

ORACLE®

THE INFORMATION COMPANY™

Инструментальная среда Oracle Warehouse Builder для разработки хранилищ данных



Андрей Пивоваров

Консультант по хранилищам данных
и аналитическим системам

Что такое Oracle Warehouse Builder?

- Warehouse Builder – средство проектирования, развертывания и управления корпоративным хранилищем данных.

Работа в OWB

1 Разработка целевых объектов

3 Разработка схем преобразований

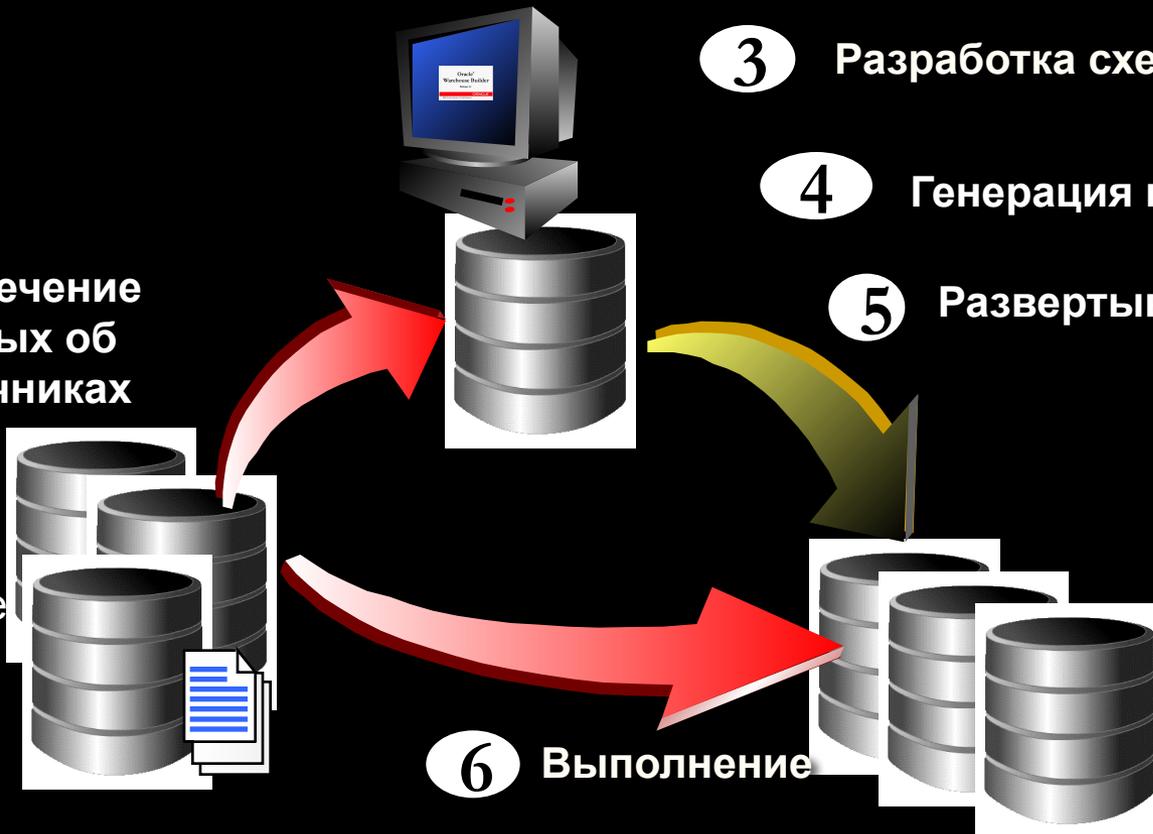
4 Генерация кода и контроль

5 Развертывание кода

6 Выполнение

2 Извлечение данных об источниках

- Реляционные
- Файлы
- Приложения



- Oracle DB
- OLAP
- Файлы
- AQ
- BI tools

ORACLE

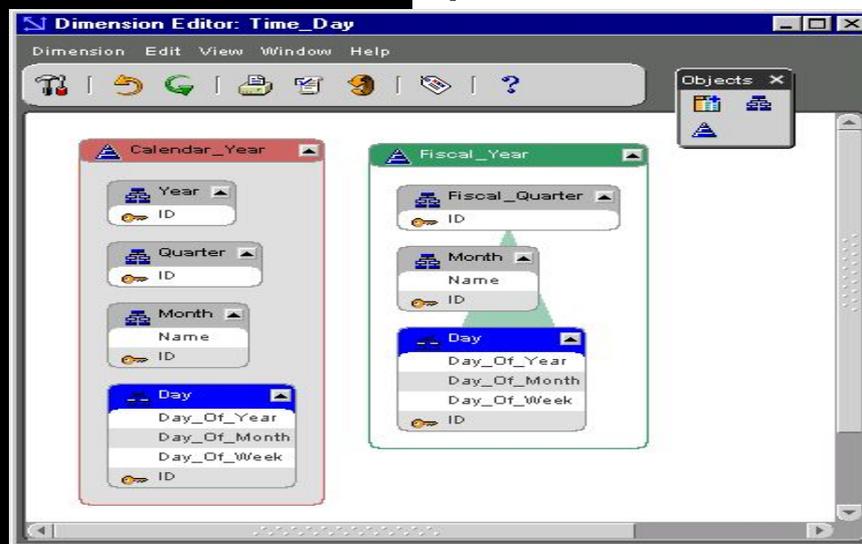
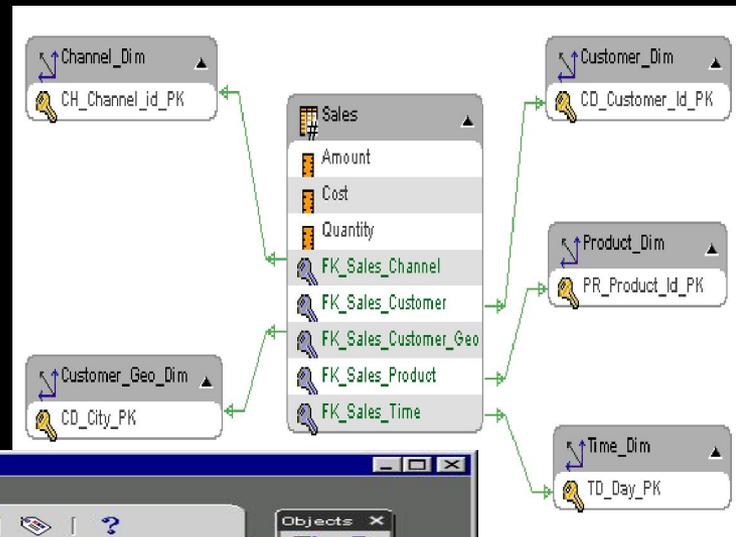
Создание хранилища данных

- Проектирование хранилища
- Спецификация источников
- Проектирование процедур загрузки данных
- Генерация кода и развертывание хранилища
- Администрирование и поддержка

Проектирование хранилища



- Средства визуального проектирования
- Реляционное моделирование
- Многомерное моделирование



Создание хранилища данных

- Проектирование хранилища
- Спецификация источников
- Проектирование процедур загрузки данных
- Генерация кода и развертывание хранилища
- Администрирование и поддержка

Информационные источники

- Реляционные СУБД:
 - Oracle
 - IBM DB2
 - SQL Server
 - ...
- Текстовые файлы
- ERP системы
 - SAP

Создание хранилища данных

- Проектирование хранилища
- Спецификация источников
- Проектирование процедур загрузки данных
- Генерация кода и развертывание хранилища
- Администрирование и поддержка

Отображение данных

Mapping Editor: MAP_CHANNELS "Read/Write"

Mapping Edit View Debug Window Help



| CHANNEL... | | |
|------------|---|-----|
| INOUTGRP1 | | |
| ID | a | b_c |
| CLASS | a | b_c |

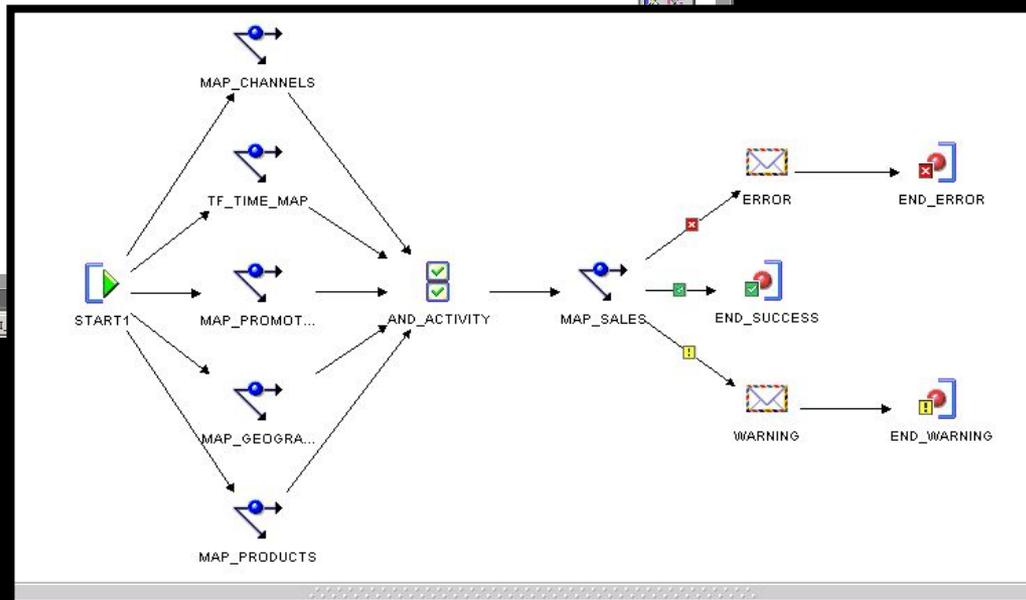
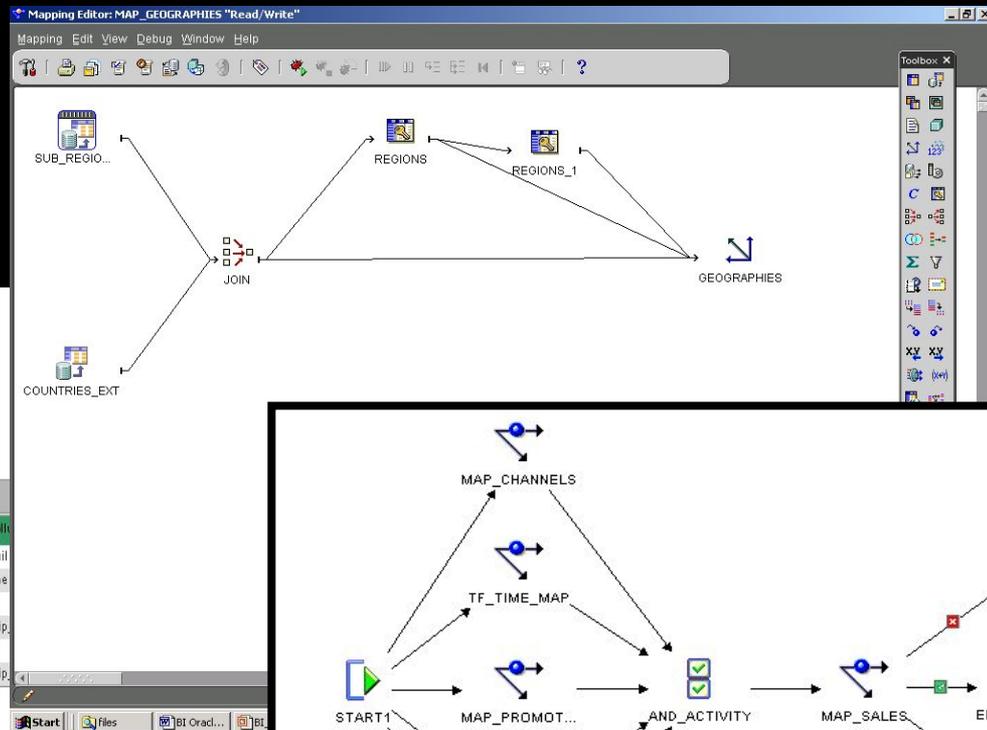
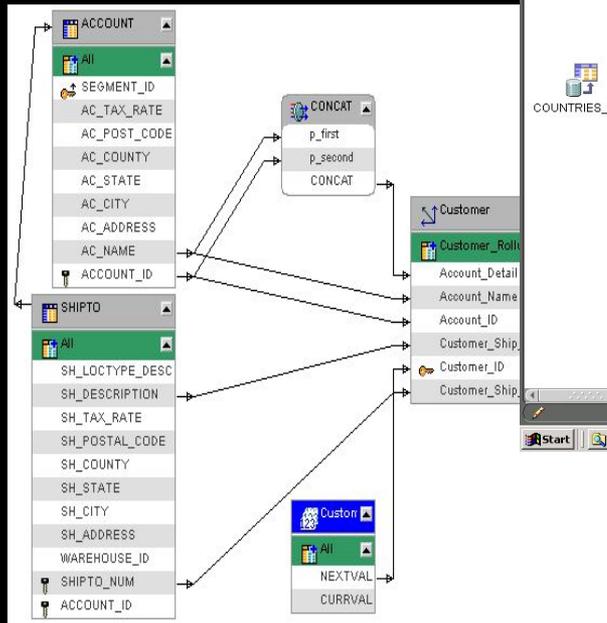
| JOIN | | |
|-------------|---|-----|
| INGRP1 | | |
| ID | a | b_c |
| CLASS | a | b_c |
| INGRP2 | | |
| ID | a | b_c |
| CLASS_ID | a | b_c |
| DESCRIPTION | a | b_c |
| OUTGRP1 | | |
| ID | a | b_c |
| CLASS | a | b_c |
| ID_1 | a | b_c |
| CLASS_ID | a | b_c |
| DESCRIPTION | a | b_c |

| CHANNELS | | |
|-----------|---|-----|
| INOUTGRP1 | | |
| CLS_ID | a | b_c |
| CLS_NAME | a | b_c |
| CNL_ID | a | b_c |
| CNL_NAME | a | b_c |
| TOT_ID | a | b_c |
| TOT_NAME | a | b_c |

| CHANNELS_1 | | |
|-------------|---|-----|
| INOUTGRP1 | | |
| ID | a | b_c |
| CLASS_ID | a | b_c |
| DESCRIPTION | a | b_c |

| CONST | | |
|---------|---|-----|
| OUTGRP1 | | |
| TOTAL | a | b_c |

Потоки данных и процессы загрузки



Создание хранилища данных

- Проектирование хранилища
- Спецификация источников
- Проектирование процедур загрузки данных
- Генерация кода и развертывание хранилища
- Администрирование и поддержка

Генерация скриптов и процедур

- DDL - скрипты
- PL/SQL пакеты
- Управляющие файлы для SQL Loader
- XML Process Definition Language (XPDL)
- ABAP code

OMB Scripting

- Доступ к метаданным OWB через командную строку
- Напоминает SQL*Plus для Oracle Database
- Создание, извлечение, изменение, удаление метаданных в пакетном режиме
- Основан на TCL
- Интеграция с JDeveloper IDE

Deployment Manager

- Привязка источников и целевых схем к физическим базам данных
- Генерация кода
- Создание объектов хранилища
- Аудит генерации и выполнения скриптов и процедур

Создание хранилища данных

- Проектирование хранилища
- Спецификация источников
- Проектирование процедур загрузки данных
- Генерация кода и развертывание хранилища
- Администрирование и поддержка

Средства администрирования

Microsoft Internet Explorer
 Address: http://127.0.0.1:8090/owb/PLSQLMapRunExecution.uik?event=navigate&p_id=3052&rcpos=1

ORACLE
 Warehouse Builder Browser - Runtime Repository

Execution | File | Trace

owb > MAP_CHANNELS > MAP_CHANNELS
MAP_CHANNELS - PL/SQL Map Run Execution Report

Map Name: **MAP_CHANNELS** Type: **PL/SQL Map** Refresh

Execution Details

| | | | |
|------------------------|-------------------------|------------------|--------------|
| Start Time | 08 Дек 2003 20:52:30 | Process Name | MAP_CHANNELS |
| End Time | 08 Дек 2003 20:52:39 | Status | Complete |
| Elapsed Time (seconds) | 9 | Data Error Count | 1 |
| Source | DEVDAVS.CHANNELS | Database Link | |
| Source | DEVDAVS.CHANNEL_CLASSES | Database Link | |
| Target | CHANNELS_1 | | |

Execution Parameters

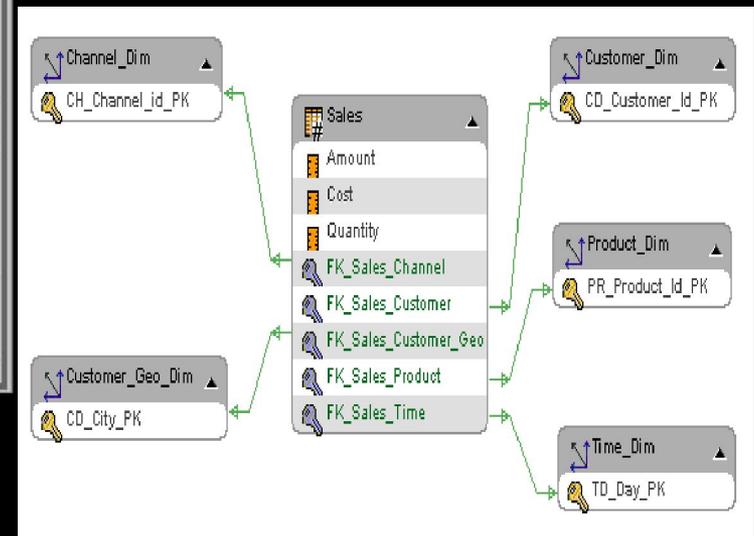
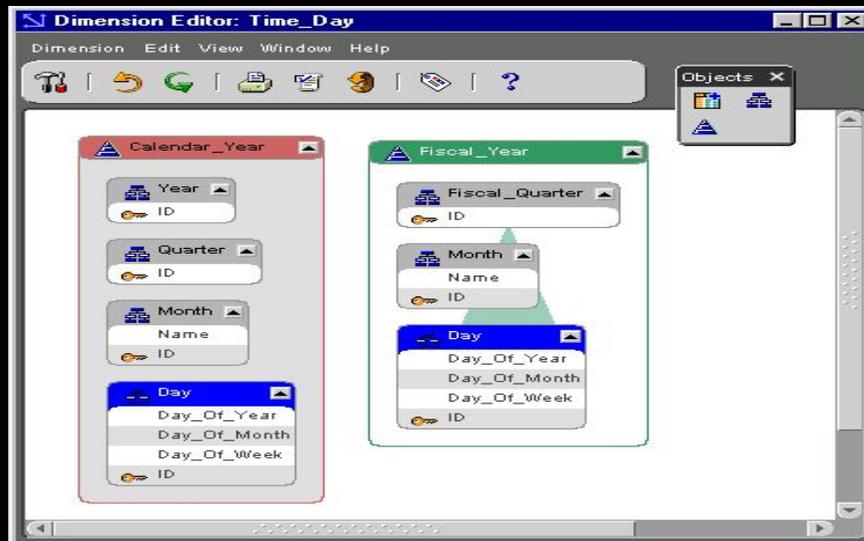
| Category | Name | Input Value | Output Value |
|----------|-------------------|-------------|--------------|
| System | Audit Level | 2 | |
| System | Bulk Size | 50 | |
| System | Commit Frequency | 1000 | |
| System | Max No. of Errors | 50 | |
| System | Operating Mode | 3 | |
| System | Procedure Name | map_r | |
| System | Prng Group | wh | |

Step Details

| Сведения | Map Step | Step Name | Step Type | Start Time | End Time | Elapsed Time (seconds) | Status | Data Error Count |
|----------|----------|------------|------------|----------------------|----------------------|------------------------|----------|------------------|
| ▶ | 0 | CHANNELS_1 | Self-Execd | 08 Дек 2003 20:52:30 | 08 Дек 2003 20:52:39 | 9 | Complete | 1 |

ДЕМОНСТРАЦИЯ

Oracle Warehouse Builder



НОВЫЕ ВОЗМОЖНОСТИ OWB 10g R2 “Paris”

Качество данных

- Data Profiling
- Data Enrichment through Oracle Data Mining
- Новые источники данных
- Real-Time Data Integration
- Transportable Modules and CDC
- Scheduled Process Execution
- Custom Metadata Capture
- Медленно меняющиеся размерности

Бизнес-анализ

- Реляционный и многомерный дизайн
- Разработка объектов Business Intelligence (Discoverer, Mining)

Эксперты

- Experts
- Personalities

ДЕМОНСТРАЦИЯ

Oracle Warehouse Builder 10g R2

The left screenshot shows the 'Profile Results Canvas' for table PROD, displaying aggregation analysis results in a table:

| Item | Maximum | Average |
|-------------------------------|---------|---------|
| Wrought Iron, Bar Stock | 498 | 221 |
| 182000 | 860 | 373 |
| SVC | 7 | 11 |
| Carlen, Mr. Paul David (Paul) | 16 | 12 |
| Vice President, Materials | | |

The right screenshot shows the 'Profile Results Canvas' for column S27CAT_NAME, displaying a bar chart of distinct values. The chart shows the frequency of distinct values for the column S27CAT_NAME. The series is 'frequency' and the group is 'Component - Electrical Part' with a value of 137.0.

The 'Data Drill Panel' in the right screenshot shows the left side below shows all distinct values of column S27CAT_NAME in table PROD. The right side below shows all rows of PROD which have the value of S27CAT_NAME that is selected on the left.

| ITEM_ID | ITEM_TYPE | ITEMNAME | ITEMORG... | ITEMPT_N... | MGREMP... | |
|---------|-----------|----------|----------------|-------------|-----------|---|
| 1 | 436 | P | Printer | 229 | CM76840 | 7 |
| 2 | 193 | P | RAM - 32 MB | 229 | CM60257 | 7 |
| 3 | 199 | P | Monitor 19" | 229 | CM31556 | 7 |
| 4 | 253 | P | Keyboard - ... | 229 | CM18761 | 7 |
| 5 | 225 | P | Monitor 17" | 229 | CM15140 | 7 |
| 6 | 245 | P | RAM - 16 MB | 210 | CM99105 | 7 |
| 7 | 199 | P | Monitor 19" | 210 | CM31556 | 7 |
| 8 | 201 | P | Keyboard - ... | 210 | CM18759 | 7 |

Data Profiling

The screenshot displays the Oracle Data Profile Manager interface, showing the results of a data profiling operation on the table PROD. The main window is titled 'Data Profile Manager: PROFILE_MOD' and features a 'Profile Results Canvas' with a bar chart and a 'Data Drill Panel'.

Profile Results Canvas: The chart displays the frequency of distinct values for the column S27CAT_NAME. The Y-axis represents frequency, ranging from 0 to 320. The X-axis lists various categories. The 'Component - Electrical Part' category has the highest frequency, with a value of 137.0.

Data Drill Panel: The panel shows the distinct values of S27CAT_NAME in table PROD. The left side lists the distinct values, and the right side shows all rows of PROD which have the value of S27CAT_NAME that is selected on the left.

Table 1: Distinct Values of S27CAT_NAME

| S27CAT_NAME | Frequency |
|-----------------------------------|-----------|
| Miscellaneous - Miscellaneous | 24 |
| Component - Electrical Part | 137 |
| Assembly - Final Assembly | 96 |
| Assembly - Subassembly | 70 |
| Component - Miscellaneous Part | 56 |
| Assembly - Miscellaneous Assem... | 44 |
| Component - Mechanical Part | 36 |
| Component - Paper Goods Part | 9 |

Table 2: Rows of PROD with S27CAT_NAME = Component - Electrical Part

| ITEM_ID | ITEM_TYPE | ITEMNAME | ITEMORG... | ITEMPT_N... | MGREMP... | |
|---------|-----------|----------|----------------|-------------|-----------|---|
| 1 | 436 | P | Printer | 229 | CM76840 | 7 |
| 2 | 193 | P | RAM - 32 MB | 229 | CM60257 | 7 |
| 3 | 199 | P | Monitor 19"... | 229 | CM31556 | 7 |
| 4 | 253 | P | Keyboard - ... | 229 | CM18761 | 7 |
| 5 | 225 | P | Monitor 17"... | 229 | CM15140 | 7 |
| 6 | 245 | P | RAM - 16 MB | 210 | CM99105 | 7 |
| 7 | 199 | P | Monitor 19"... | 210 | CM31556 | 7 |
| 8 | 201 | P | Keyboard - ... | 210 | CM18759 | 7 |

Вложенные преобразования (Nested mappings)

- Возможность выносить часто встречающиеся элементы преобразований в отдельные компоненты
- Упрощение создания новых преобразований
- Уменьшение сложности для понимания преобразований

Эксперты (Experts)

The screenshot displays the Oracle Expert Editor interface for a workflow named "FILE_TO_TABLE_1 'Read/Write'".

Workflow Diagram: The main workspace shows a sequence of tasks: Start Task → Select Source → Import → Select Target. From "Select Target", the flow branches into two paths: one leading to a "table" icon and another to a "file" icon. Both paths converge into a "Data_viewer Task", which then leads to an "End Task".

Explorer: A tree view on the left lists various tasks and their parameters. The "host" parameter is currently selected.

Object Inspector: A table below the Explorer shows details for the selected "host" object.

| Property | Value |
|-------------|------------|
| NAME | host |
| DIRECTION | OUT |
| DATATYPE | STRING |
| BINDING | g_src_host |
| VALUE | |
| DESCRIPTION | |

Task Editor: Select Source: A window at the bottom shows the configuration for the "Select Source" task. The "Instructions" tab is active, displaying the following code:

```
1 puts $g_src_loc
2 puts $g_src_object_type
3 puts $g_dir
4 puts $g_src_module
5
```

Palette: A vertical toolbar on the right contains various task icons such as "Advanced Queue", "Cube", "Dimension", "External Table", "File", "Function", "Mapping", "Materialized View", "Procedure", "Process Flow", "Sequence", "Sub-map", "Table", "View", "OMB", "Java", "Sub-task-flow", "Custom Dialog", "Object Selector", "Define Source", "Define Target", "Commit", "Data Viewer", "Deploy", "Derivation", "Generation", "Impact Analysis", "Change Manager", "Source Import", and "Validation".

Metadata Integration Views

MIV Source Definition Editor: Ascential MIV "Read/Write"

File Help

Ascential MIV

- Data Object Views
 - Tables
 - Views
 - Sequences
- Supporting Views
 - Columns
 - Keys
 - Foreign Keys
 - Key Columns
 - Foreign Key Tab
- Functional Area View
 - Functional Areas
 - Functional Area
 - Functional Area
- Advanced Options
- Verification

Choose whether it is SQL based or File based:

SQL based File based

Enter your Tables view definition:

```
select TABLE_NAME, BUSINESS_NAME, DESCRIPTION
from ...<Parameter name="dblink" />
where ...
```

Number of rows shown: 20 Test

Test Results:

| TABLE_NAME | BUSINESS_NAME | DE |
|------------|---------------|----|
| | | |
| | | |

Real-Time warehousing

- Использует механизм Change Data Capture
- Поддержка Oracle Streams
- Возможность частого обновления хранилища (даже раз в минуту)

Технология OWB для BI

- Создание Аналитических Пространств (AW)
- Заполнение OLAP каталога
- Генерация CWM/CWM2 и AWXML скриптов
- Поддержка BI Beans
- Интеграция с EUL для Discoverer
- Поддержка Oracle Data Mining

НОВЫЕ ВОЗМОЖНОСТИ OWB 10g R2 “Paris”

Качество данных

- Data Profiling
- Data Enrichment through Oracle Data Mining
- Новые источники данных
- Real-Time Data Integration
- Transportable Modules and CDC
- Scheduled Process Execution
- Custom Metadata Capture
- Медленно меняющиеся размерности

Бизнес-анализ

- Реляционный и многомерный дизайн
- Разработка объектов Business Intelligence (Discoverer, Mining)

Эксперты

- Experts
- Personalities

Дополнительная информация

<http://otn.oracle.com/products/bi/index.html>

Andrey.Pivovarov@oracle.com

A large, stylized logo in the background consisting of a grey 'Q', a red ampersand '&', and a grey 'A'. The text 'ВОПРОСЫ И ОТВЕТЫ' is overlaid on the ampersand.

**ВОПРОСЫ И
ОТВЕТЫ**

ORACLE®

THE INFORMATION COMPANY™