



**ORACLE®**

**Oracle Data Integrator 介绍**

# 议程

- 为什么需要数据集成
- 什么是**Oracle Data Integrator**
- 架构介绍
- 扩展功能
- **ODI**应用场景
- 总结



# 为什么需要数据集成？

需要...

## Information How and Where you Want It



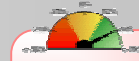
Business Intelligence



Corporate Performance Management



Business Process Management



Business Activity Monitoring

## Data Integration

Migration



Data Warehousing



Master Data Management



Data Synchronization



Federation

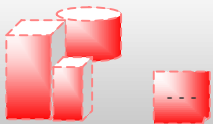


Real Time Messaging



有...

## Data in Disparate Sources



Legacy



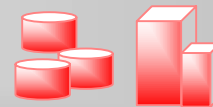
ERP



CRM

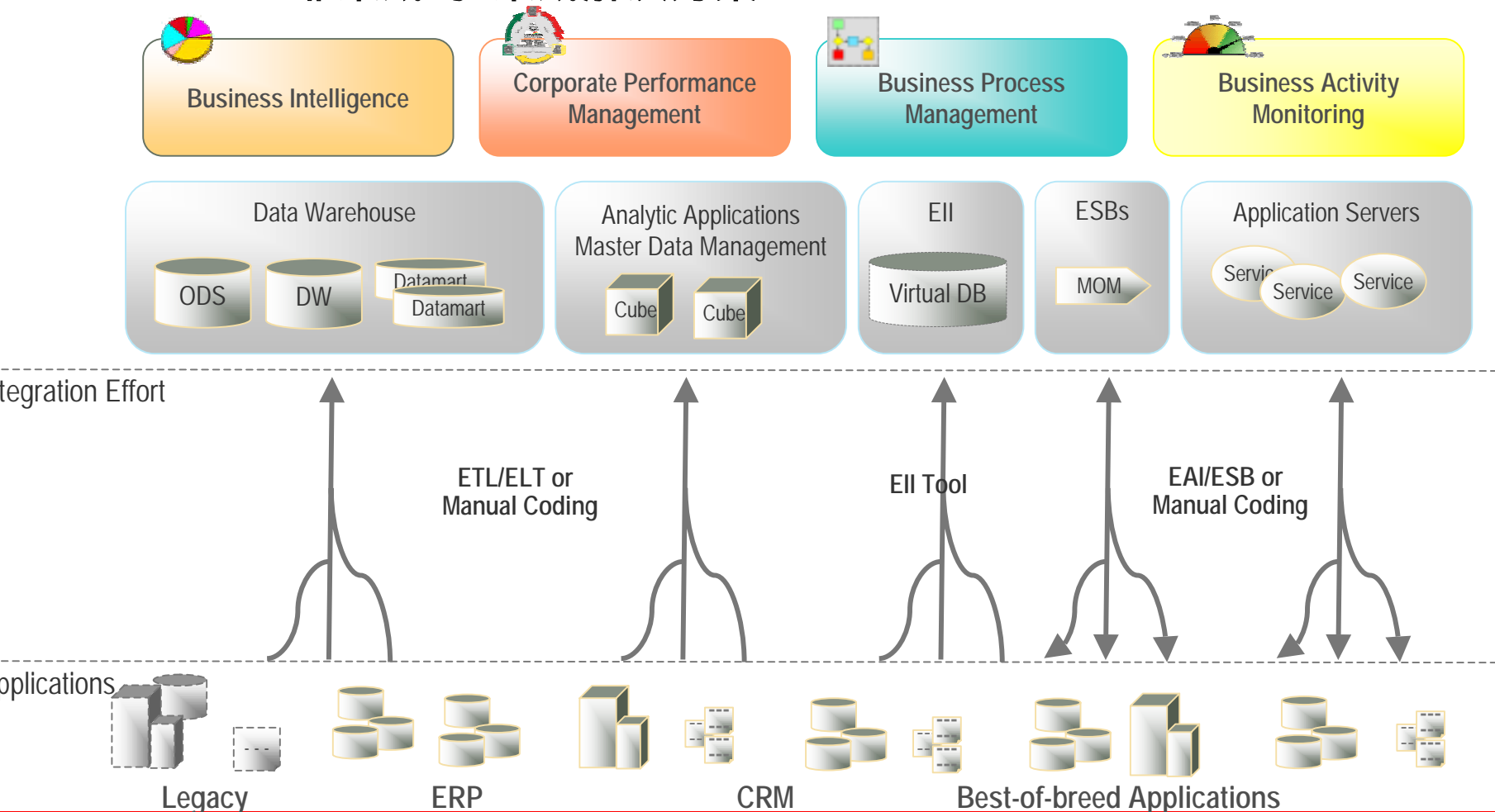


Best-of-breed Applications



# 传统的集成架构

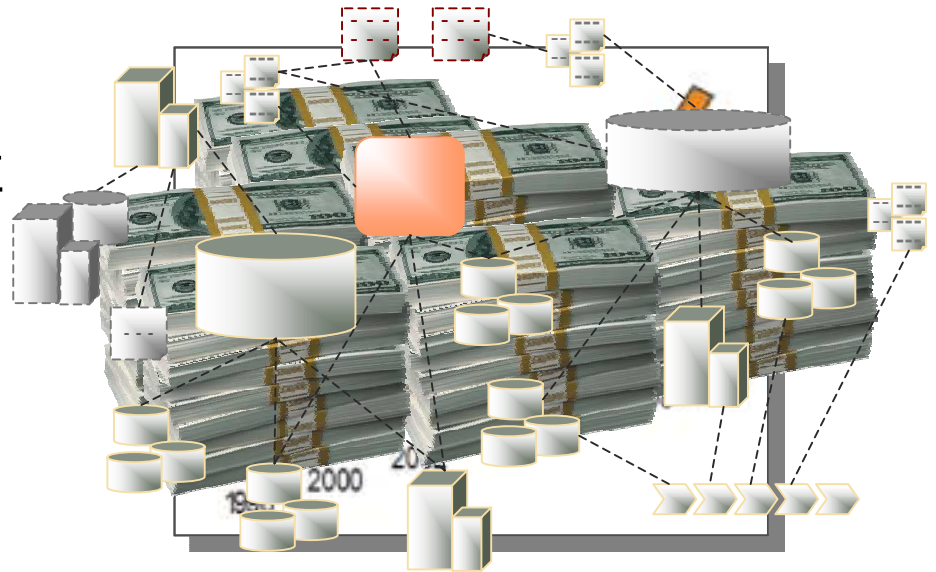
- 多种集成工具和方案
- 非集成的“集成解决方案”



# 集成的挑战

## 挑战

- 复杂
- 快速增长的集成数据量
- 资源的限制



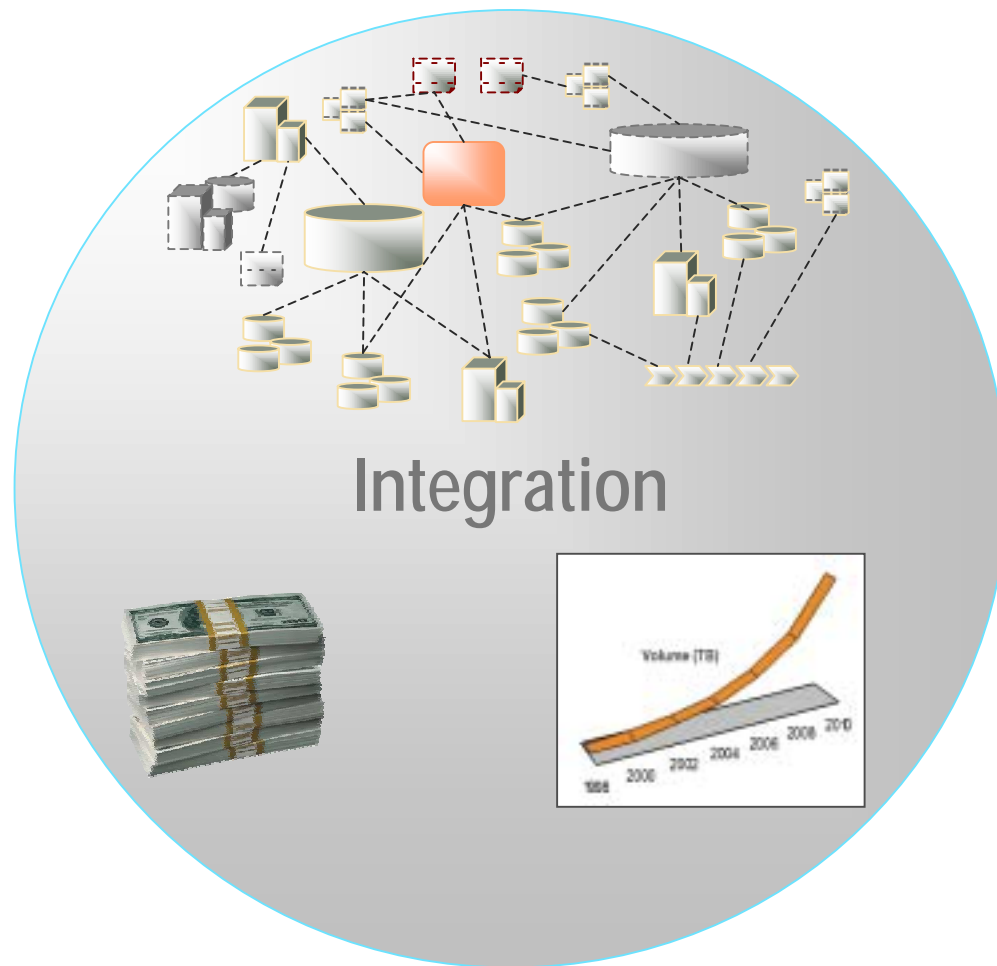
# 理想的数据集成平台

- 高性价比

- 高效率

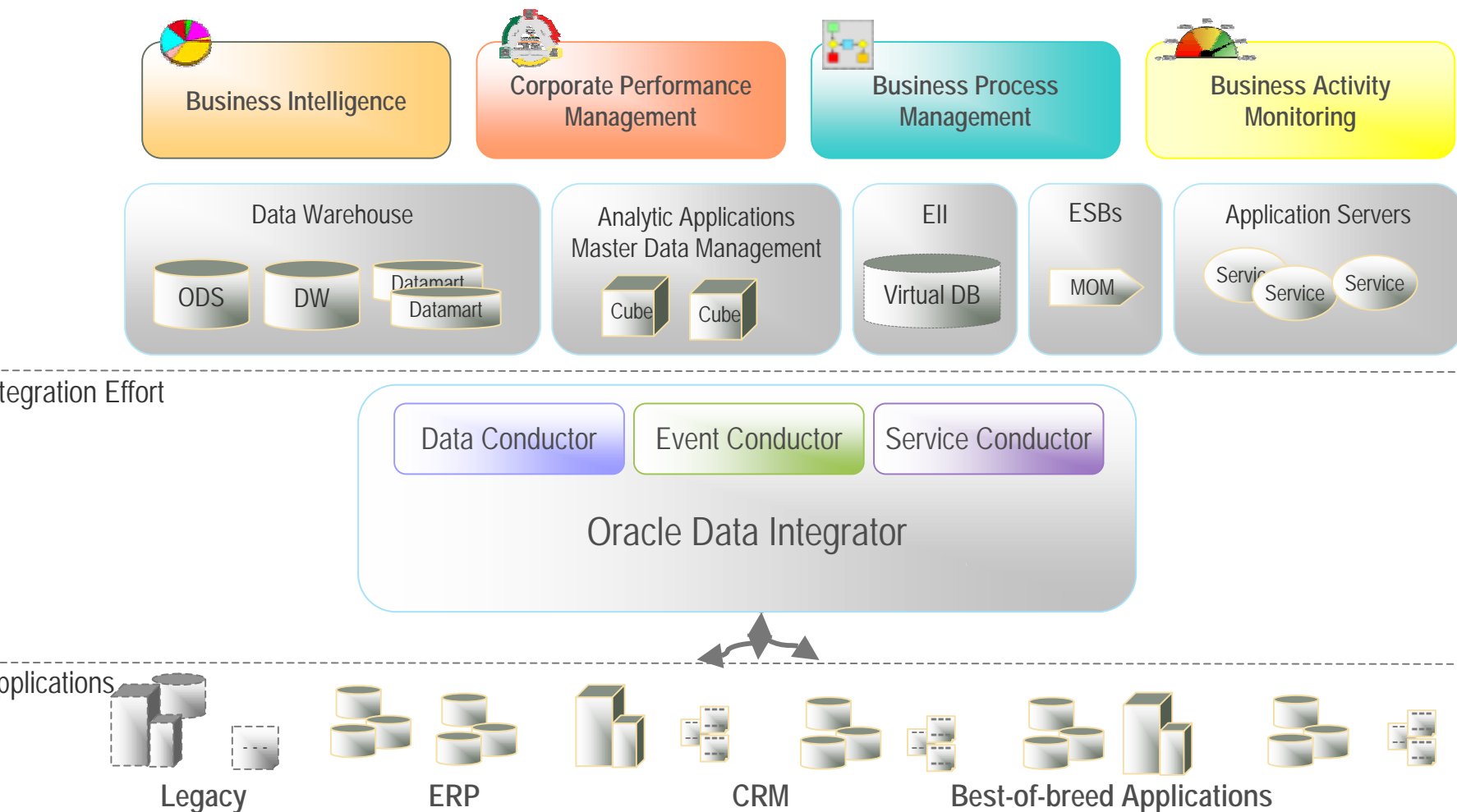
- 单一集成平台

- 数据仓库
- 数据迁移
- 企业数据集成



# Oracle Data Integrator

- 单一集成平台



# 什么是Oracle Data Integrator

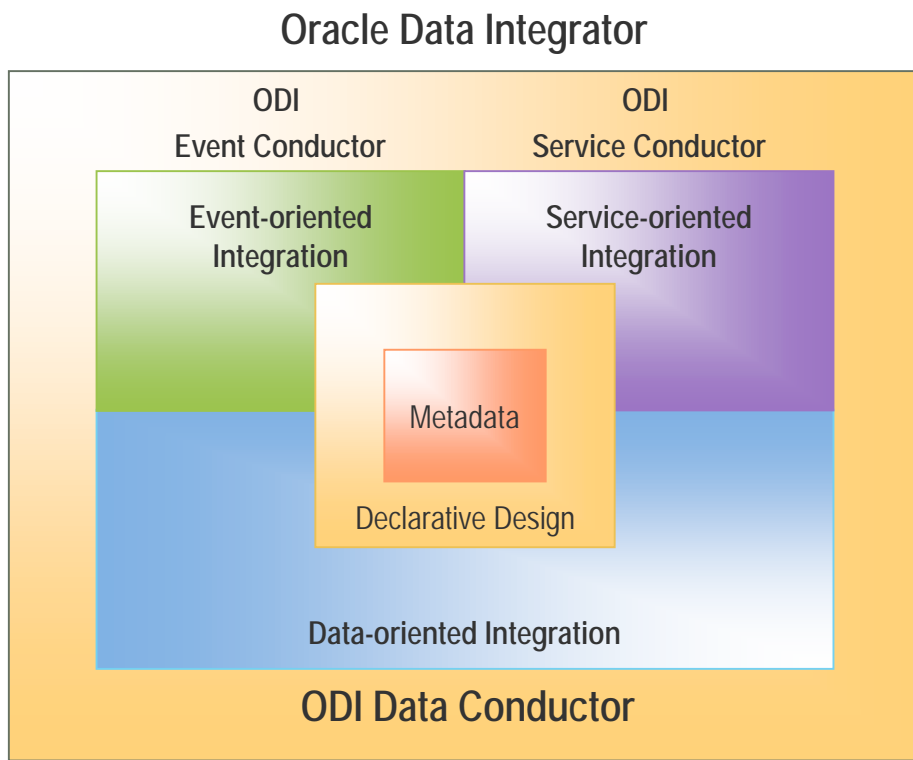


# 什么是Oracle Data Integrator ?

- **Oracle 从Sunopsis收购而来**
  - 完善Oracle的数据集成方案
  - 加强Oracle的BI, SOA和 MDM 产品
- **关于Sunopsis**
  - 被证明的异种数据集成技术的领先者
  - 支持异种数据库、文件系统和ERP应用
  - 关键的不同点: E-LT, Knowledge Modules, Declarative Tools
- **Oracle的策略**
  - 和Oracle中间件和数据库集成
  - 完善Oracle的产品线 : BI, SOA, & MDM
  - 扩展已有的数据集成 (OWB) 到异种目标
  - 作为Oracle应用软件的基础平台

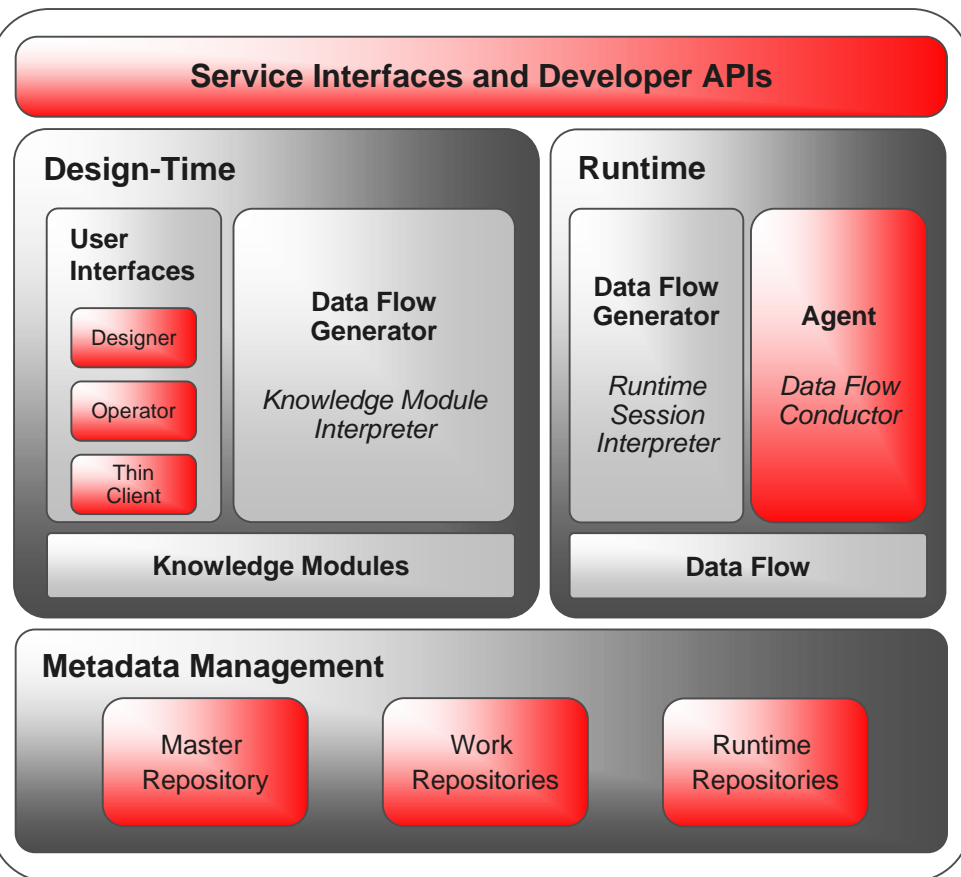
# 什么是Oracle Data Integrator ?

**Oracle Data Integrator 是一个完整的数据集成平台，能够满足所有数据集成的需求。**



# Oracle Data Integrator 架构介绍

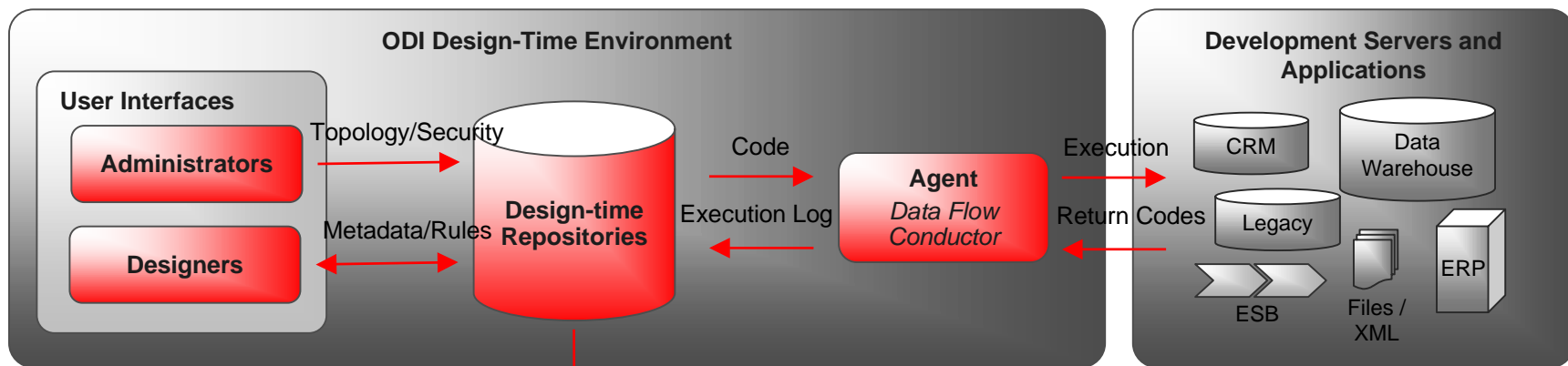
# 架构: 概念视图



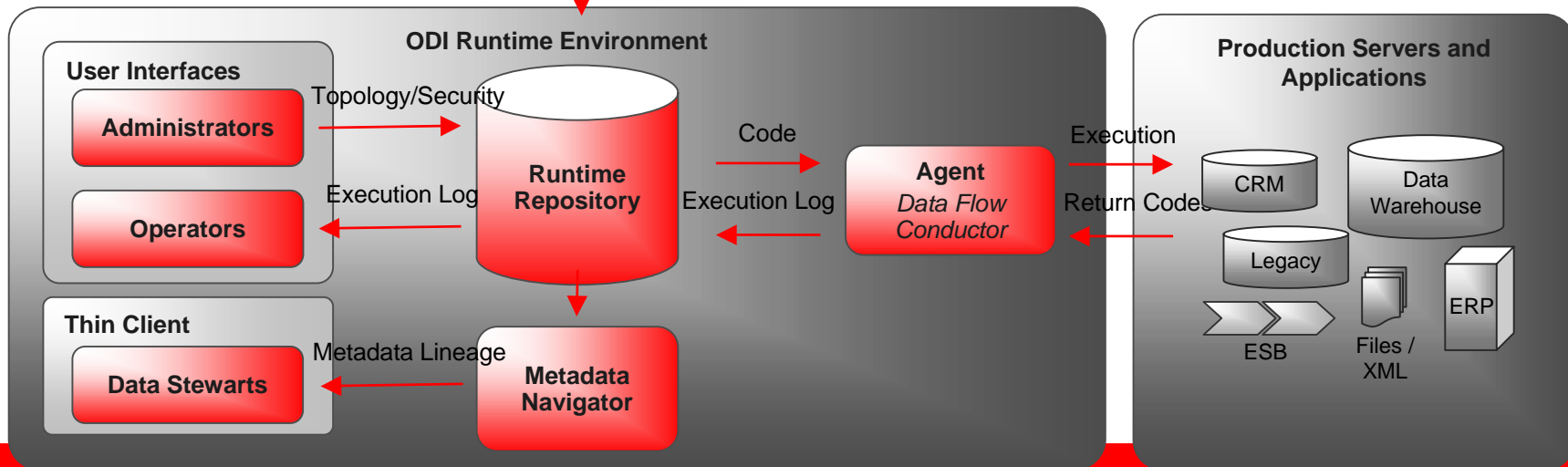
- **纯Java的设计环境**
  - 可以在任何平台上运行
  - 通过瘦客户端浏览元数据
- **纯Java的运行环境**
  - 可以在任何平台上运行
  - 协调执行数据流
- **元数据库**
  - 可以装在任何关系数据库上
  - 便于开发
  - 组件化的、可扩展的元数据

# 架构: 概念视图

## Development

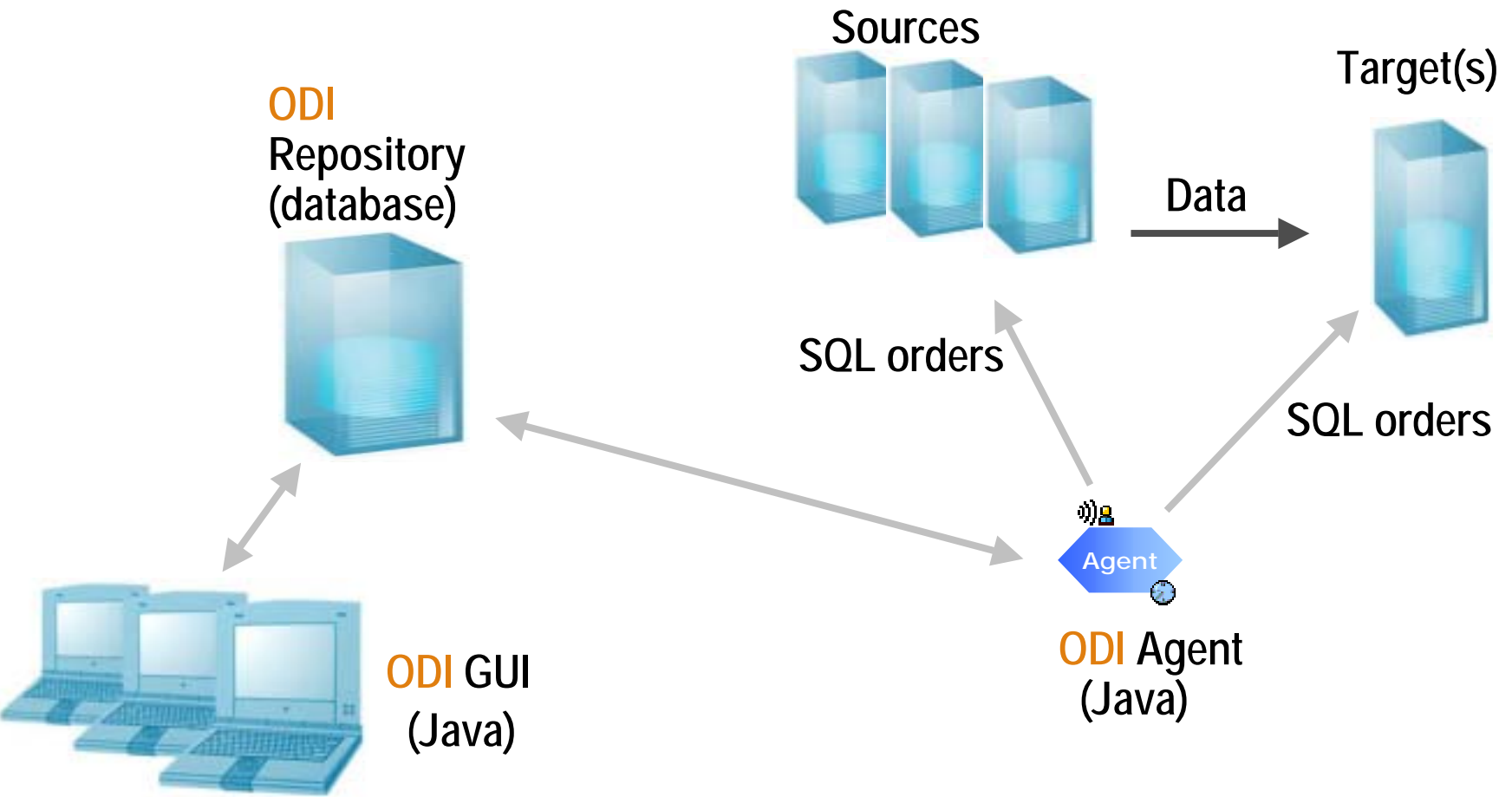


## Production



# Oracle Data Integrator

## 架构示例



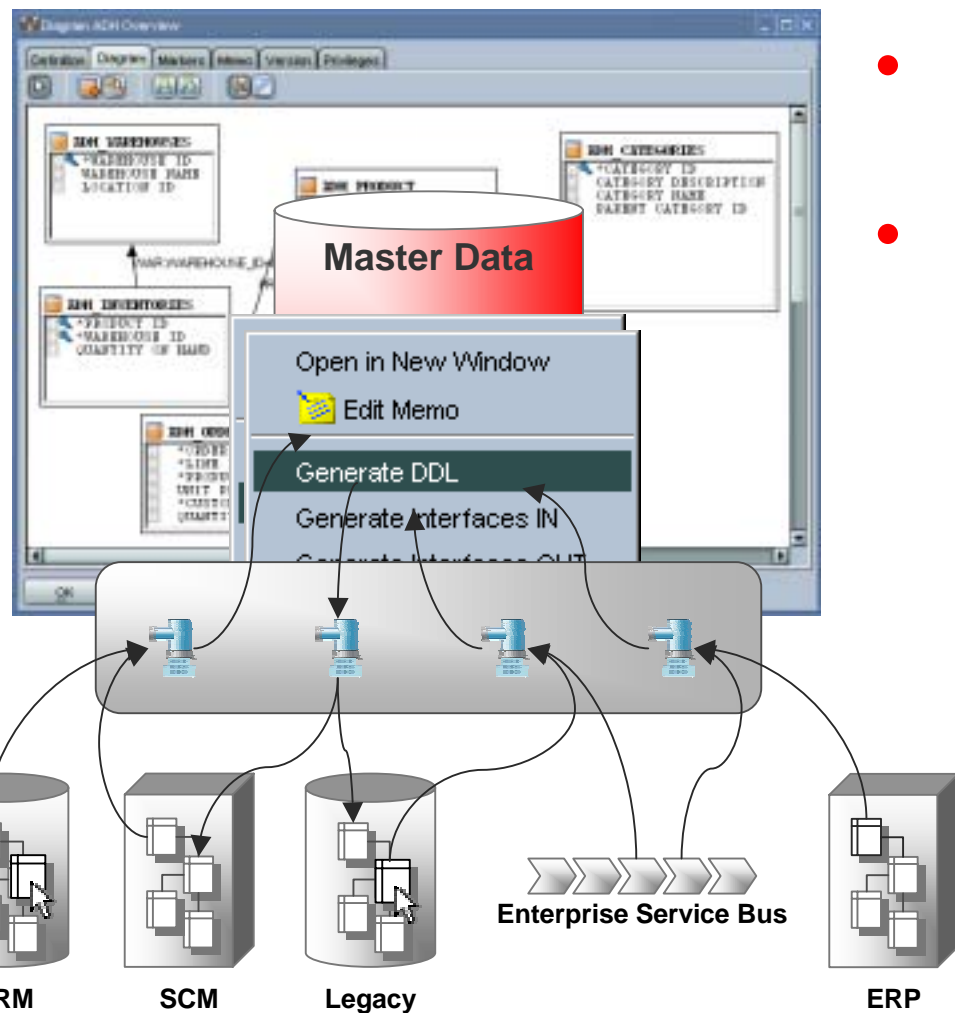
# Oracle Data Integrator 扩展功能

# 扩展的功能

- **主数据管理**
  - 通用格式设计
  - 自动生成权威格式和转换
  - 内建的数据集成
- **实时数据**
  - 变化数据捕获
  - 基于消息的集成
- **SOA功能**
  - 生成数据服务
  - 生成转换服务
- **扩展能力**
  - 知识模块的框架设计
  - 脚本语言
  - 开放的工具

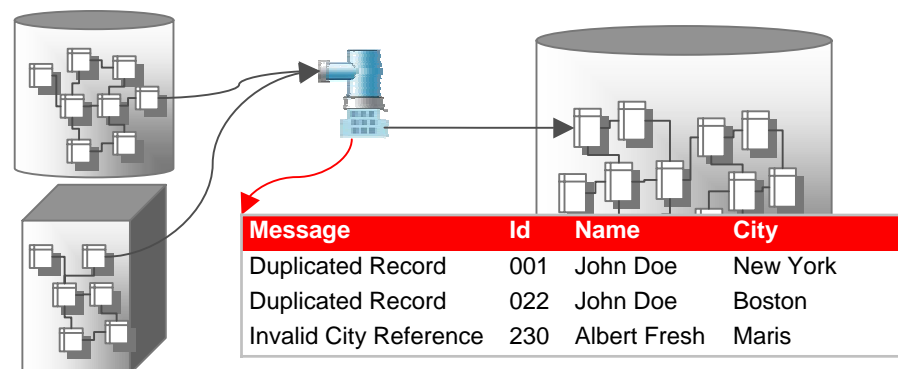


# 主数据管理：通用格式设计



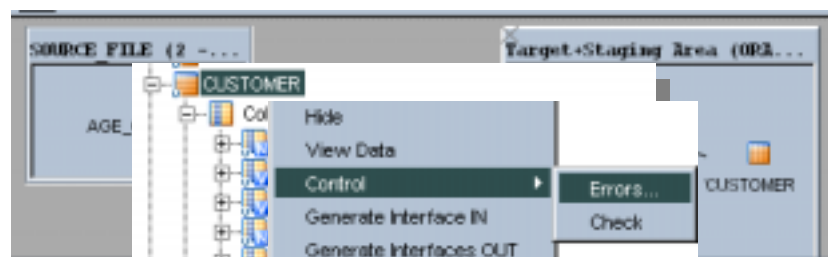
- 与打包的主数据管理应用集成
- 设计和推广典型的格式
  1. 使用现存的元数据来设计主数据管理的应用（实体、领域、关系）
  2. 生成和管理主数据结构
  3. 使用元数据生成和部署转换

# 主数据管理：内建数据集成

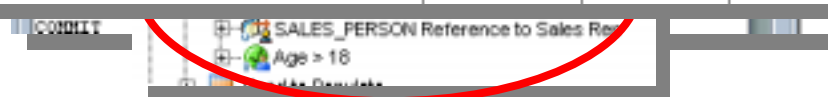


- 数据完整性防火墙
- 审计，清洗和回收

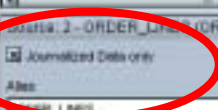
1. 定义表级约束
2. 设计映射和检查流程完整性
3. 审计、清洗或者回收被拒绝的记录



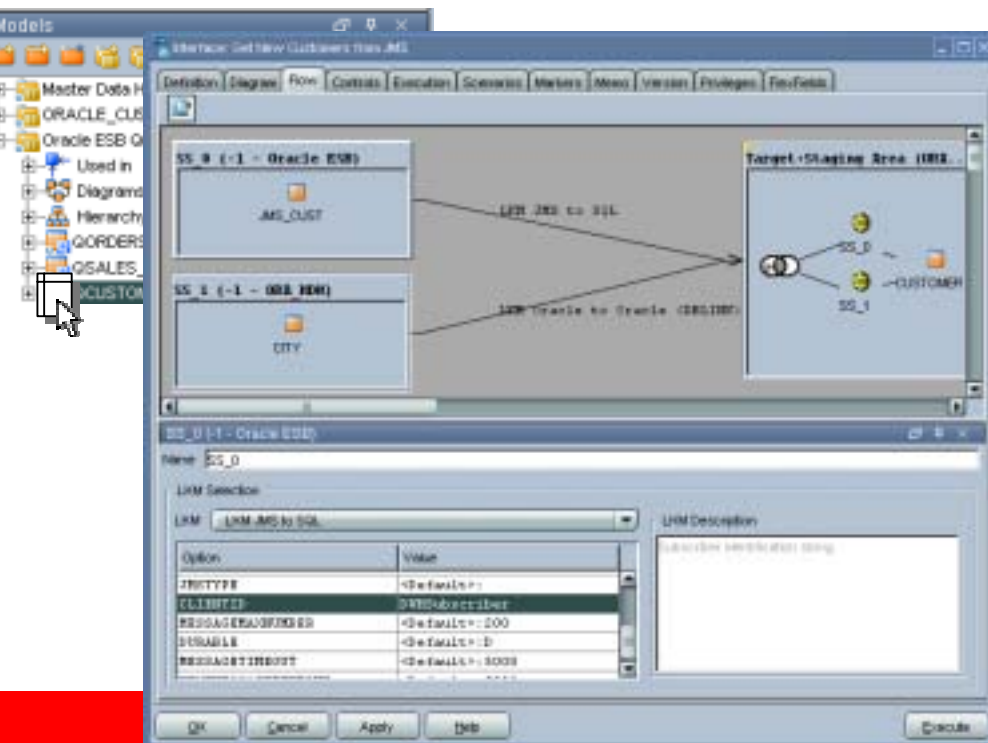
Error Message	Customer Id	First Name	Last Name	Age	City	Sales Rep	Phone	Address	Date Checked
Age should be greater than 18	107	Jacky	Swenson	17	19		(202) 555	64 Imagination Drive	12/23/05
Duplicated Values: "Unique Name and Sales Rep"	508	Philippe	Michaud	19	23	22	123 456 1	Oakland street 15	12/23/05
Duplicated Values: "Unique Name and Sales Rep"	202	Philippe	Michaud	22	23	22	78 21 86 2	197 impasse Renc	12/23/05
Missing Reference (Reference to City) To DEMO	103	Peter	Travis	33	342		(510) 555	7835 Hartford Drive	12/23/05
Missing Reference (Reference to Sales Rep) To	105	Tony	Gold	55	19	231	(619) 555	91 Torre drive	12/23/05



100

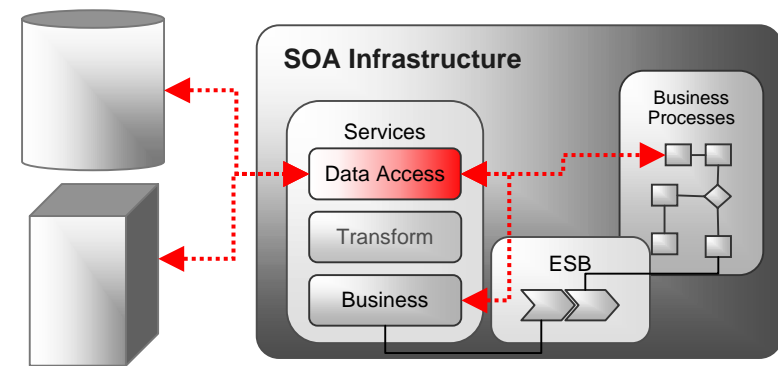


1. 设计和生成映射
2. 选择增量数据
3. 开始捕获



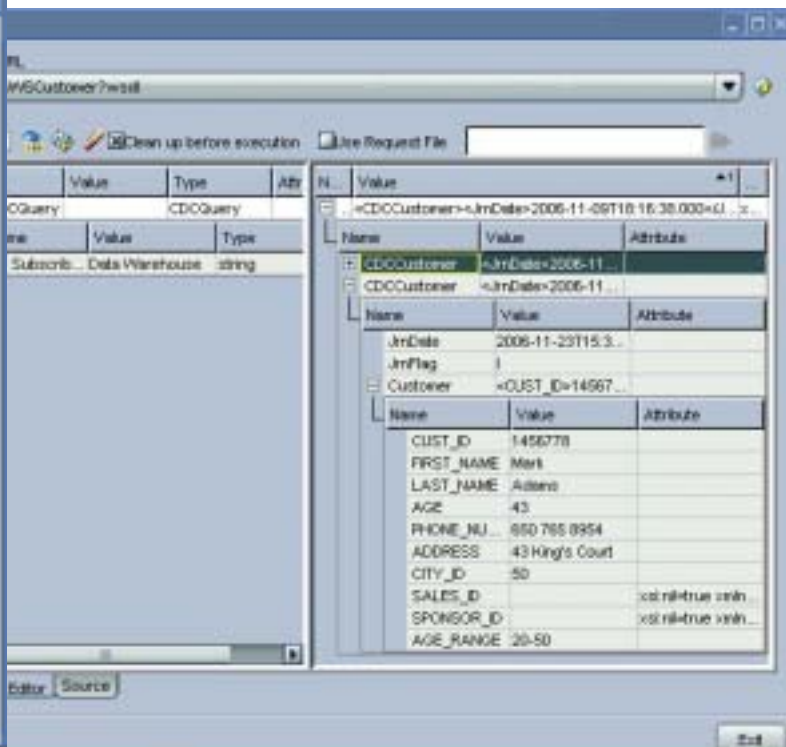
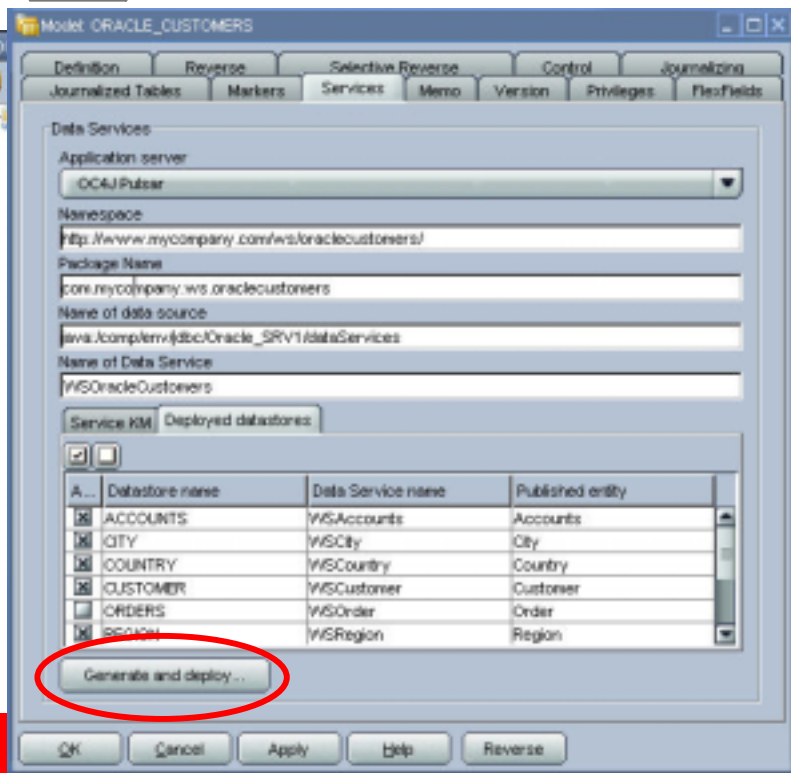
- 与发布和订阅JMS消息的软件集成
- 保证消息传递的事务完整性
- 大数据量的批量转换
  1. 设计复杂的批量转换，混合了队列、数据库和应用
  2. 使用JMS队列和主题作为源或者目标

# SOA功能：数据访问服务

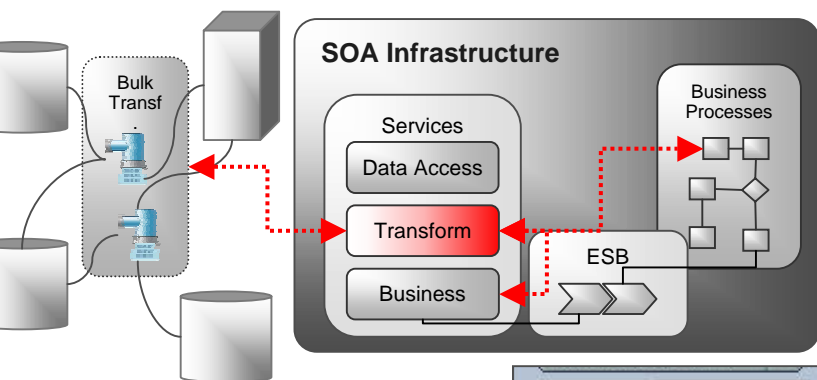


## 生成和共享数据访问服务

1. 生成和部署数据服务
2. 测试数据服务
3. 让数据服务在SOA架构中发挥作用



# SOA功能：数据流程服务



- 使得数据转换成为一个Web Service

1. 编排数据流程
2. 以Web Service形式发布数据流程到你的SOA架构中

The screenshot shows the Oracle SOA Suite configuration interface. The left pane displays the 'Package: UPDATE\_SALES' with a 'Toolbox' containing various components like 'SnpsExportScen', 'SnpsFileAppend', etc. The right pane shows the configuration for the 'getWebServiceVersion' and 'invokeScenario' services. The 'invokeScenario' section is expanded, showing a table of parameters.

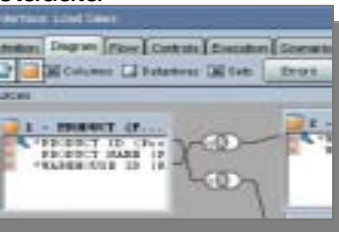
Name	Value	Type	Attribute
RepositoryConnection			
Repository...			
Name			
OdiUser	SUPERVISOR	string	
OdiPass...	*****	string	
Command			
Scenario...			
Name			
ScenName	UPDATE_SALES	string	
ScenVer...	001	string	
Context	PRODUCTION	string	
Session...	Update DW Sales	string	
Keywords	Prd, PTY1, Sales	string	
Variable...			
Variabl...			
Name			
Name	REFRESH_DATE	string	
Value	2007-10-18	string	
Agent			
AgentType			
Name			
Host	coruscant	string	
Port	20040	int	

# 扩展性：知识模块

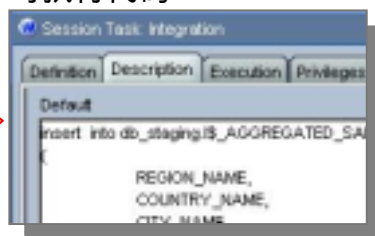
M的元代码

```
getInfo() {  
    return snpRef.getTable  
}  
  
getCollist() {  
    return snpRef.getCollist  
}  
  
getPop("DIST") {  
    return snpRef.getPop("DIST")  
}  
  
getCollist("HAS_JPM") {  
    return snpRef.getCollist("HAS_JPM")  
}  
}
```

Metadata



可执行代码

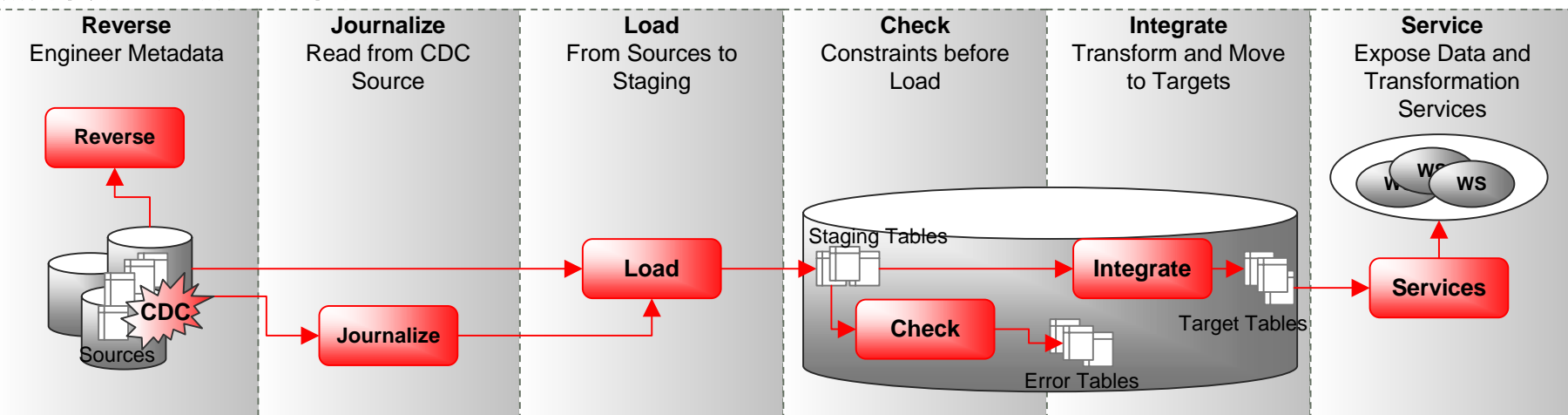


- 120多个知识模块（非黑盒的）

- ✓ 开发和利用最佳实践
- ✓ 简化管理工作
- ✓ 减少拥有成本

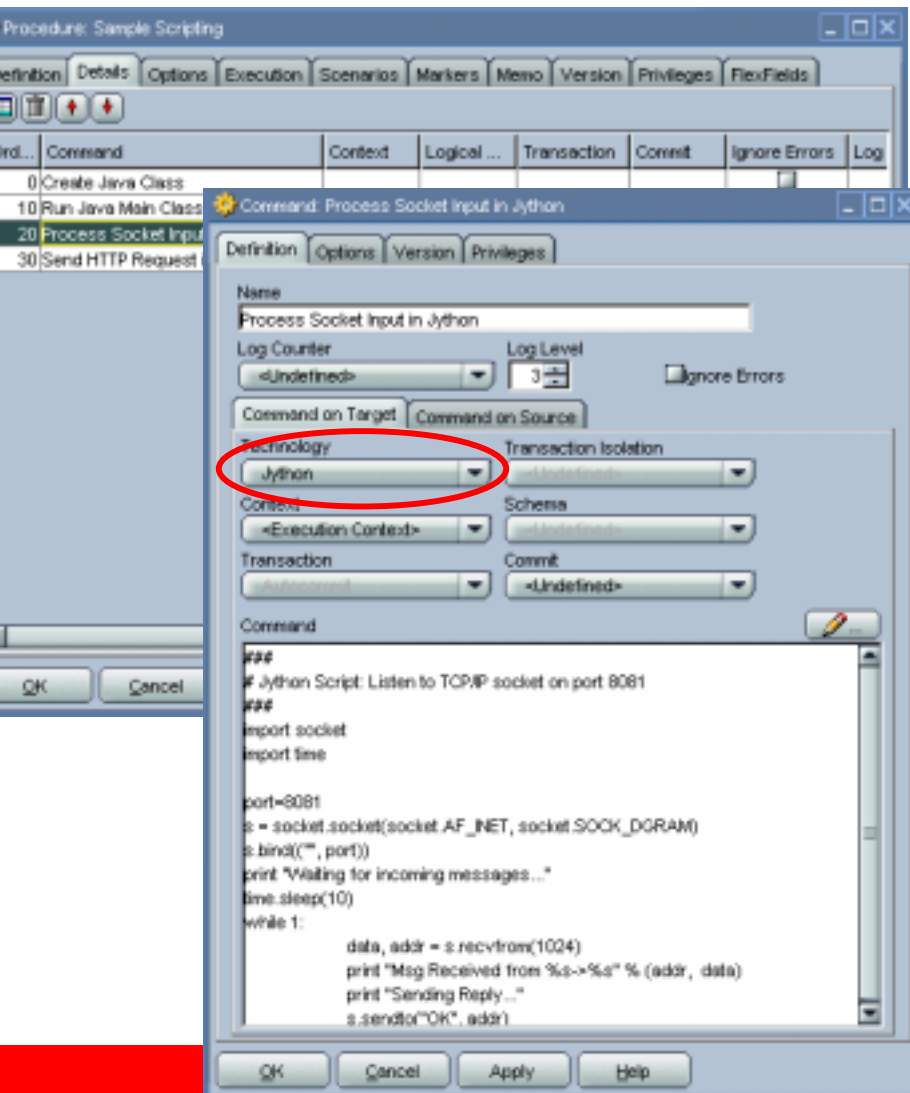
- 客制化和扩展性

## 热插拔的知识模块架构





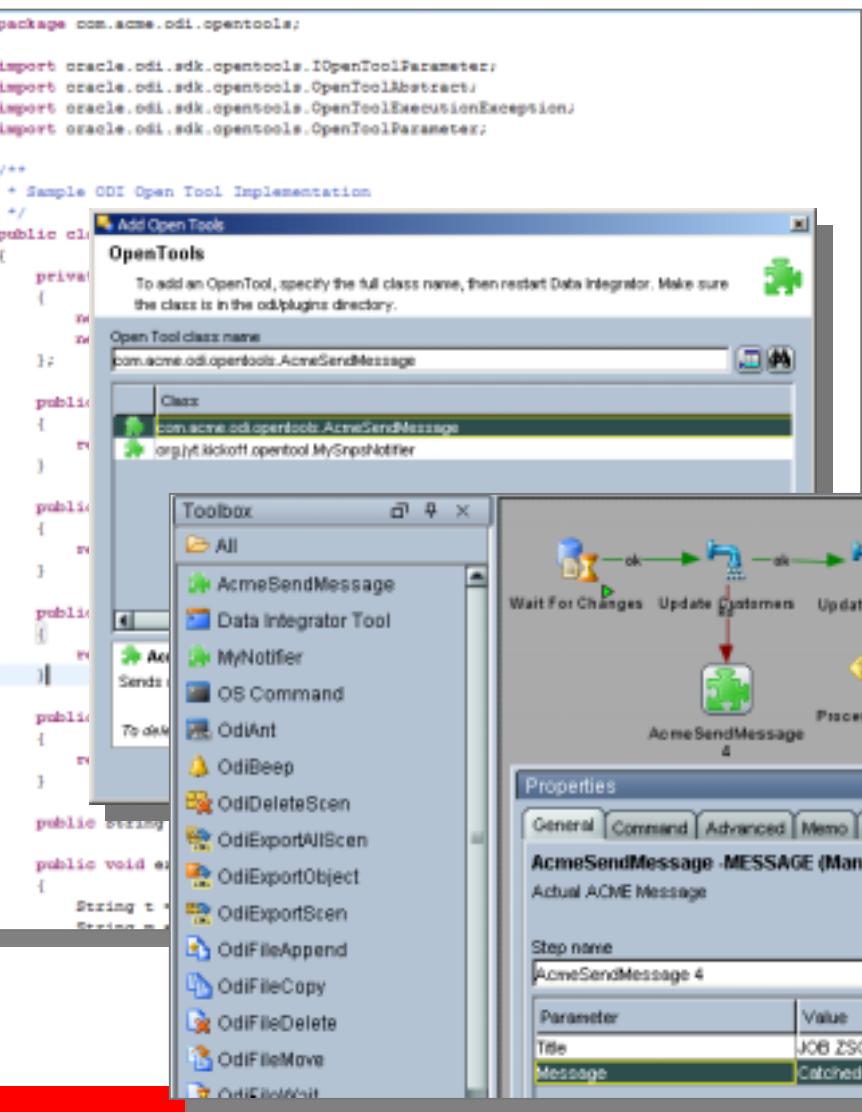
# 扩展性：脚本的架构



- 通过脚本过程扩展数据流程
- 利用到所有的数据库语言
  - SQL, PL/SQL, Transact SQL, etc.
- 使用操作系统脚本
  - Win32 DOS, sh, ksh, csh, OS400 commands, JCL, etc.
- 可选兼容的Bean Scripting Framework语言
  - Java, JavaScript, Jython (Java Python), Perl, etc.



# 扩展性：开放的工具

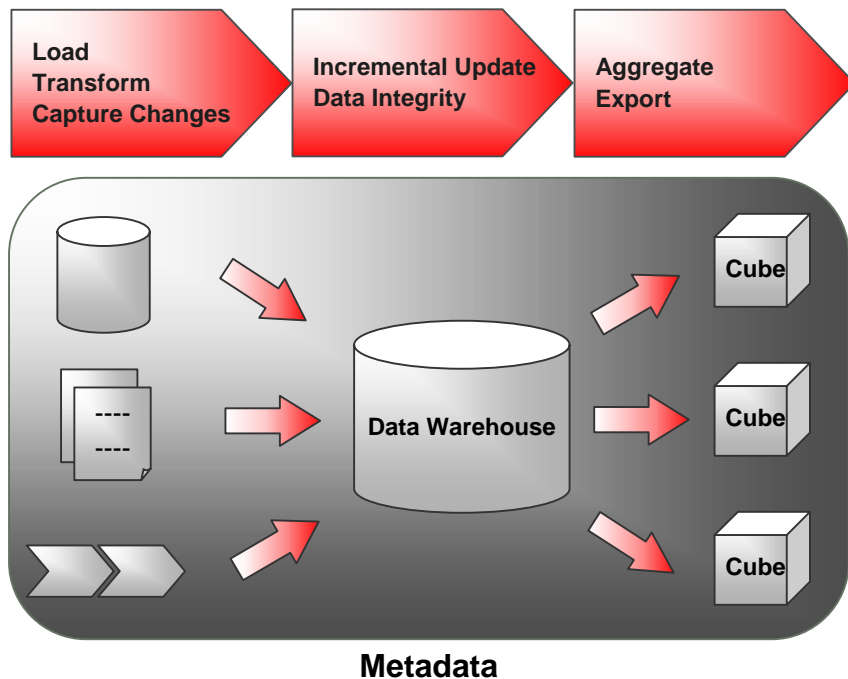


- 扩展ODI工具
- 可以增加你自己的工具到设计面板中
  1. 继承  
OdiOpenToolAbstract  
Java接口
  2. 在ODI中注册开放的工具
  3. 在你的设计环境下使用开放工具

# ODI应用场景

# 数据仓库的E-LT功能

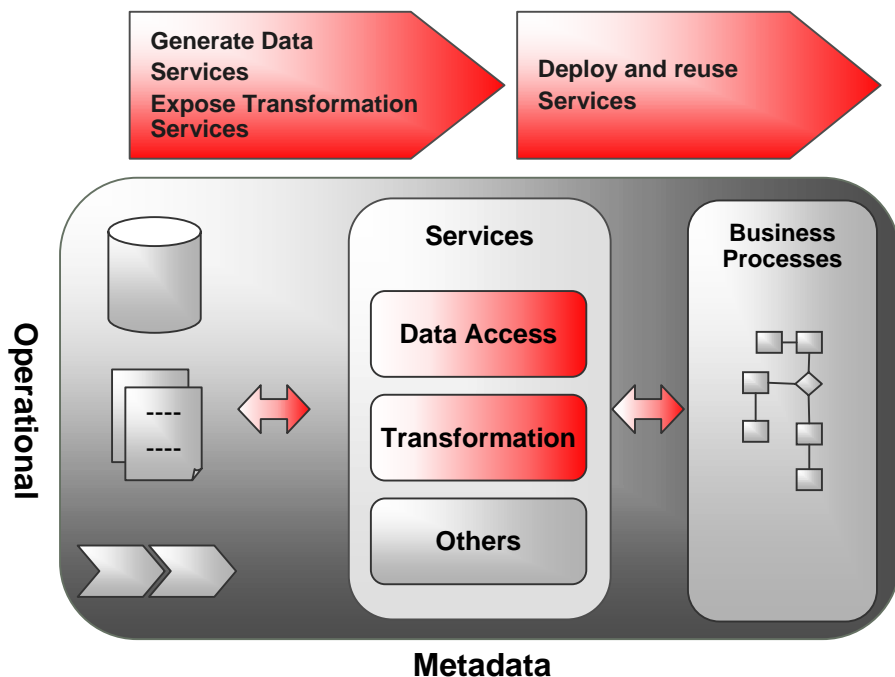
## 用ODI来推广高性能的数据仓库



- ✓ 异构的源和目标
- ✓ 增量加载
- ✓ 缓慢变化维
- ✓ 数据完整性和一致性
- ✓ 变化数据捕获
- ✓ 数据关联分析

# SOA交互

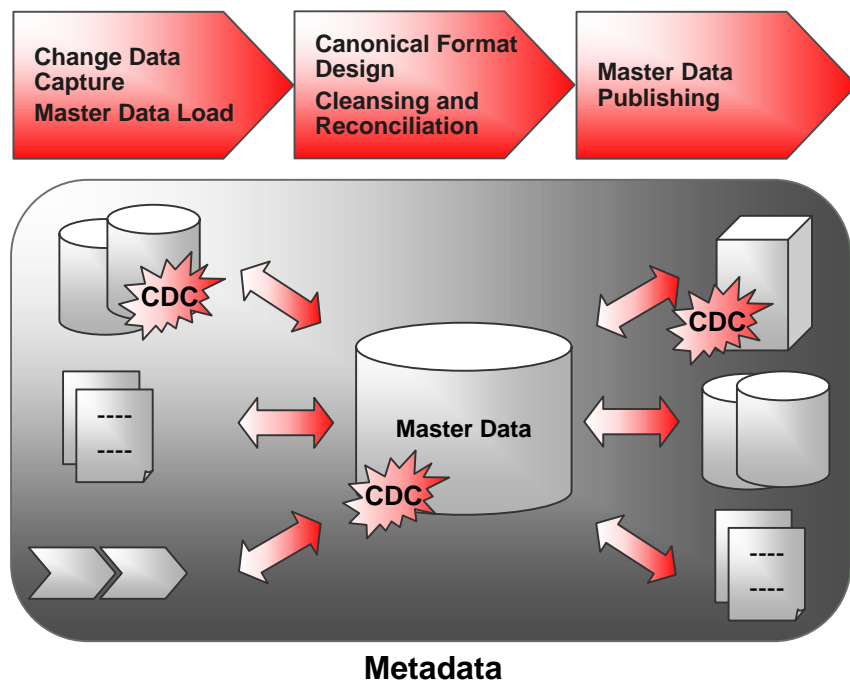
为数据集成建立消息架构，在ODI中合并了数据批处理功能



- ✓ 调用外部服务做数据集成
- ✓ 部署数据访问服务
- ✓ 部署数据转换服务
- ✓ 集成数据访问和转换服务到你的SOA架构中

# 主数据管理

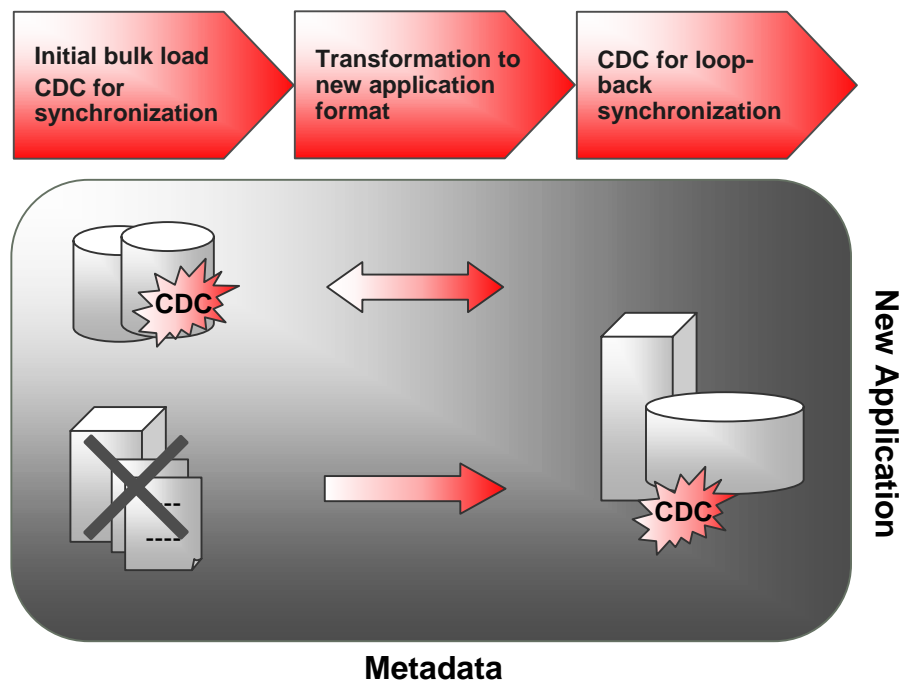
创建事实的单一视图，通过ODI同步数据



- ✓ 与打包的主数据管理应用集成
- ✓ 用作基础架构，设计你自己的HUB
- ✓ 创建声明式的数据流程
- ✓ 变化数据捕获 (CDC)
- ✓ 规范并清洗数据
- ✓ 发布和共享主数据
- ✓ 扩展元数据定义

# 数据迁移

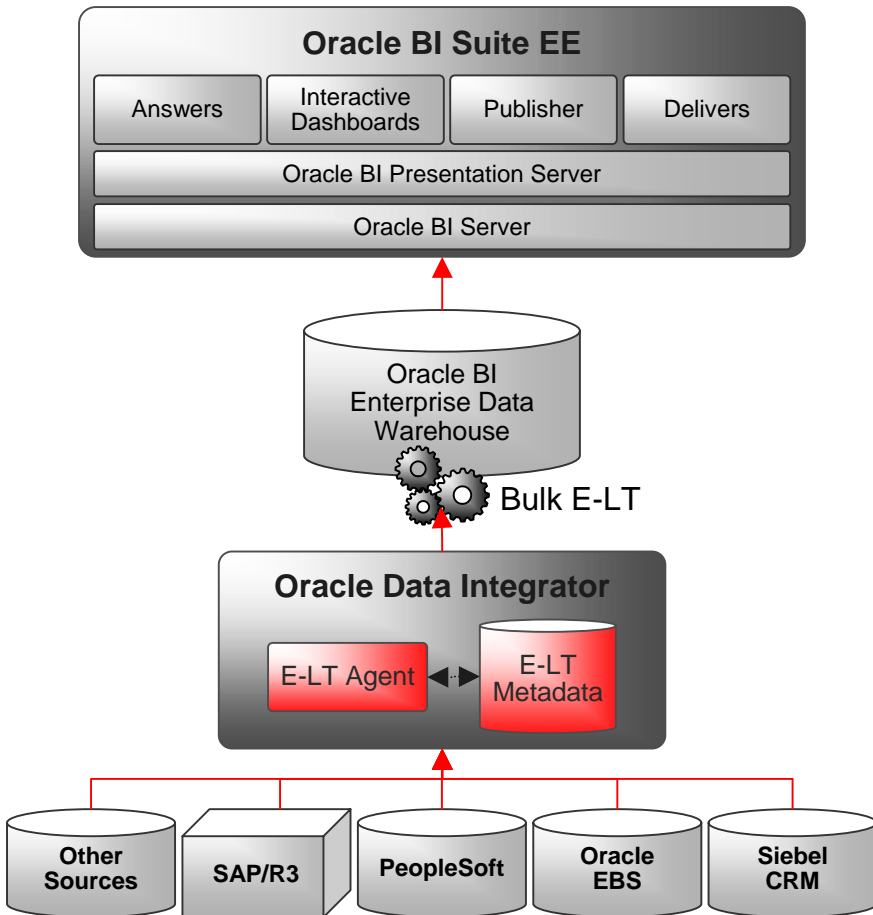
更新应用或者移植到一个新的方案中，用ODI一次性迁移数据并保持同步



- ✓ 批量加载历史数据到新的应用中
- ✓ 转换源格式到目标
- ✓ 两个系统并存时，同步新旧应用间的数据
- ✓ 双向捕获变化数据 (CDC)

# ODI增强了Oracle BI

## 结合ODI来推广高性能数据仓库



## Oracle Business Intelligence Suite EE:

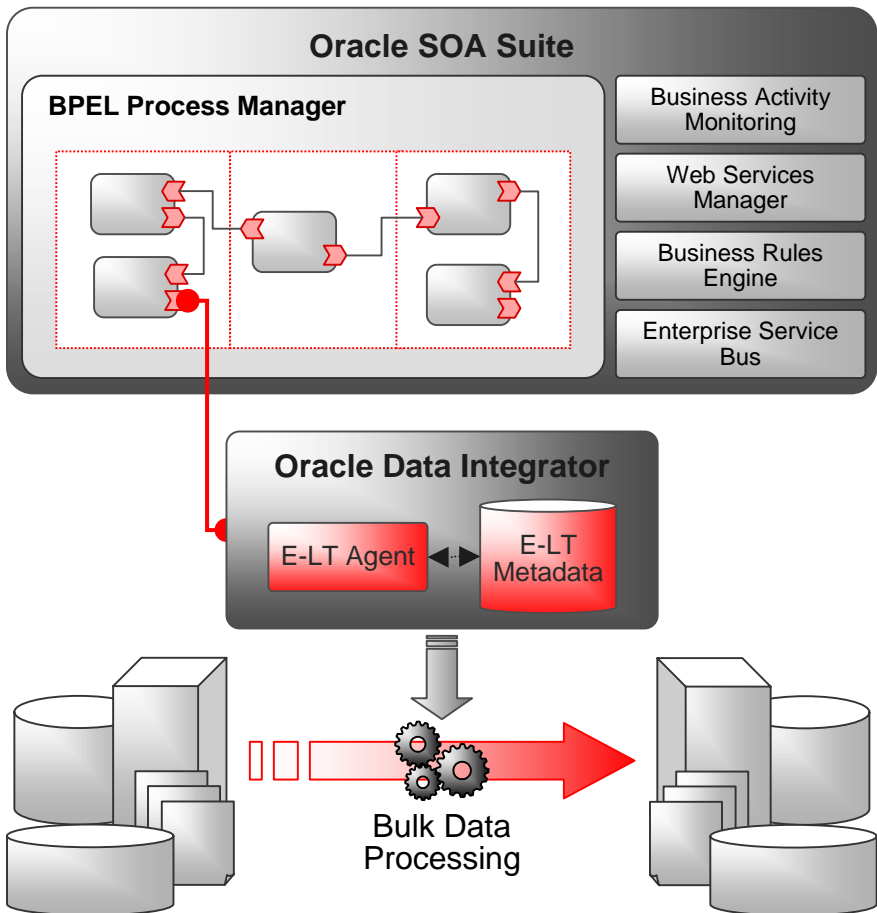
- ✓ 简单的业务模型视图
- ✓ 高级计算和计算引擎
- ✓ 智能需求生成
- ✓ 优化的数据访问

## Oracle Data Integrator:

- ✓ 推广数据仓库
- ✓ 优化的加载和转换性能
- ✓ 可扩展的预打包的E-LT上下文

# ODI增强了Oracle SOA套件

## 给BPEL增加了批量处理转换数据的能力



### Oracle SOA套件:

- ✓ BPEL过程管理能力来管理商业流程

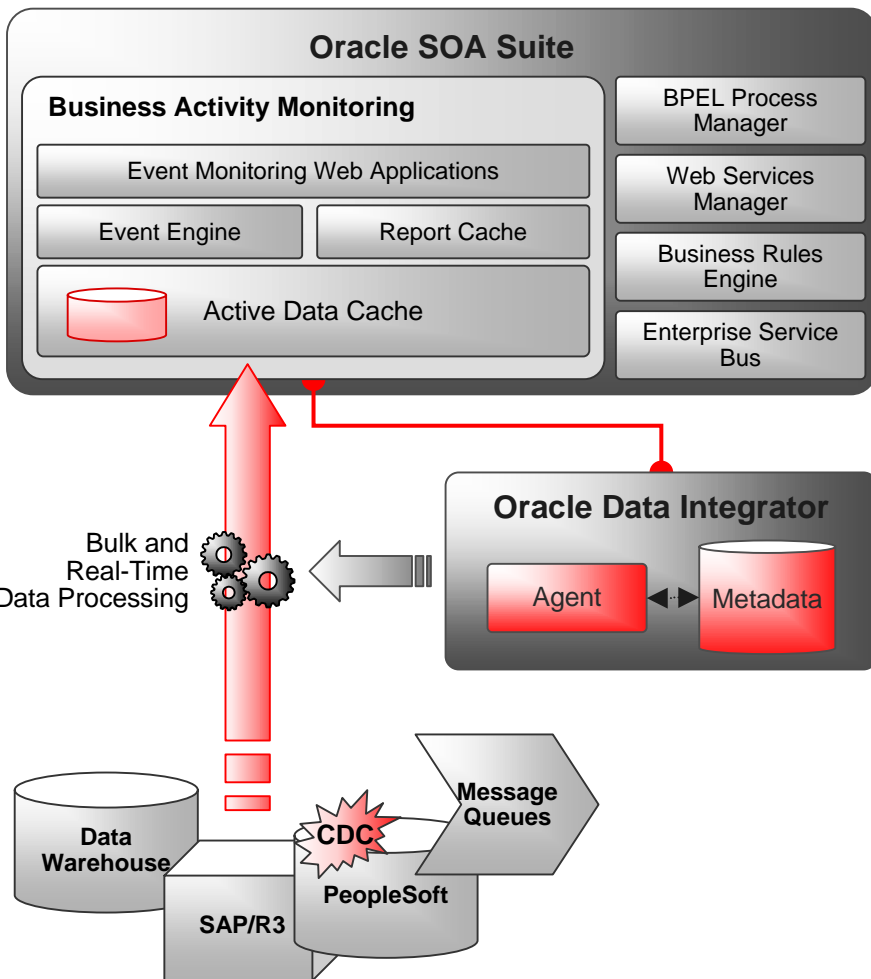
### Oracle 数据集成器:

- ✓ 有效的批量数据作为商业流程的一个步骤
- ✓ 通过数据服务和转换服务交互



# ODI增强Oracle SOA套件

## 使得BAM数据缓存更有效



## Oracle SOA 套件:

- ✓ 业务监控，为实时的洞察力

## Oracle Data Integrator:

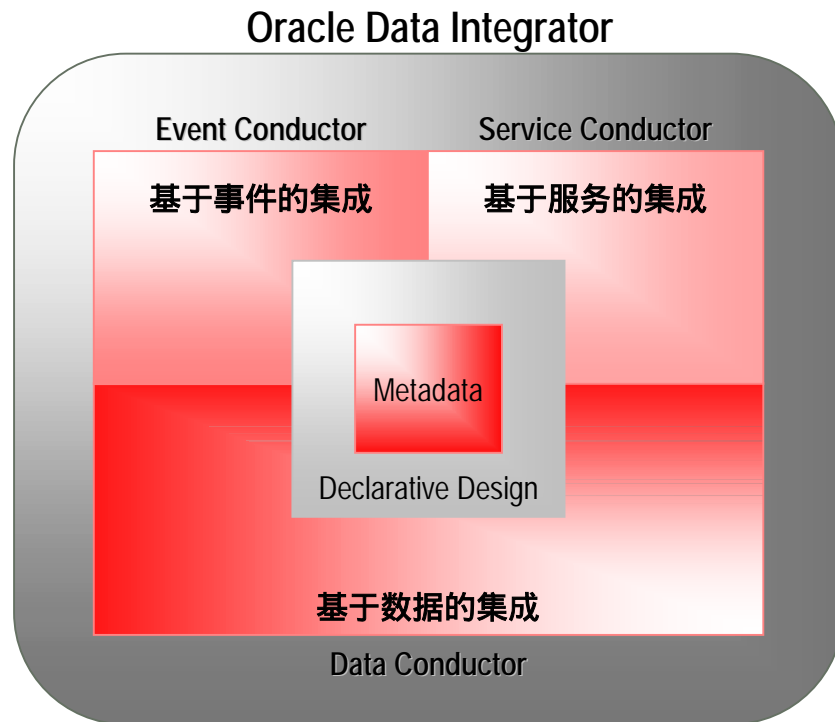
- ✓ 高性能的BAM缓存数据加载
- ✓ 预建的和集成的



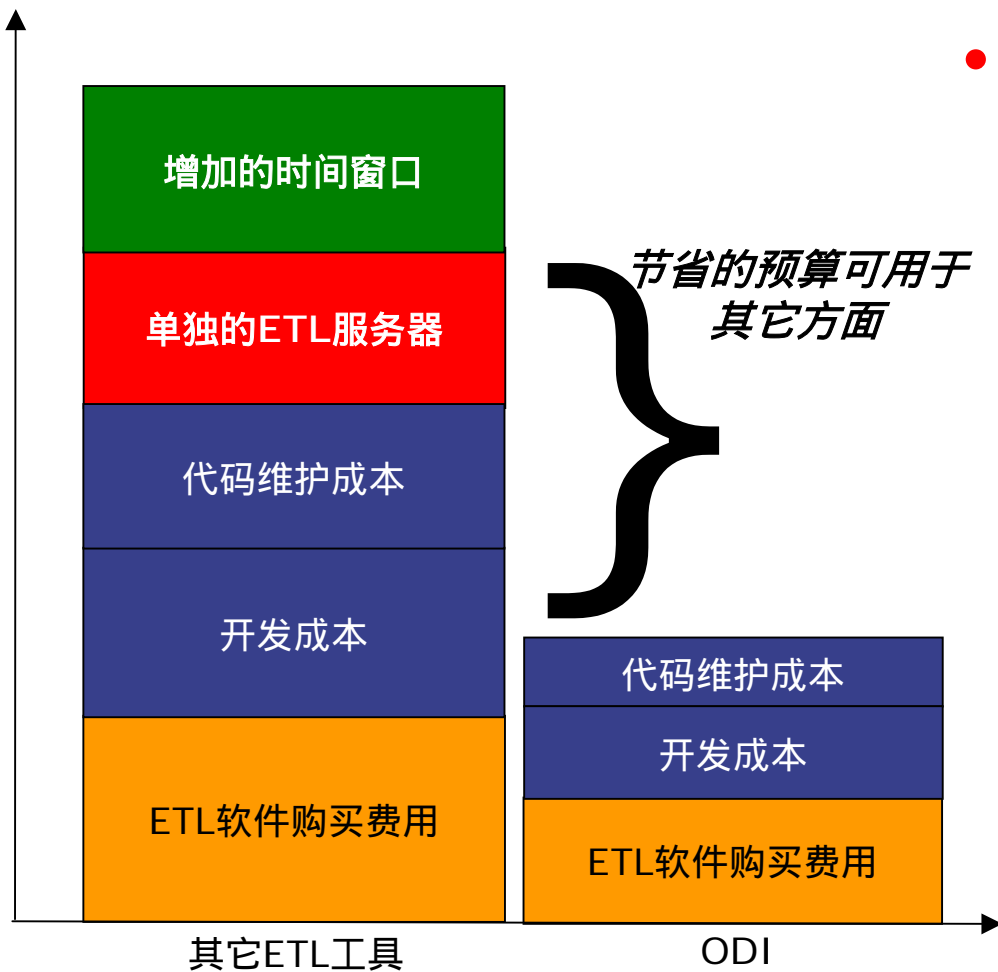
- **Productivity**
  - 实时的集成平台
  - 高效的声明式设计
  - 知识模块，热插拔
- **Performance**
  - E-LT 高性能架构
- **Price**
  - 无中间服务器，TCO低

# 完整的信息集成方案

- 批量：
  - 定时执行
- 事件触发
  - 变化数据捕获
  - 准实时的数据采集
- 面向服务
  - 作为Web服务嵌入到SOA应用



# 更低的总拥有成本



- 对客户而言：

- **软件**：使用已有的关系数据库，而部署专有的ETL引擎
- **技能**：利用已有的关系数据库和SQL技能
- **硬件**：不需要添加专有的ETL服务器
- **机会成本**：增加的时间窗口

# 更多信息

- 下载地址：

- <http://www.oracle.com/technology/software/products/ias/htdocs/101310.html>

- OTN上的技术信息：

- <http://www.oracle.com/technology/products/oracle-data-integrator/index.html>

- 技术讨论组：

- <http://forums.oracle.com/forums/forum.jspa?forumID=374>



Q&A

ORACLE®

甲骨文