图形模块

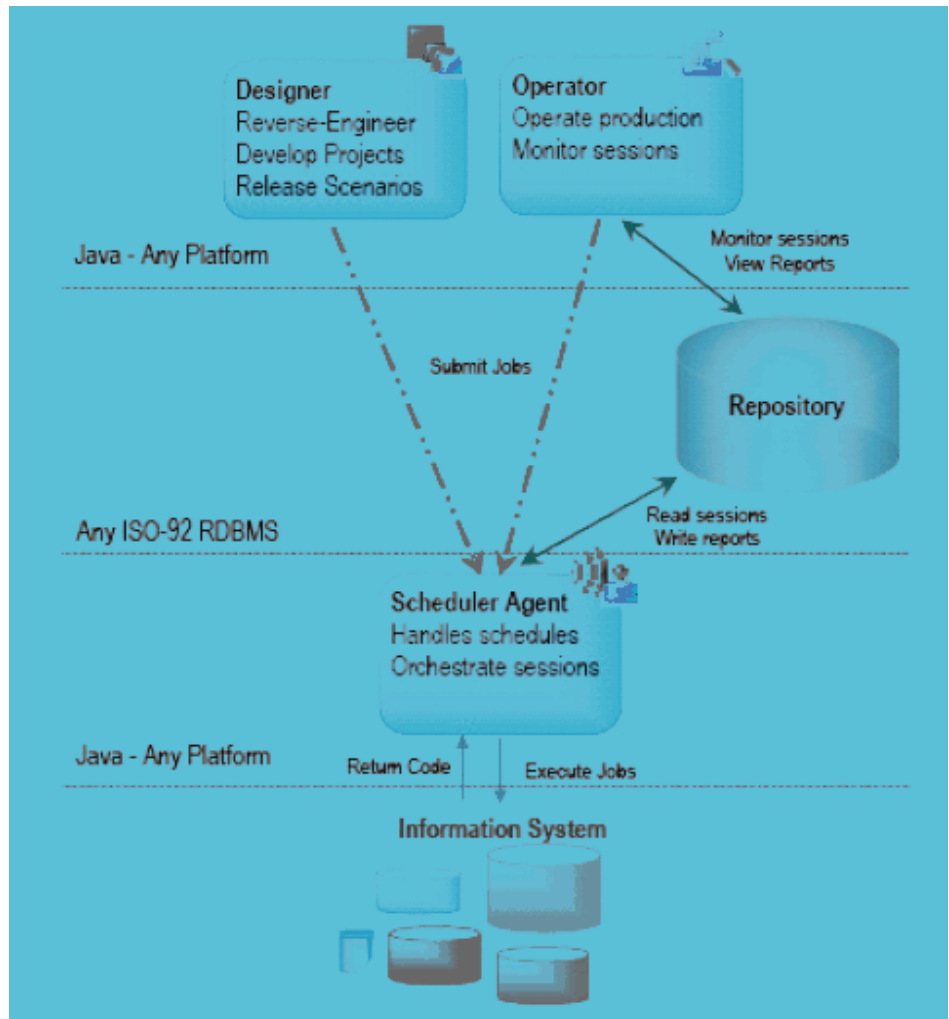


图2：运行时组件

存储库

存储库由一与主存储库和几个工作存储库组成。这些存储库是存储在关系数据库管理系统中的数据库。模块配置、开发或使用的全部对象被存储在其中的一个存储库中，并且通过各种体系结构的组件、以客户服务器模式得到访问。

通常，存储库中仅有一个主存储库，其中包含安全信息（用户资料及权限）、拓扑信息（技术及服务的定义）、以及目标的各版本。利用Topology Manager以及Security Manager使包含在主存储库中的信息得到保留。由于上述模块中均存储拓扑和饿安全信息，所以，它们都具备对主要储存库的访问权限。

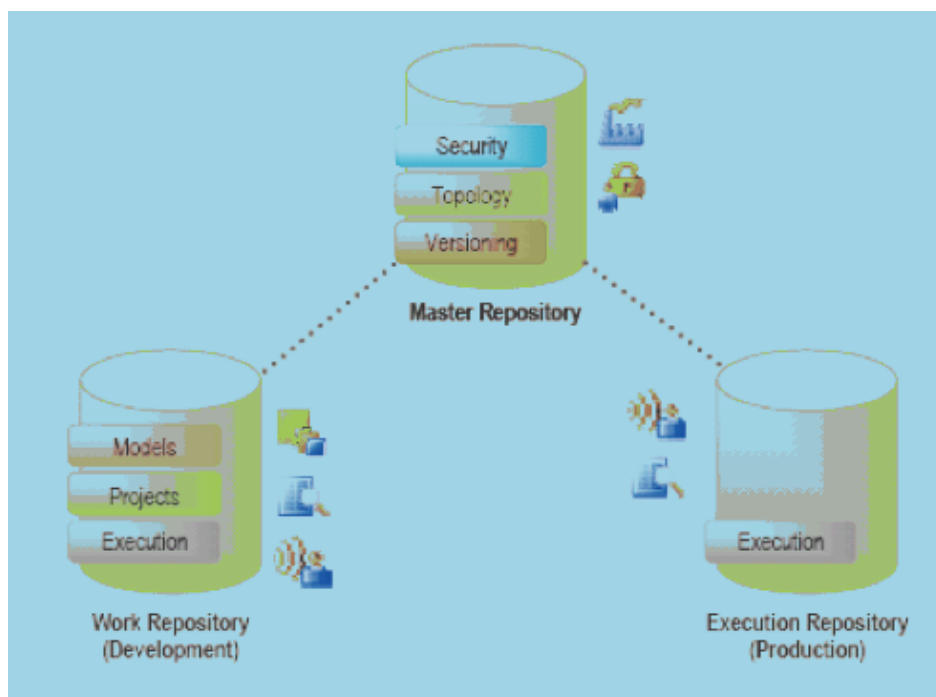


图3：主存储库及工作存储库

项目对象被存储在工作存储库上。不同的工作存储库可以在同一安装中共存。这对于维持隔离环境或反映特定版本生命周期 - - 诸如开发、资质以及生产环境，是非常有用的。

工作存储库中存储关于以下项目的信息：

- 模块 - - 包括数据存储、列、数据完整性约束、交叉索引以及数据沿袭（志）
- 项目 - - 包括声明式规则、软件包、程序、文件夹、知识模块以及变量。
- 运行时信息 - - 包括场景、排程信息以及日志

用户利用Designer以及Operator模块对工作存储库的内容进行管理。还可以通过运行时的Agent（代理程序）对工作存储库进行访问。

在工作存储库仅被用于存储执行信息时（通常出于生产目的），可以将其称为执行存储库。在运行时，利用Operator界面以及通过Agent（代理程序）可以对执行存储库进行访问。

切记所有工作存储库始终附属一个、并且唯一一个主存储库。

Metadata Navigator

Metadata Navigator是一种Java 2 Enterprise Edition（J2EE）应用，可以提供对存储库的Web访问。它使得用户能够对对象进行浏览，包括项目、模块以及执行日志。Metadata Navigator可以被安装在应用服务器上，诸如Oracle Container for Java（OC4J）或Apache Tomcat。

企业用户、开发人员、操作人员以及管理人员可以通过Web浏览器访问Metadata Navigator。凭借其全面的Web界面，用户能够看到流程图、追踪所有数据源、甚至深入到字段级了解构建数据所使用的转换。用户还可以通过Metadata Navigator从Web浏览器中发布和监控场景。

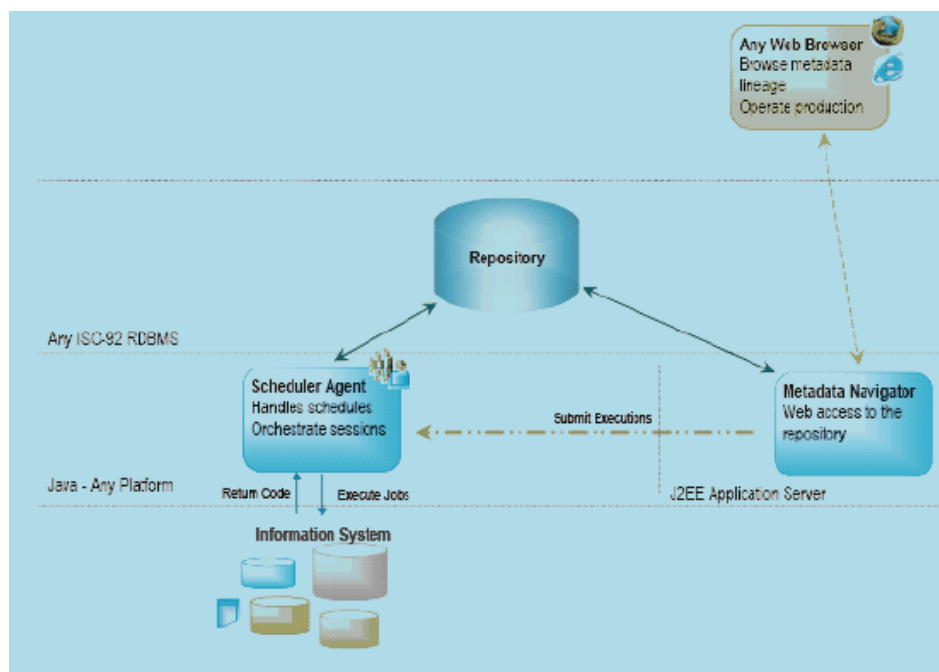


图4：通过Metadata Navigator，用户能够从Web浏览器中访问执行元数据

其他组件及特征

Oracle Data Integrator还包括以下可选择性组件及特征：

- 知识模块可以使技术、数据库以及应用程序快速便捷集成的实现成为可能。它们存在于范围广泛的平台中，包括Oracle、Teradata、Sybase IQ、Netezza、SAP/R3、Oracle Applications、Siebel、LDAP、以及XML。
- 具备负载均衡的Advanced Parallel Option特性，通过自动平衡几种Agent（代理程序）间的工作负载，可以使数据的大批量处理成为可能。
- 高级版本管理可以提供一种管理、保护以及复制工作单元修订的界面，即使在最大程度开发环境下。
- Common Format Designer（CFD）特性使用户能够设计或从其他数据模型中快速组装数据模型，并随后自动生成用于加载和从该模型中提取数据的流程。诸如，用户能够利用Common Format Designer、通过集合异构源生成操作性数据存储、数据中心或数据规范格式。它还可以被用来设计数据仓库模型（诸如，Star或Snowflake Schema，3NF）。

企业用户、开发人员、操作人员以及管理人员可以通过Web浏览器访问Metadata Navigator。凭借其全面的Web界面，用户能够看到流程图、追踪所有数据源、甚至深入到字段级了解构建数据所使用的转换。用户还可以通过Metadata Navigator从Web浏览器中发布和监控场景。

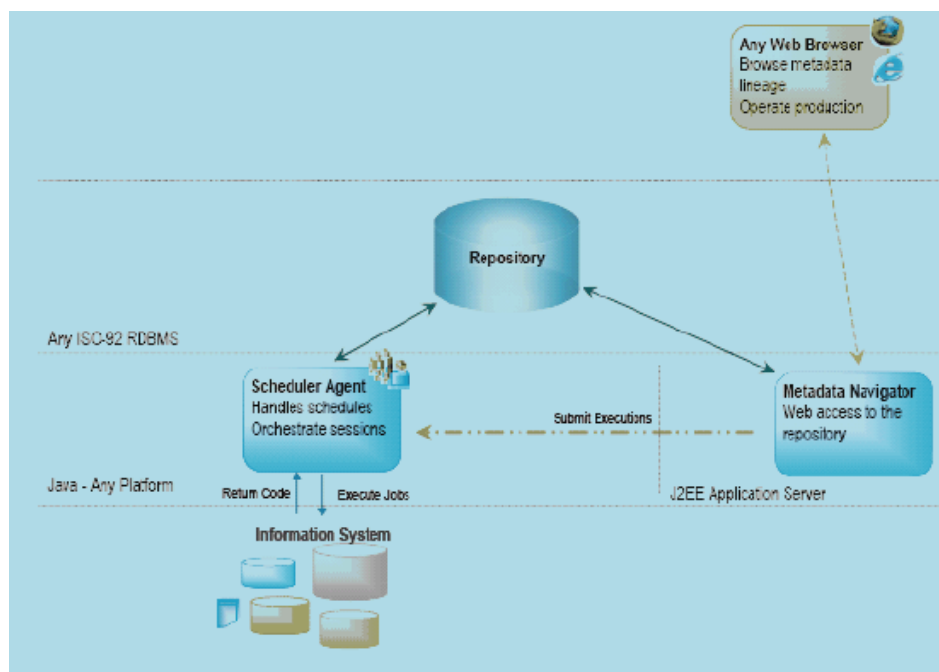


图4：通过Metadata Navigator，用户能够从Web浏览器中访问执行元数据

其他组件及特征

Oracle Data Integrator还包括以下可选择性组件及特征：

- 知识模块可以使技术、数据库以及应用程序快速便捷集成的实现成为可能。它们存在于范围广泛的平台中，包括Oracle、Teradata、Sybase IQ、Netezza、SAP/R3、Oracle Applications、Siebel、LDAP、以及XML。
- 具备负载均衡的Advanced Parallel Option特性，通过自动平衡几种Agent（代理程序）间的工作负载，可以使数据的大批量处理成为可能。
- 高级版本管理可以提供一种管理、保护以及复制工作单元修订的界面，即使在最大程度开发环境下。
- Common Format Designer（CFD）特性使用户能够设计或从其他数据模型中快速组装数据模型，并随后自动生成用于加载和从该模型中提取数据的流程。诸如，用户能够利用Common Format Designer、通过集合异构源生成操作性数据存储、数据中心或数据规范格式。它还可以被用来设计数据仓库模型（诸如，Star或Snowflake Schema，3NF）。

- Publish and Subscribe Changed Data Capture (CDC) (发布与订阅变化数据捕获) 特性通过仅提取变化数据 , 实现对源数据变化的追踪 , 并降低处理数据的容量。
- Publish and Subscribe Messaging (分布与订阅消息) 特性提供利用面向消息中间件 (MOM) 的能力 , 旨在执行一种异步、时间驱动集成体系结构。

总结

Oracle Data Integrator是一种代表当前最高技术水平、非传统性 (Legacy-free) 的轻量级集成平台。所有组件可以在任何Java兼容系统上单独运行。由于其全新的体系结构 , 使Oracle Data Integrator可以在任何平台上得到快速安装。



Oracle Data Integrator结构体系
2006年12月

Oracle公司
世界总部
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A

全球咨询服务：
电话：+1.650.506.7000
传真：+1.650.506.7200
Oracle.com

版权 2006, Oracle. 版权所有

本文件仅供参考之用，对其内容的更改无需另行通知。本文件中所提供内容均“按其现状”提供，无任何口头或书面的担保，包括是否适合销售、是否适用于某一特定用途方面的担保。我们特别声明概不承担任何与本文件相关的责任，以及由本文件直接或间接构成的合同义务。事先未经我方的书面允许，不得借助任何方式、以任何形式，无论是电子或机械，再造或转发本文件。

Oracle、JD Edwards、PeopleSoft以及Siebel为Oracle Corporation以及/或其附属机构的注册商标。其他名称可能是各自所有者的商标。