

Tecnologias Oracle para DW

Oracle Data Integrator - ETL

Processamento Analítico de Dados
Profa. Dra. Cristina Dutra de Aguiar Ciferri
Bruno Tomazela
Renata Miwa Tsuruda

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Oracle – Edições dos DBs Oracle

- SE1 – Standard Edition One
- SE – Standard Edition
- EE – Enterprise Edition
- PE – Personal Edition
- XE – Express Edition

- Obs: Apenas a edição EE possui os atributos necessários para a utilização plena de Data Warehouse.

O que se Espera de uma Ferramenta ETL e o que o Oracle Pode Atender

- Extrair dados de diversas fontes e plataformas
- Alta performance e processamento distribuído
- Alta escalabilidade
- Compactação de dados e criptografia
- Manipular todos os tipos de dados
- Repositório de metadados
- Schedule / Cargas Event-Driven

O que se Espera de uma Ferramenta ETL e o que o Oracle Pode Atender

- Interface gráfica de desenvolvimento
- Arquitetura aberta
- Geração / Reengenharia de códigos

Oracle Data Integration Suite

- Cria, implanta e gerencia data warehouses complexos
- Informatizar a migração e a movimentação dos dados em lote
- Assegura que as informações sejam oportunas, precisas e consistentes entre sistemas complexos

Oracle Data Integration Suite

- Produtos ETL da suite
 - Oracle Data Integrator (ODI)
 - Oracle Data Integration Enterprise Edition
 - Oracle Data Quality and Profiling
 - Oracle Changed Data Capture (CDC)

Agenda

- Oracle
- **Oracle Data Integrator (ODI)**
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Oracle Data Integrator (ODI)

- Flexibilidade
 - SGBD, XML, File, Excel e muitos outros
 - Tratamento para módulos de aplicações comerciais existentes
- Alta Performance:
 - E-LT – Extract, Load and Transform
 - Changed Data Capture

Oracle Data Integrator (ODI)

- Controle de restrições e consistência
 - Geração de regras de integridade
- Permite a verificação do fluxo de dados processados pelas interfaces
 - Isolamento de erros e/ou reciclagem
- Identifica falta de entrada de dados

Oracle Data Integrator (ODI)

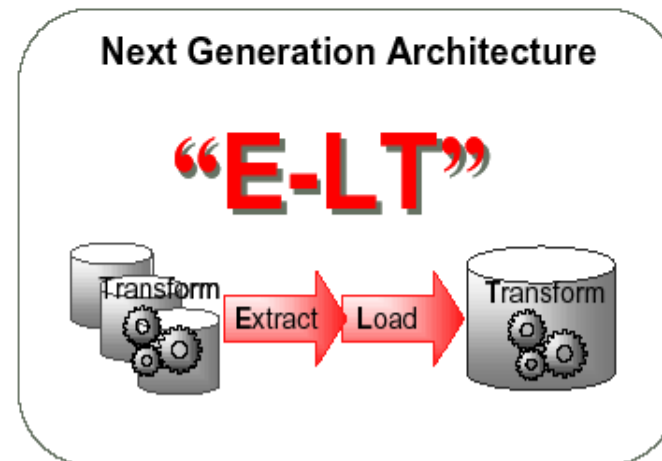
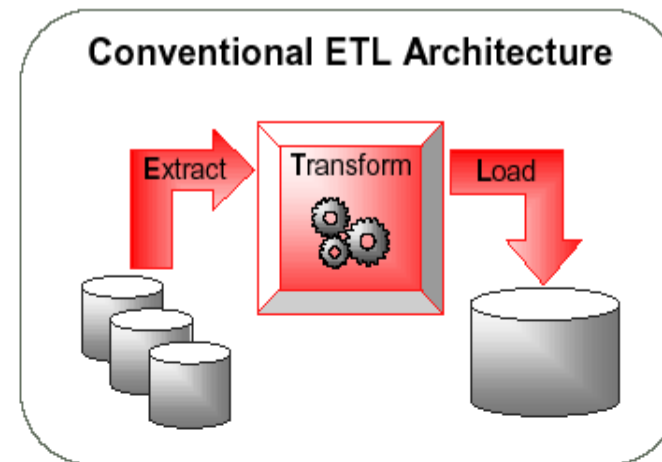
- Interoperabilidade
 - Suporta todos os padrões Java e SOA
 - Hot-Pluggable (Knowledge Modules)
- Produtividade
 - Declarative Design
 - Knowledge Modules

Oracle – Arquitetura E-LT

- Código nativo para todos os DBs suportados
- Extrai, transforma e carrega utilizando a potência dos próprios bancos source / target
- Eficiência sem paralelo
- Reduz efetivamente TCO

Oracle E-LT

- Extraction
 - Extrai e executa parte da transformação utilizando a engine do DB source. Em seguida, carrega na memória do servidor de extração
- Load
 - Carrega os dados já transformados no DW
- Transformation
 - Para cada tupla carregada, a engine do ETL finaliza o processo de transformação



ETL vs. E-LT Approach

Diferenciais – Declarative Design e Knowledge Modules

- Ambiente virtual de desenvolvimento
 - Declarative Design
 - “O que” será extraído e transformado
 - Knowledge Modules (Templates)
 - “Como” essas operações deverão ser realizadas face às várias plataformas e bancos de dados envolvidos no processo

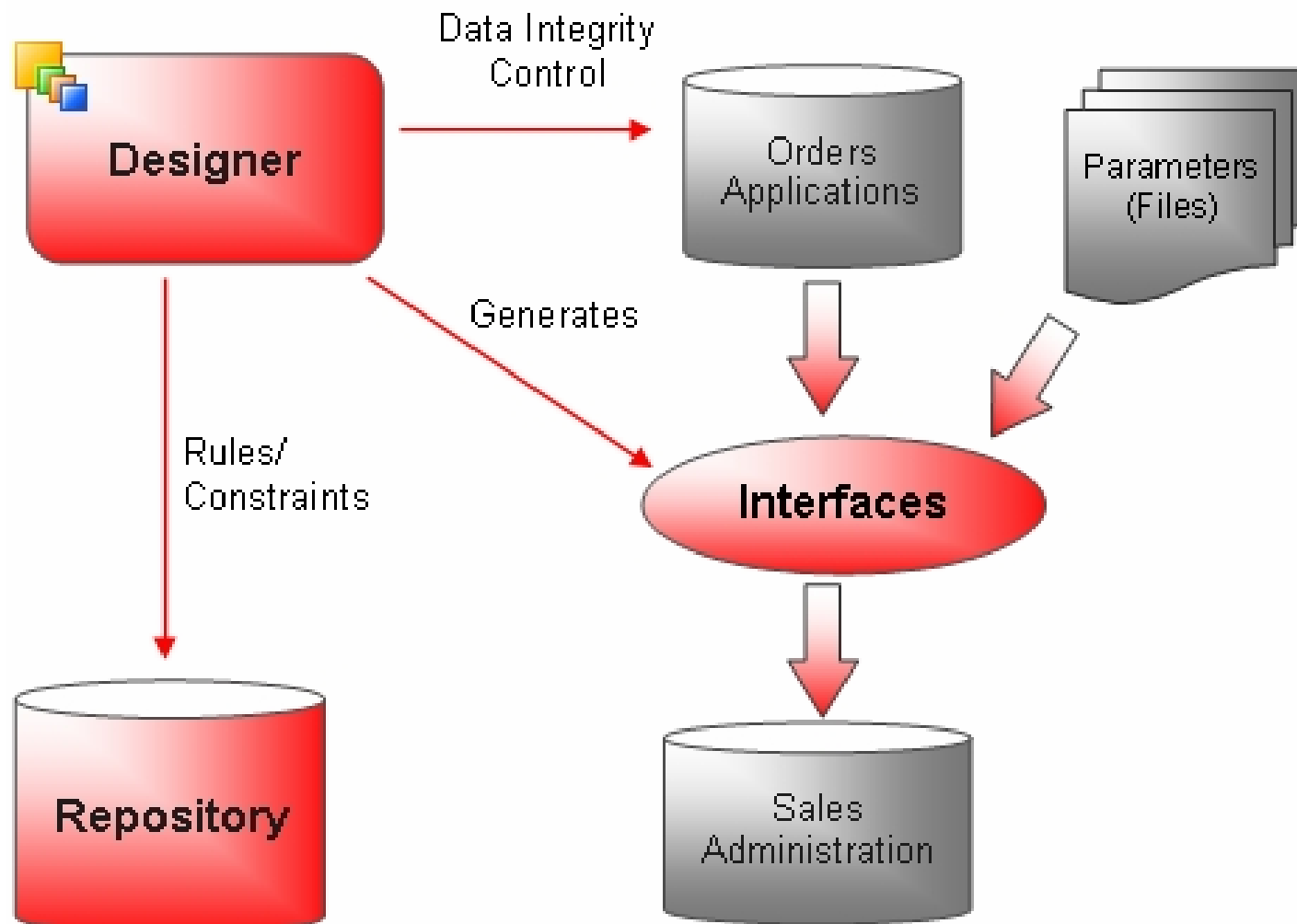
Agenda

- Oracle
- Oracle Data Integrator (ODI)
- **Estudo de Caso**
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Demonstração do Oracle Data Integrator

- Tutorial *Oracle Data Integrator: Getting Started with na ETL Project* [Oracle, 2008]
 - Limpeza dos dados por meio de regras de integridade
 - Transformação e carga de dados em um DW usando interfaces
 - Automatização do processo de ETL

Estudo de Caso: Administração de Vendas



Orders Application (HSQLDB)

SRC_REGION			
<u>REGION_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
REGION	VARCHAR(50)		null
COUNTRY_ID	NUMERIC(10)		null
COUNTRY	VARCHAR(50)		null

SRC_CITY			
<u>CITY_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
CITY	VARCHAR(50)		null
REGION_ID	NUMERIC(10)		null
POPULATION	NUMERIC(10)		null

SRC_ORDER_LINES			
<u>ORDER_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
<u>LORDER_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
PRODUCT_ID	NUMERIC(10)		null
QTY	NUMERIC(10)		null
AMOUNT	NUMERIC(10,2)		null

SRC_CUSTOMER			
<u>CUSTID</u>	NUMERIC(10)	<pk>	<u>not null</u>
DEAR	NUMERIC(1)		null
LAST_NAME	VARCHAR(50)		null
FIRST_NAME	VARCHAR(50)		null
ADDRESS	VARCHAR(100)		null
CITY_ID	NUMERIC(10)		null
PHONE	VARCHAR(50)		null
AGE	NUMERIC(3)		null
SALES_PERS_ID	NUMERIC(10)		null

SRC_ORDERS			
<u>ORDER_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
STATUS	VARCHAR(3)		null
CUST_ID	NUMERIC(10)		null
ORDER_DATE	DATE		null
CUSTOMER	VARCHAR(35)		null

SRC_PRODUCT			
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk>	<u>not null</u>
PRODUCT	VARCHAR(50)		null
PRICE	NUMERIC(10,2)		null
FAMILY_NAME	VARCHAR(50)		null

Parameters (arquivo texto CSV)

SRC_SALES_PERSON			
<u>SALES_PERSON_ID</u>	<u>NUMERIC(10)</u>	<u><pk></u>	<u>not null</u>
FIRST_NAME	VARCHAR(50)		null
LAST_NAME	VARCHAR(50)		null
HIRE_DATE	DATE		null

SRC_AGE_GROUP			
<u>AGE_MIN</u>	<u>NUMERIC(3)</u>	<u><pk></u>	<u>not null</u>
<u>AGE_MAX</u>	<u>NUMERIC(3)</u>	<u><pk></u>	<u>not null</u>
AGE_RANGE	VARCHAR(50)		null

Sales Administration (DW)

TRG_COUNTRY			
<u>COUNTRY_ID</u>	NUMERIC(10)	<pk>	not null
COUNTRY	VARCHAR(50)		null

FK_REGION_COUNTRY

TRG_REGION			
<u>REGION_ID</u>	NUMERIC(10)	<pk>	not null
COUNTRY_ID	NUMERIC(10)	<fk>	not null
REGION	VARCHAR(50)		null

FK_CITY_REGION

TRG_CITY			
<u>CITY_ID</u>	NUMERIC(10)	<pk>	not null
REGION_ID	NUMERIC(10)	<fk>	not null
CITY	VARCHAR(50)		null
POPULATION	NUMERIC(10)		null

FK_CUST_CITY

TRG_CUSTOMER			
<u>CUST_ID</u>	NUMERIC(10)	<pk>	not null
DEAR	VARCHAR(50)		null
CUST_NAME	VARCHAR(50)		null
ADDRESS	VARCHAR(50)		null
CITY_ID	NUMERIC(10)	<fk>	not null
PHONE	VARCHAR(50)		null
AGE	NUMERIC(10)		null
AGE_RANGE	VARCHAR(50)		null
SALES_PERS	VARCHAR(50)		null
CRE_DATE	DATE		null
UPD_DATE	DATE		null

TRG_PROD_FAMILY			
<u>FAMILY_ID</u>	VARCHAR(50)	<pk>	not null
FAMILY_NAME	VARCHAR(50)		null

FK_PROD_PROD_FAMILY

TRG_PRODUCT			
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk>	not null
FAMILY_ID	VARCHAR(50)	<fk>	not null
PRICE	NUMERIC(10)		null
PRODUCT	VARCHAR(50)		null

FK_SALES_PROD

TRG_SALES			
<u>CUST_ID</u>	NUMERIC(10)	<pk, fk1>	not null
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk, fk2>	not null
FIRST_ORD_ID	NUMERIC(10)		not null
FIRST_ORD_DATE	DATE		not null
LAST_ORD_ID	NUMERIC(10)		not null
LAST_ORD_DATE	DATE		not null
QTY	NUMERIC(10)		not null
AMOUNT	NUMERIC(10)		not null
PROD_AVG_PRICE	NUMERIC(10)		not null

FK_SALES_CUST

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- **Inicializando o ODI**
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Inicializando o Oracle Data Integrator

- Iniciar o ambiente de demonstração
 - Iniciar > Programas > Oracle > Oracle Data Integrator > Examples > Start Demo Environment
- Iniciar o Designer
 - Iniciar > Programas > Oracle > Oracle Data Integrator > Designer
- Para fechar o ambiente
 - Iniciar > Programas > Oracle > Oracle Data Integrator > Examples > Stop Demo Environment

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- **Limpeza dos Dados**
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Limpeza dos Dados

- Verificar a consistência dos dados das fontes
 - Restrições de integridade
 - Regras de negócio
- Permite analisar os erros identificados
- Corrigir erros caso necessário

Limpeza dos Dados

- Restrições do estudo de caso
 - Idade dos clientes deve ser maior que 21 anos
 - Restrição de chave estrangeira:

SRC_CITY			
<u>CITY_ID</u>	NUMERIC(10)	<pk>	not null
CITY	VARCHAR(50)		null
REGION_ID	NUMERIC(10)		null
POPULATION	NUMERIC(10)		null

FK

SRC_CUSTOMER			
<u>CUSTID</u>	NUMERIC(10)	<pk>	not null
DEAR	NUMERIC(1)		null
LAST_NAME	VARCHAR(50)		null
FIRST_NAME	VARCHAR(50)		null
ADDRESS	VARCHAR(100)		null
<u>CITY_ID</u>	NUMERIC(10)		null
PHONE	VARCHAR(50)		null
AGE	NUMERIC(3)		null
SALES_PERS_ID	NUMERIC(10)		null

Limpeza dos Dados

- Criar as restrições na tabela SRC_CUSTOMER:
 - Condição: Age > 21
 - Referência: FK_SRC_CUSTOMER_SRC_CITY
- Executar as restrições
- Verificar o relatório de erros

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- **Transformação dos Dados**
- Interface de Integração dos Dados
- Automatização
- Conclusão
- Referências

Transformação dos Dados

- Quais transformações queremos:
 - Obter TRG_CUSTOMER.AGE_RANGE a partir de SRC_AGE_GROUP.AGE_RANGE e SRC_CUSTOMER.AGE

Idade do cliente
(SRC_CUSTOMER - HSQL)

X



Faixa de idade do cliente
(TRG_CUSTOMER - HSQL)

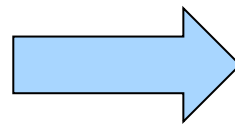
Faixa de Idade
(SRC_SALES_PERS - txt)

Transformação dos Dados

- Quais transformações queremos:
 - Obter TRG_CUSTOMER.SALES_PERS de SRC_SALES_PERS.FIRST_NAME e SRC_SALES_PERS.LAST_NAME a partir de SRC_CUSTOMER.SALES_PERS_ID

Primeiro e último nome do representante de vendas
(SRC_SALES_PERS - txt)

X



Nome do representante
(TRG_CUSTOMER - HSQL)

Id do representante de vendas
(SRC_CUSTOMER - HSQL)

Transformação dos Dados

- Quais transformações queremos:
 - Transformar o valor numérico (0, 1, 2) presente em SRC_CUSTOMER.**DEAR** em uma das strings padrão (Mr, Mrs ou Ms) e salvar em TRG_CUSTOMER.**DEAR**

0, 1, 2
(SRC_CUSTOMER - HSQL)

X

0 – Mr
1 – Mrs
2 – Ms



Pronome de tratamento
(TRG_CUSTOMER - HSQL)

Transformação dos Dados

- Quais transformações queremos:
 - Obter TRG_CUSTOMER.**CUST_NAME** a partir de SRC_CURSTOMER.**FIRST_NAME** e SRC_CURSTOMER.**LAST_NAME**

Primeiro e último nome do cliente
(SRC_CUSTOMER - HSQL)



Nome do cliente
(TRG_CUSTOMER - HSQL)

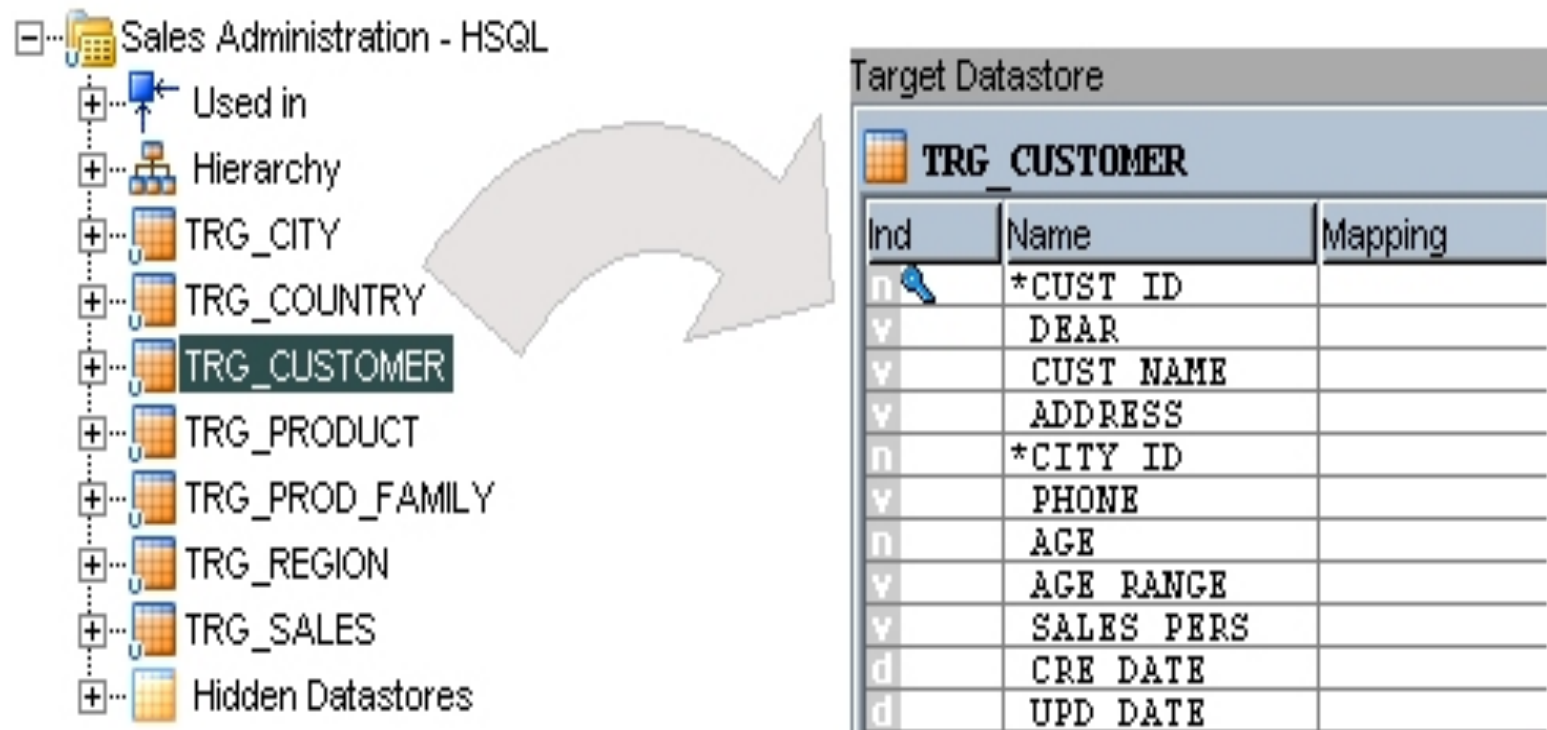
Transformação dos Dados

The image illustrates the process of creating a new interface for data transformation in SAP Business Objects. On the left, a tree view shows the project structure under 'Demo' > 'Sales Administration' > 'Interfaces'. A context menu is open over the 'Interfaces' folder, with 'Open in New Window' and 'Insert Interface' options. Below the tree, a small preview window shows the 'Definition' tab of a new interface named 'Pop. TRG_CUSTOMER'. On the right, the main 'Interface: NEW' dialog box is shown, with the 'Definition' tab selected. The 'Sources' area is empty, and the 'Target Datastore' area contains an 'Untitled' table with the following structure:

Ind	Name	Mapping
-----	------	---------

The 'Properties' section at the bottom of the dialog is empty, with the text 'Select an element in the diagram to display its properties'. The dialog includes 'OK', 'Cancel', 'Apply', 'Help', and 'Execute' buttons.

Transformação dos Dados

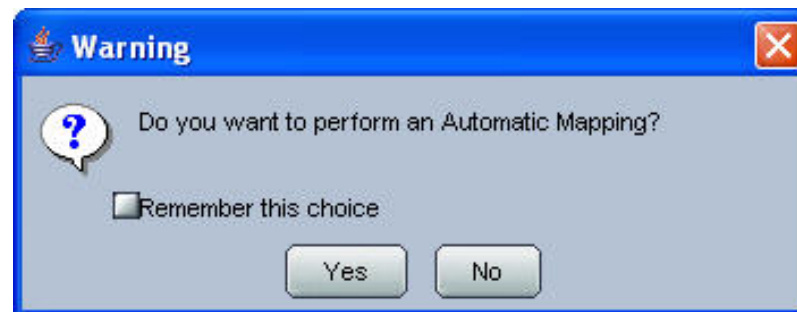
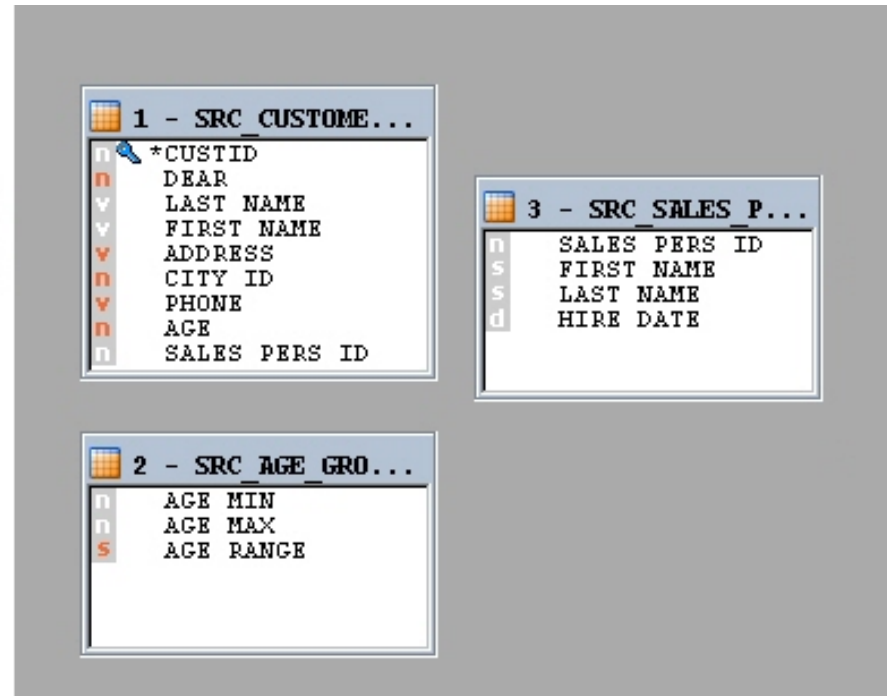
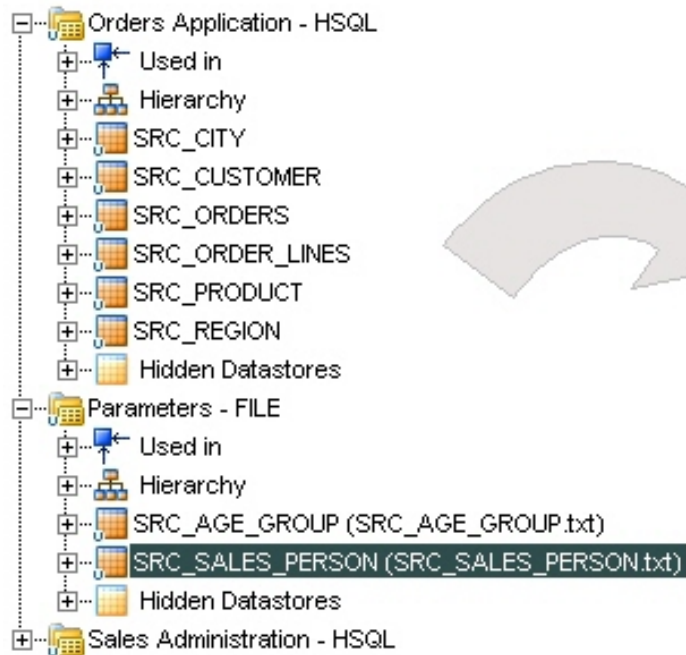


The image shows a software interface for data transformation. On the left, a tree view displays a folder named "Sales Administration - HSQL" containing several datastores: "Used in", "Hierarchy", "TRG_CITY", "TRG_COUNTRY", "TRG_CUSTOMER" (highlighted), "TRG_PRODUCT", "TRG_PROD_FAMILY", "TRG_REGION", "TRG_SALES", and "Hidden Datastores". A large grey arrow points from the "TRG_CUSTOMER" datastore to a detailed view on the right.

The detailed view, titled "Target Datastore", shows the structure of the "TRG_CUSTOMER" datastore. It is a table with three columns: "Ind", "Name", and "Mapping". The rows are as follows:

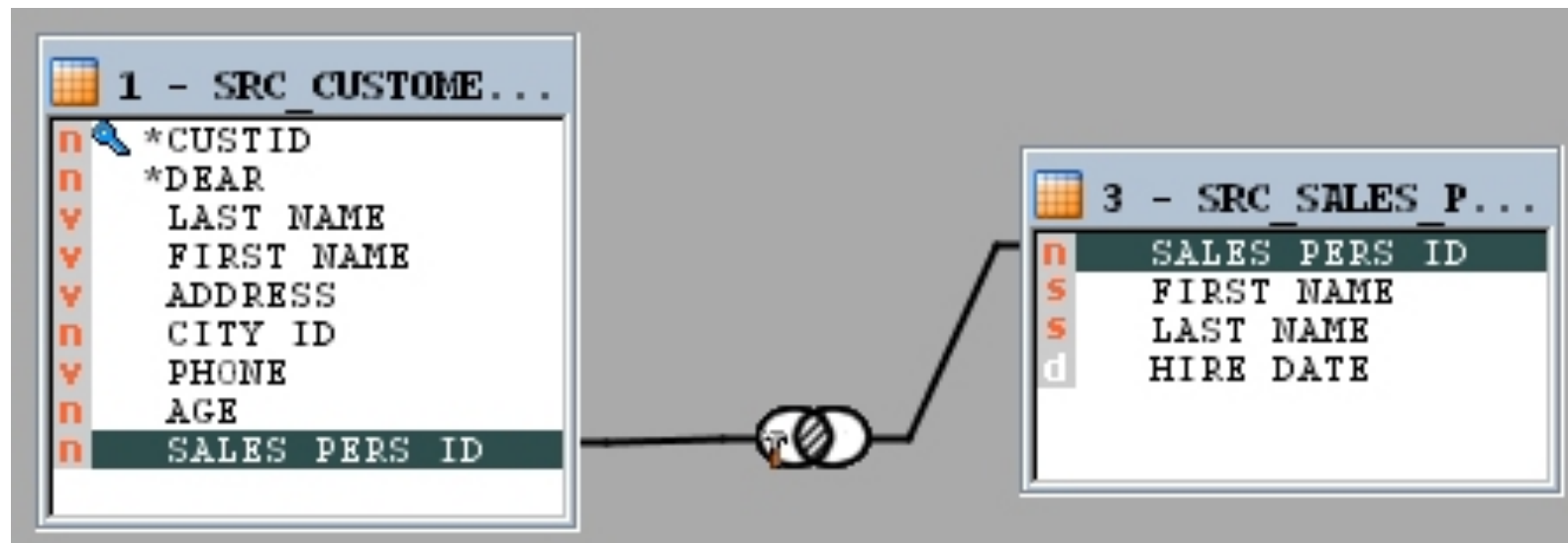
Ind	Name	Mapping
n	*CUST ID	
v	DEAR	
v	CUST NAME	
v	ADDRESS	
n	*CITY ID	
v	PHONE	
n	AGE	
v	AGE RANGE	
v	SALES PERS	
d	CRE DATE	
d	UPD DATE	

Transformação dos Dados



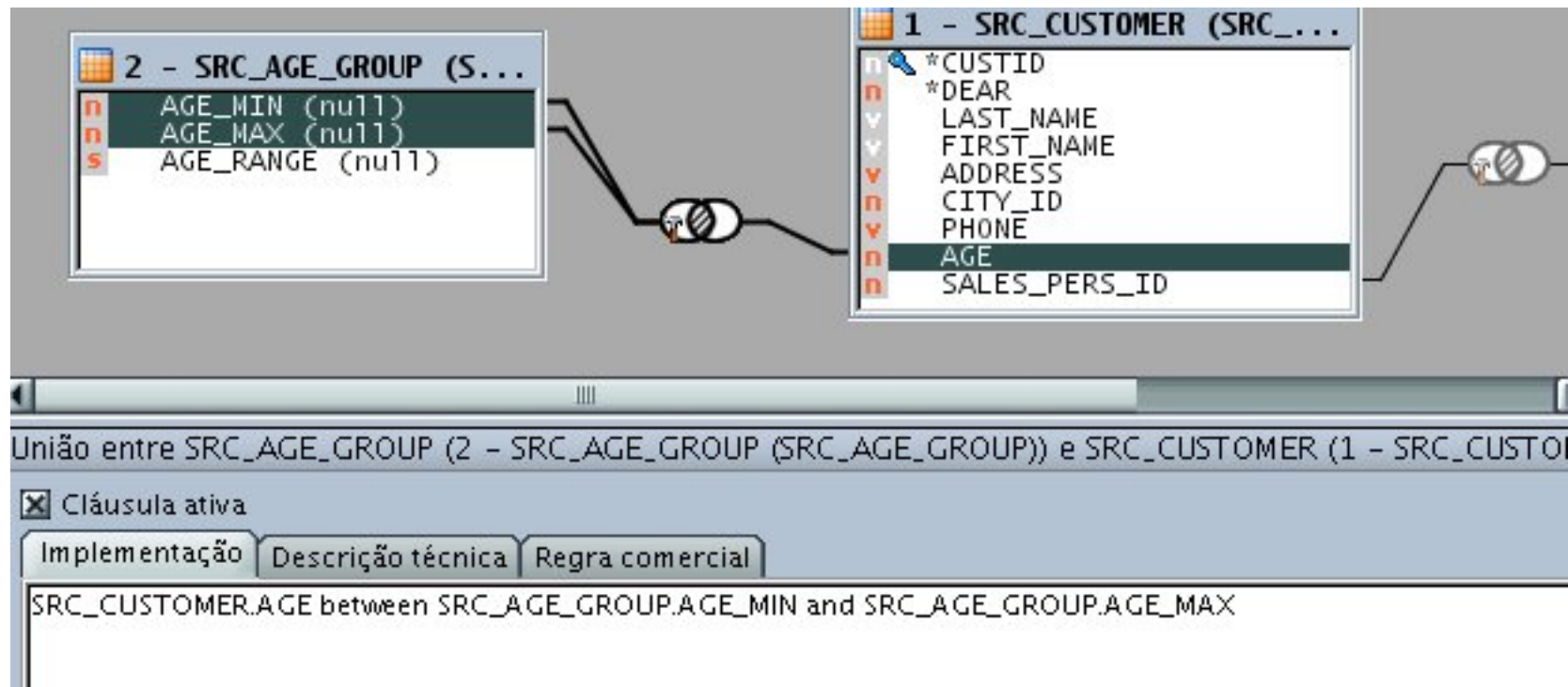
Transformação dos Dados

- Definindo os JOINS entre as fontes de dados



Transformação dos Dados

- Definindo os JOINS entre as fontes de dados



Transformação dos Dados

- Definindo as regras de transformação

Definir mapeamentos para os atributos em “Target datastore” que não foram mapeados automaticamente.

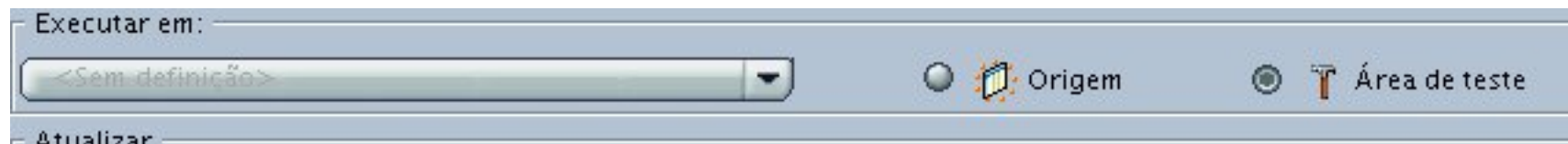
- **CUST_ID:** clicar em CUST_ID e arrastar SRC_CUSTOMER.CUSTID para o campo de texto da aba “implementação”.

- **DEAR:** compor a seguinte expressão:

- CASEWHEN (SRC_CUSTOMER.DEAR=0, 'MR',
CASEWHEN (SRC_CUSTOMER.DEAR=1, 'MRS',
'MS'))

Transformação dos Dados

- Definindo as regras de transformação
 - **CUST_NAME:** Compor a seguinte expressão:
 - SRC_CUSTOMER.FIRST_NAME || ' ' || UCASE(SRC_CUSTOMER.LAST_NAME)
 - **SALES_PERS:** Compor a seguinte expressão:
 - SRC_SALES_PERSON.FIRST_NAME || ' ' || UCASE(SRC_SALES_PERSON.LAST_NAME)



Transformação dos Dados

- Definindo as regras de transformação
 - **CRE_DATE**: Como queremos apenas que o mapeamento funcione apenas no Insert, desmarcar “Update”. Depois, compor a seguinte expressão:
 - `CURDATE ()`
 - **UPD_DATE**: Como queremos apenas que o mapeamento funcione apenas no Update, desmarcar “Insert”. Depois, compor a seguinte expressão:
 - `CURDATE ()`

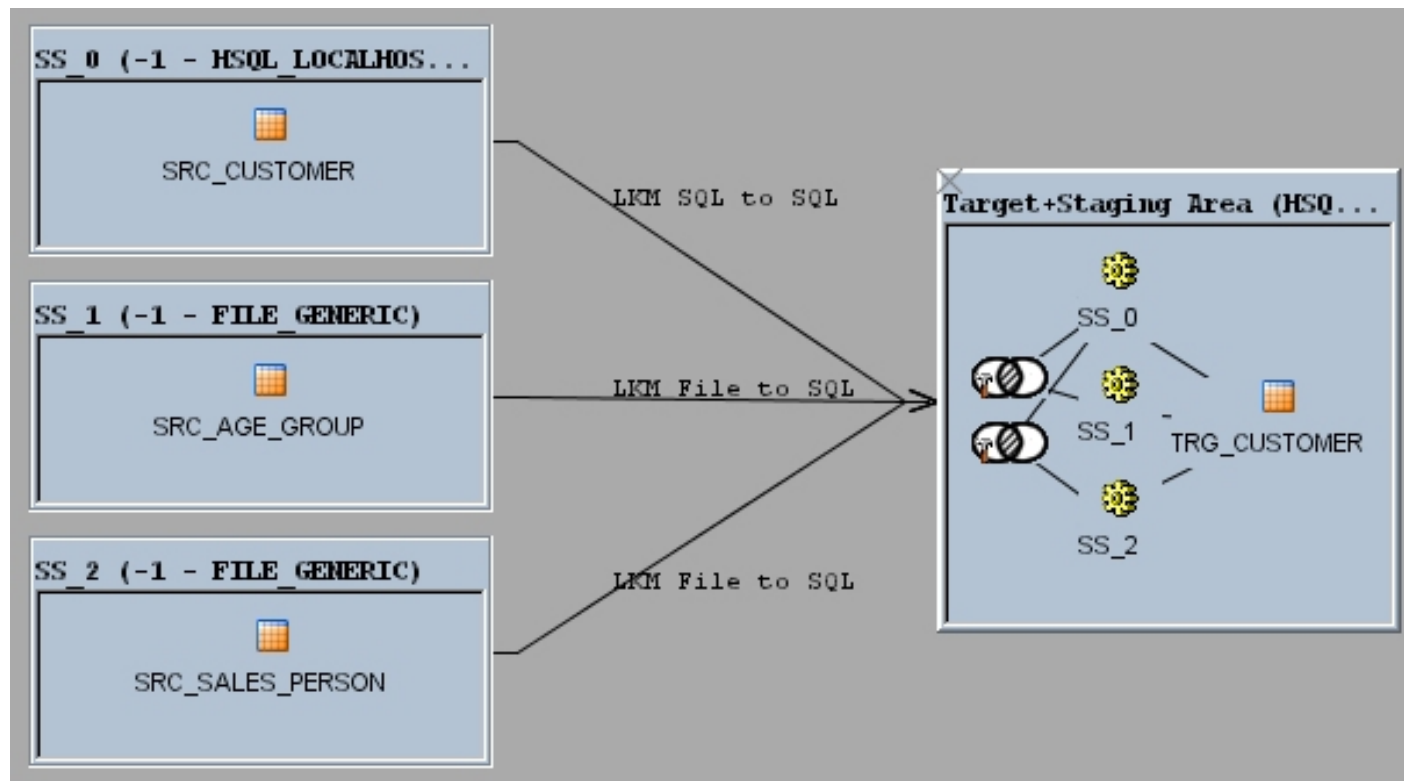
Transformação dos Dados

- Neste ponto, deveremos ter esta tabela de mapeamento.

TRG_CUSTOMER		
Ind	Name	Mapping
n	*CUST ID	SRC CUSTOMER.CUSTID
v	DEAR	CASEWHEN(SRC CUSTOMER.DEA...
v	CUST NAME	SRC CUSTOMER.FIRST NAME ...
v	ADDRESS	SRC CUSTOMER.ADDRESS
n	*CITY ID	SRC CUSTOMER.CITY ID
v	PHONE	SRC CUSTOMER.PHONE
n	AGE	SRC CUSTOMER.AGE
v	AGE RANGE	SRC AGE GROUP.AGE RANGE
v	SALES PERS	SRC SALES PERSON.FIRST NA...
d	CRE DATE	CURDATE()
d	UPD DATE	CURDATE()

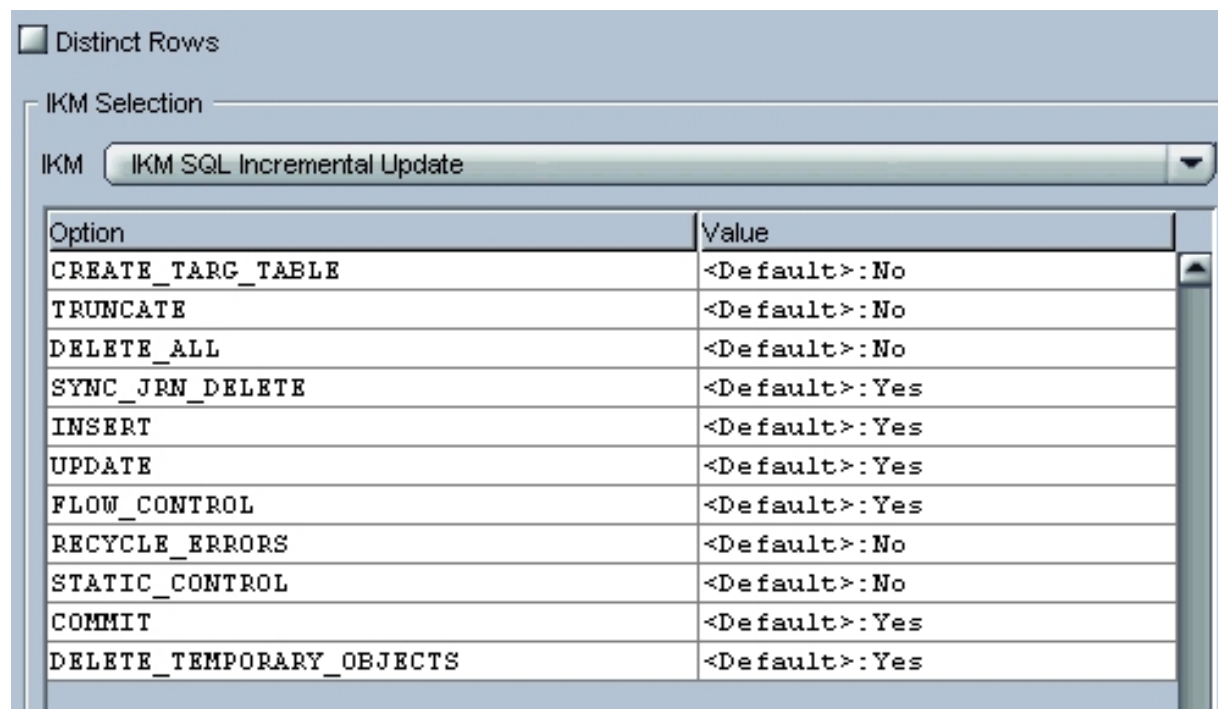
Transformação dos Dados

- Escolhendo estratégias de carregamento dos dados
 - Na aba “**Flow**” são visualizados os passos para realizar a transformação dos dados



Transformação dos Dados

- Escolhendo estratégias de integração dos dados
 - Clique na janela “Target + Staging Area” e escolha “IKM SQL Incremental Update”



Transformação dos Dados

- Escolhendo estratégias de controle dos dados
 - Vá para a aba “Control”

Maximum number of errors allowed %

Select your KM


CKM HSQL

Option	Value
DROP_ERROR_TABLE	<Default>:No
DROP_CHECK_TABLE	<Default>:No

Constraints

Constraint Name	Value
AGE >21	<Default>:Yes
PK_TRG_CUSTOMER	<Default>:Yes
FK_CUST_CITY	<Default>:Yes

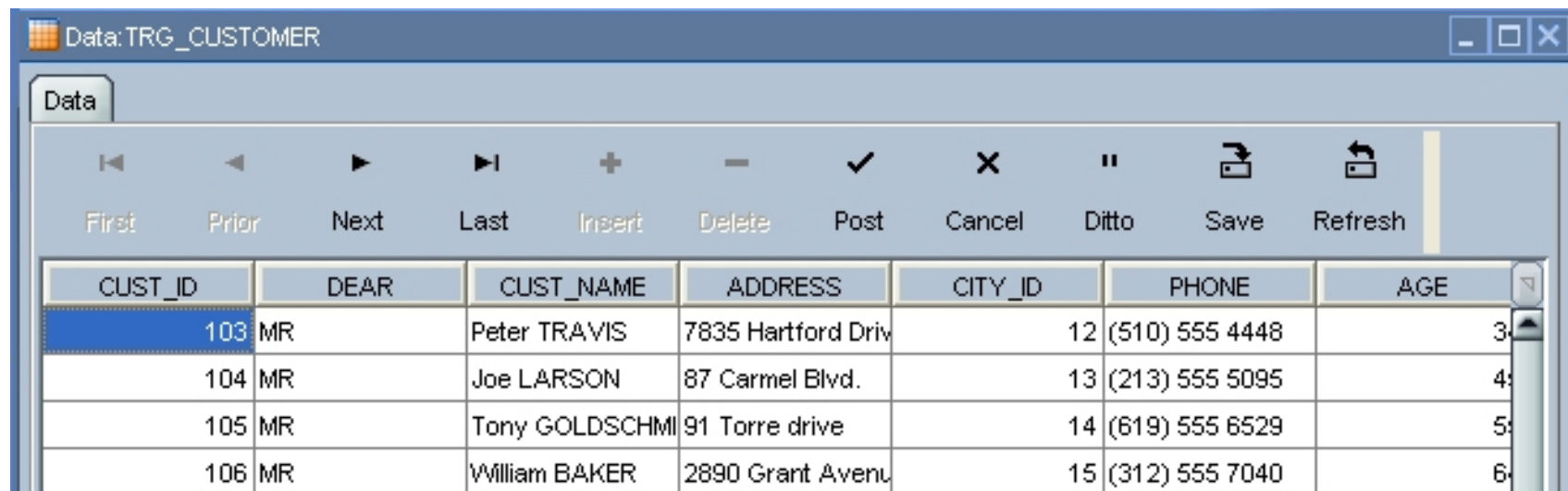
Transformação dos Dados

- Agora estamos prontos para executar nossa interface! Clique em “Execute”, salve a interface
- Abra a janela “Operator”  . Dê um clique duplo em “Pop. TRG_CUSTOMER”, vá para a aba “Execution” e verifique os números.

Nb of Inserts	25	Nb of Updates	0
Nb of Deletes	0	Nb of Errors	9

Transformação dos Dados

- Verificando o resultado
 - Verifique a tabela resultante. Na perspectiva “Models”, expanda “Sales Administrator – HQSL”, clique com o botão direito em TRG_CUSTOMER e em “Data”.

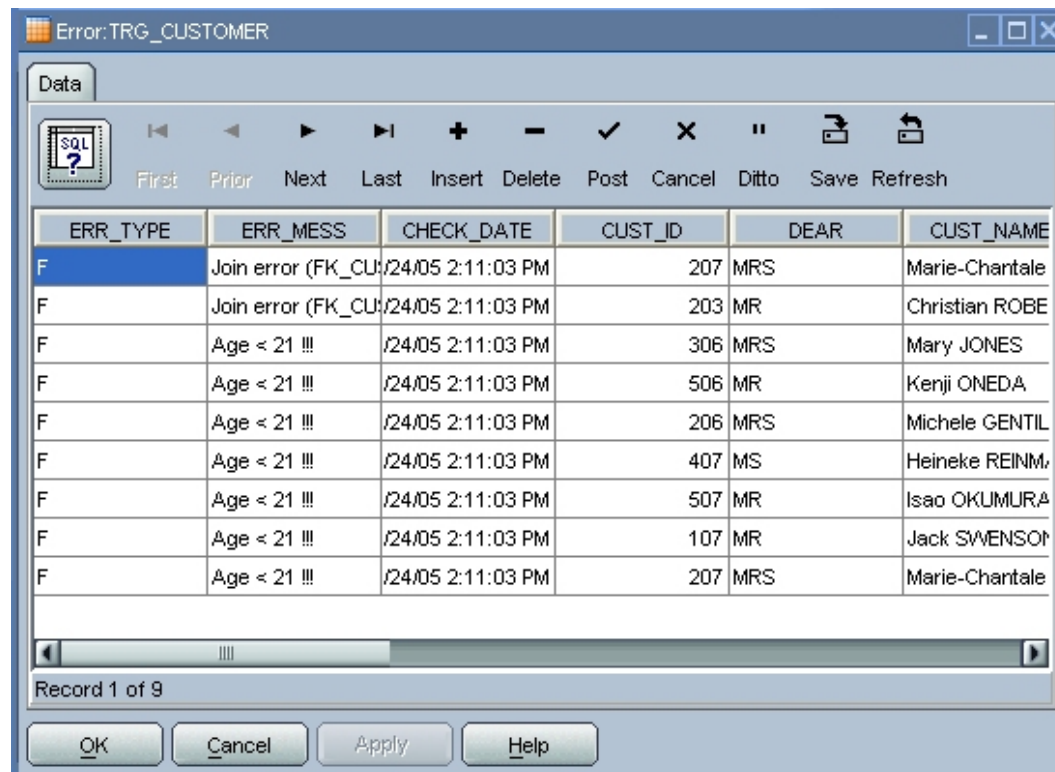


The screenshot shows a data grid window titled "Data:TRG_CUSTOMER". The window has a toolbar with the following icons and labels: First, Prior, Next, Last, Insert, Delete, Post, Cancel, Ditto, Save, and Refresh. The data table below has the following columns and rows:

CUST_ID	DEAR	CUST_NAME	ADDRESS	CITY_ID	PHONE	AGE
103	MR	Peter TRAVIS	7835 Hartford Drive	12	(510) 555 4448	3
104	MR	Joe LARSON	87 Carmel Blvd.	13	(213) 555 5095	4
105	MR	Tony GOLDSCHMIDT	91 Torre drive	14	(619) 555 6529	5
106	MR	William BAKER	2890 Grant Avenue	15	(312) 555 7040	6

Transformação dos Dados

- Verificando erros
 - Vimos que ocorreram 9 erros durante a execução da integração de dados. Para visualizá-los, clique com o botão direito em “TRG_CUSTOMER” > “Control” > “Errors”



The screenshot shows a window titled "Error: TRG_CUSTOMER" with a "Data" tab. The window contains a table with 6 columns: ERR_TYPE, ERR_MESS, CHECK_DATE, CUST_ID, DEAR, and CUST_NAME. The table lists 9 error records. The first record is highlighted in blue. Below the table is a scrollbar and a status bar indicating "Record 1 of 9". At the bottom of the window are buttons for "OK", "Cancel", "Apply", and "Help".

ERR_TYPE	ERR_MESS	CHECK_DATE	CUST_ID	DEAR	CUST_NAME
F	Join error (FK_CU	/24/05 2:11:03 PM	207	MRS	Marie-Chantale
F	Join error (FK_CU	/24/05 2:11:03 PM	203	MR	Christian ROBE
F	Age < 21 !!!	/24/05 2:11:03 PM	306	MRS	Mary JONES
F	Age < 21 !!!	/24/05 2:11:03 PM	506	MR	Kenji ONEDA
F	Age < 21 !!!	/24/05 2:11:03 PM	206	MRS	Michele GENTIL
F	Age < 21 !!!	/24/05 2:11:03 PM	407	MS	Heineke REINM,
F	Age < 21 !!!	/24/05 2:11:03 PM	507	MR	Isao OKUMURA
F	Age < 21 !!!	/24/05 2:11:03 PM	107	MR	Jack SWENSOP
F	Age < 21 !!!	/24/05 2:11:03 PM	207	MRS	Marie-Chantale

Transformação dos Dados

- Corrigindo erros: alterar os valores diretamente na fonte de dados

CUSTID	DEAR	LAST_NAME	FIRST_NAME	ADDRESS	CITY_ID	PHONE
101	0	Brendt	Paul	10 Jasper Blvd.	107	(212) 555 2146
102	0	McCarthy	Robin	27 Pasadena Drive	11	(214) 555 3075
103	0	Travis	Peter	7835 Hartford Drive	12	(510) 555 4448
104	0	Larson	Joe	87 Carmel Blvd.	13	(213) 555 5095
105	0	Goldschmidt	Tony	91 Torre drive	14	(619) 555 6529
106	0	Baker	William	2890 Grant Avenue	15	(312) 555 7040
107	0	Swenson	Jack	64 Imagination Drive	19	(202) 555 8125
201	0	Sartois	Jean	71 rue Rousseau	25	79 23 26 23
202	0	Michaud	Philippe	197 impasse Rendu	23	78 21 86 20
203	0	Robert	Christian	1rue Cezanne	107	42 25 27 29
204	1	Martin	Christine	12 allée Victor Hugo	24	25 26 46 26
205	0	Piaget	Luc	38 allée des Saules	29	53 42 24 28
206	1	Gentil	Michele	17montee des Chevaliers	25	65 62 26 13
207	1	Dupont	Marie-Chantale	37 rue Murat	20	46 72 23 53
301	1	Edwards	Caroline	68 Downing Street	35	243867945

Nb of Inserts: 1 Nb of Updates: 0
Nb of Deletes: 0 Nb of Errors: 8

- **Por exemplo:** o cliente com Id 203 está registrado com uma cidade não registrada (208), quando na verdade a cidade correta é a com Id 107. Basta alterar este valor e clicar em “Execute” na janela “Interface”.

Agenda

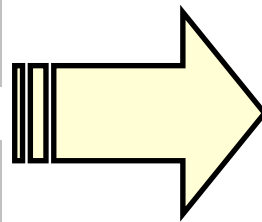
- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- **Interface de Integração dos Dados**
- Automatização
- Conclusão
- Referências

Interface para a Tabela TRG_SALES

- Definir junção entre
 - SRC_ORDERS
 - SRC_ORDER_LINES

SRC_ORDER_LINES			
<u>ORDER_ID</u>	NUMERIC(10)	<pk>	not null
<u>LORDER_ID</u>	NUMERIC(10)	<pk>	not null
PRODUCT_ID	NUMERIC(10)		null
QTY	NUMERIC(10)		null
AMOUNT	NUMERIC(10,2)		null

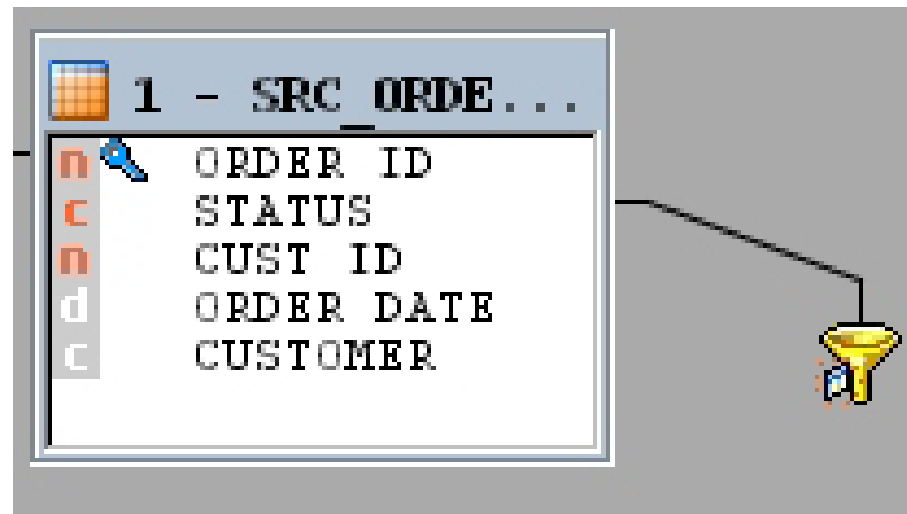
SRC_ORDERS			
<u>ORDER_ID</u>	NUMERIC(10)	<pk>	not null
STATUS	VARCHAR(3)		null
CUST_ID	NUMERIC(10)		null
ORDER_DATE	DATE		null
CUSTOMER	VARCHAR(35)		null



TRG_SALES			
<u>CUST_ID</u>	NUMERIC(10)	<pk, fk1>	not null
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk, fk2>	not null
FIRST_ORD_ID	NUMERIC(10)		not null
FIRST_ORD_DATE	DATE		not null
LAST_ORD_ID	NUMERIC(10)		not null
LAST_ORD_DATE	DATE		not null
QTY	NUMERIC(10)		not null
AMOUNT	NUMERIC(10)		not null
PROD_AVG_PRICE	NUMERIC(10)		not null

Interface para a Tabela TRG_SALES

- Regras de transformação de dados:
 - Somente pedidos com status = 'CLO' (closed)
 - Definir filtro no atributo status:
 - SRC_ORDERS.STATUS = 'CLO'



Interface para a Tabela TRG_SALES

- Regras para transformação de dados:
 - FIRST_ORD_ID: menor ORDER_ID
 - MIN(SRC_ORDERS.ORDER_ID)
 - FIRST_ORD_DATE: menor data de pedido
 - MIN(SRC_ORDERS.ORDER_DATE)
 - LAST_ORD_ID: maior ORDER_ID
 - MAX(SRC_ORDERS.ORDER_ID)
 - LAST_ORD_DATE: maior data de pedido
 - MAX(SRC_ORDERS.ORDER_DATE)

Interface para a Tabela TRG_SALES

- Regras para transformação de dados:
 - QTY: total da quantidade de produtos vendidos
 - SUM(SRC_ORDER_LINES.QTY)
 - AMOUNT: total de venda dos produtos
 - SUM(SRC_ORDER_LINES.AMOUNT)
 - PROD_AVG_PRICE: preço médio de venda dos produtos
 - AVG(SRC_ORDER_LINES.AMOUNT)

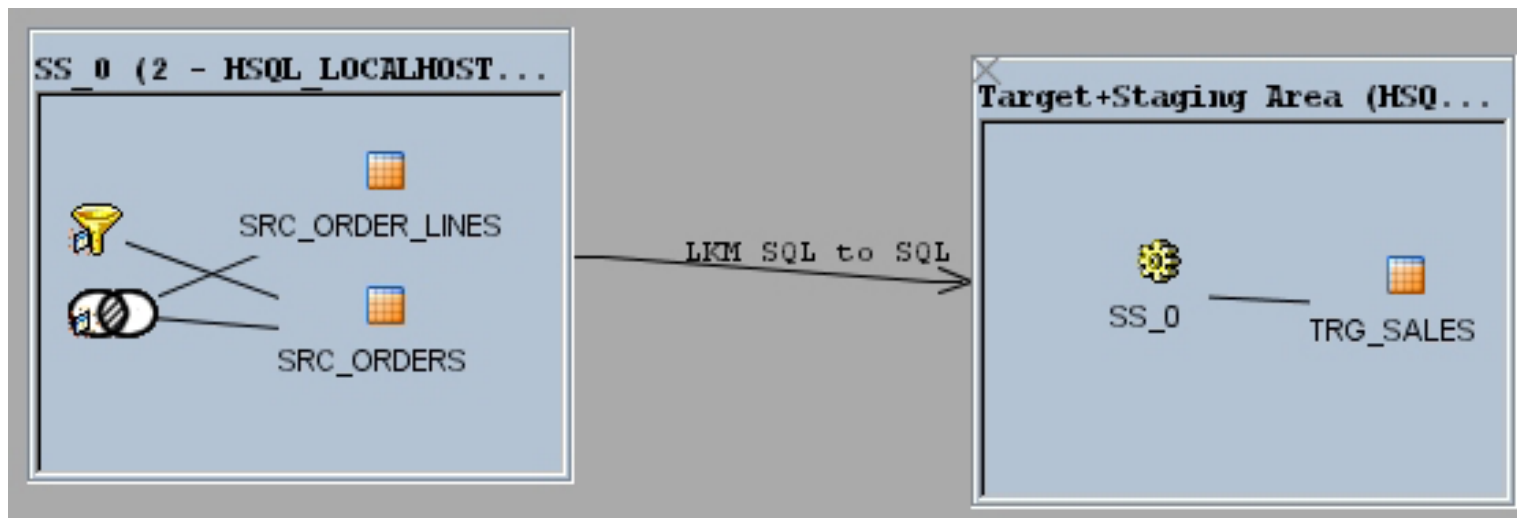
Interface para a Tabela TRG_SALES

- Regras para transformação de dados:

TRG_SALES		
Ind	Name	Mapping
n	*CUST ID	SRC ORDERS.CUST ID
n	*PRODUCT ID	SRC ORDER LINES.PRODUCT ID
n	*FIRST ORD ID	MIN(SRC ORDERS.ORDER ID)
d	*FIRST ORD...	MIN(SRC ORDERS.ORDER DATE)
n	*LAST ORD ID	MAX(SRC ORDERS.ORDER ID)
d	*LAST ORD ...	MAX(SRC ORDERS.ORDER DATE)
n	*QTY	SUM(SRC ORDER LINES.QTY)
n	*AMOUNT	SUM(SRC ORDER LINES.AMOUNT)
n	*PROD AVG ...	AVG(SRC ORDER LINES.AMOUNT)

Interface para a Tabela TRG_SALES

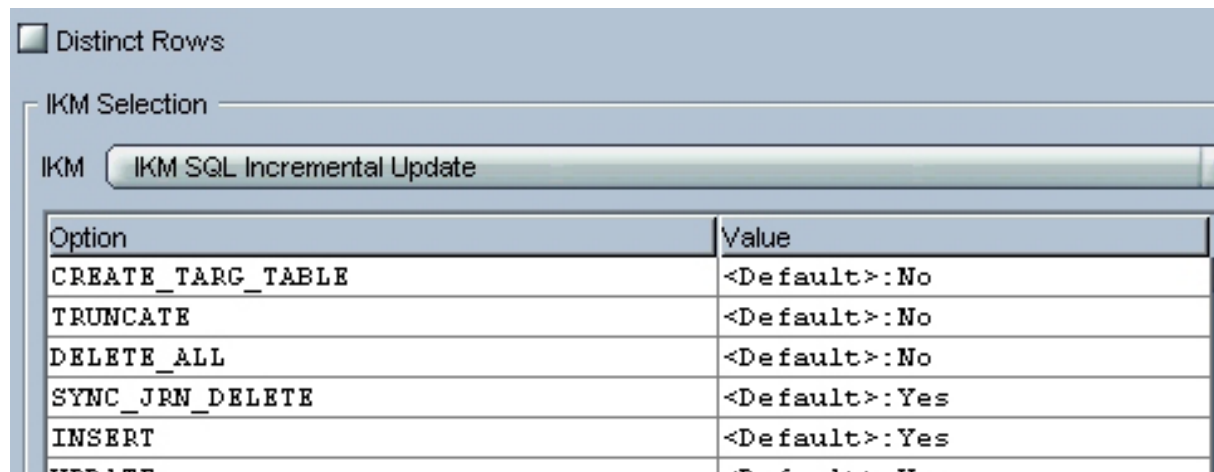
- Configurar Knowledge Module (KM) para exportar e carregar dados: wrapper
 - Aba Flow/Fluxo
 - Data Loading Strategy (LKM): LKM SQL to SQL



- Mais informações sobre LKM em [Oracle, 2009]

Interface para a Tabela TRG_SALES

- Configurar KIM para integração dos dados
 - Aba Flow/Fluxo
 - Data Integration Strategy (IKM):
 - IKM SQL Incremental Update



- Mais informações sobre LKM em [Oracle, 2009]

Interface para a Tabela TRG_SALES

- Configurar KM para verificação dos dados:
 - Aba Control/Controle
 - Data Control Strategy: CKM HSQL
 - Restrições
 - PK_TRG_SALES
 - FK_SALES_CUST (chave estrangeira SRC_CUSTOMER)
 - FK_SALES_PROD (chave estrangeira SRC_PRODUCT)
 - Mais informações sobre CKM em [Oracle, 2009]

Interface para a Tabela TRG_SALES

- Executar interface - ver resultado e erros

The screenshot shows a data grid interface for the table TRG_SALES. The table contains the following data:

CUST_ID	PRODUCT_ID	FIRST_ORD_ID	FIRST_ORD_DATE	LAST_ORD_ID	LAST_ORD_DATE	QTY
103	7	3	3/23/90	3	3/23/90	1
103	8	3	3/23/90	71	3/18/99	1
103	9	3	3/23/90	71	3/18/99	1
104	6	4	4/15/90	77	9/12/00	4
104	11	4	4/26/91	4	4/26/91	1
104	12	4	4/15/90	77	9/12/00	5
106	7	6	6/23/98	38	7/23/99	2
106	13	6	6/23/98	74	7/23/99	4
106	14	6	6/23/98	74	7/23/99	2
201	2	8	8/18/00	40	9/10/00	2
201	4	8	8/18/00	8	8/18/00	1
201	10	8	8/18/00	40	9/10/00	3
203	4	10	5/20/99	10	5/20/99	1
203	7	10	5/20/99	10	5/20/99	1
203	9	10	5/20/99	10	5/20/99	1

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- **Automatização**
- Conclusão
- Referências

Automatização da Carga dos Dados no DW

- Definir o fluxo correto para a carga de dados no DW
 - Usar interfaces previamente definidas
 - Atenção para a ordem de carga:
 - Dependências de chaves estrangeiras

TRG_COUNTRY			
<u>COUNTRY_ID</u>	NUMERIC(10)	<pk>	not null
COUNTRY	VARCHAR(50)		null

1

FK_REGION_COUNTRY

TRG_REGION			
<u>REGION_ID</u>	NUMERIC(10)	<pk>	not null
COUNTRY_ID	NUMERIC(10)	<fk>	not null
REGION	VARCHAR(50)		null

2

FK_CITY_REGION

TRG_CITY			
<u>CITY_ID</u>	NUMERIC(10)	<pk>	not null
REGION_ID	NUMERIC(10)	<fk>	not null
CITY	VARCHAR(50)		null
POPULATION	NUMERIC(10)		null

3

FK_CUST_CITY

TRG_CUSTOMER			
<u>CUST_ID</u>	NUMERIC(10)	<pk>	not null
DEAR	VARCHAR(50)		null
CUST_NAME	VARCHAR(50)		null
ADDRESS	VARCHAR(50)		null
CITY_ID	NUMERIC(10)	<fk>	not null
PHONE	VARCHAR(50)		null
AGE	NUMERIC(10)		null
AGE_RANGE	VARCHAR(50)		null
SALES_PERS	VARCHAR(50)		null
CRE_DATE	DATE		null
UPD_DATE	DATE		null

6

FK_SALES_CUST

TRG_PROD_FAMILY			
<u>FAMILY_ID</u>	VARCHAR(50)	<pk>	not null
FAMILY_NAME	VARCHAR(50)		null

4

FK_PROD_PROD_FAMILY

TRG_PRODUCT			
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk>	not null
FAMILY_ID	VARCHAR(50)	<fk>	not null
PRICE	NUMERIC(10)		null
PRODUCT	VARCHAR(50)		null

5

FK_SALES_PROD

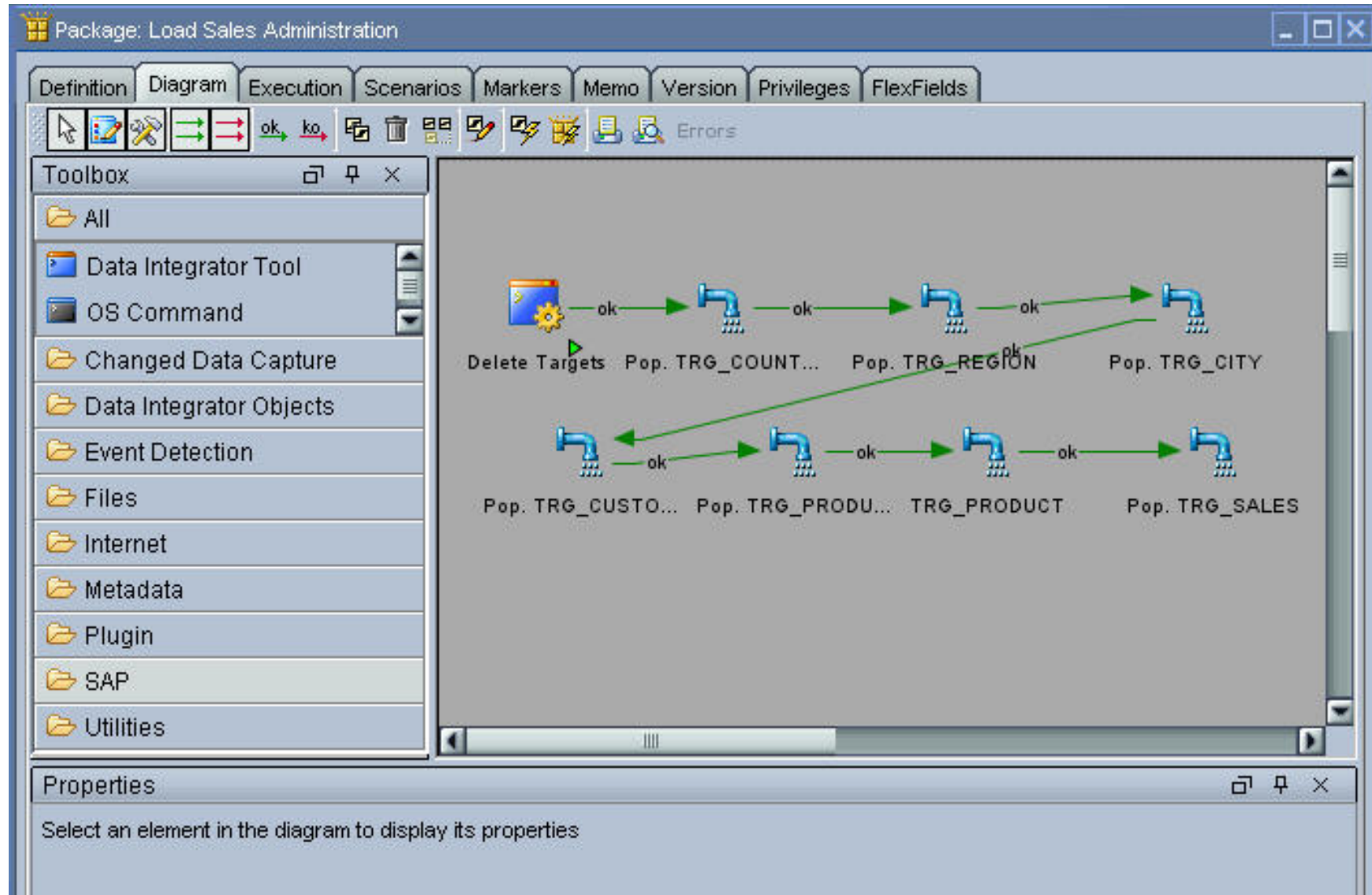
TRG_SALES			
<u>CUST_ID</u>	NUMERIC(10)	<pk,fk1>	not null
<u>PRODUCT_ID</u>	NUMERIC(10)	<pk,fk2>	not null
FIRST_ORD_ID	NUMERIC(10)		not null
FIRST_ORD_DATE	DATE		not null
LAST_ORD_ID	NUMERIC(10)		not null
LAST_ORD_DATE	DATE		not null
QTY	NUMERIC(10)		not null
AMOUNT	NUMERIC(10)		not null
PROD_AVG_PRICE	NUMERIC(10)		not null

7

Automatização da Carga dos Dados no DW

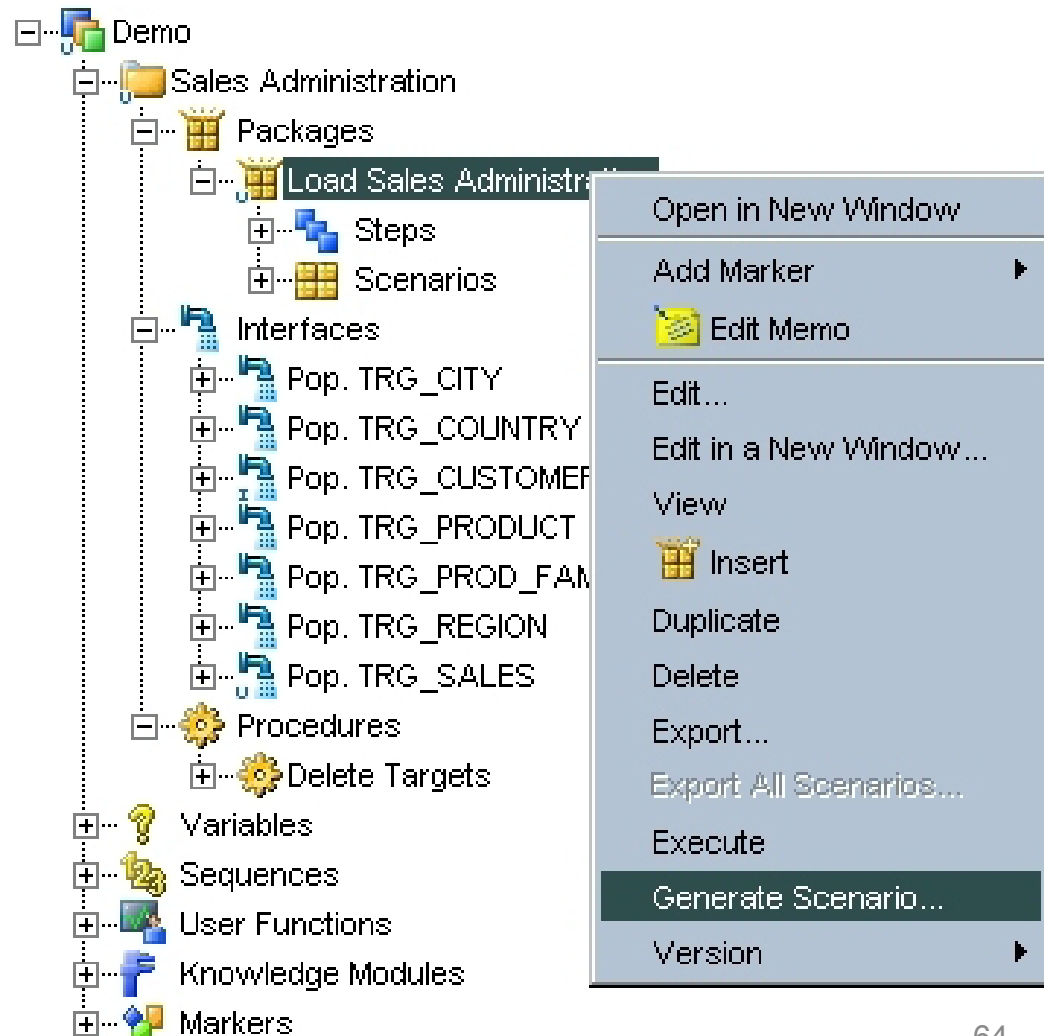
- Criar package “Load Sales Administration”:
 - Procedimento “Delete Targets”
 - “Pop. TRG_COUNTRY” interface
 - “Pop. TRG_REGION” interface
 - “Pop. TRG_CITY” interface
 - “Pop. TRG_PROD_FAMILY” interface
 - “Pop. TRG_PRODUCT” interface
 - “Pop. TRG_CUSTOMER” interface
 - “Pop. TRG_SALES” interface

Automatização da Carga dos Dados no



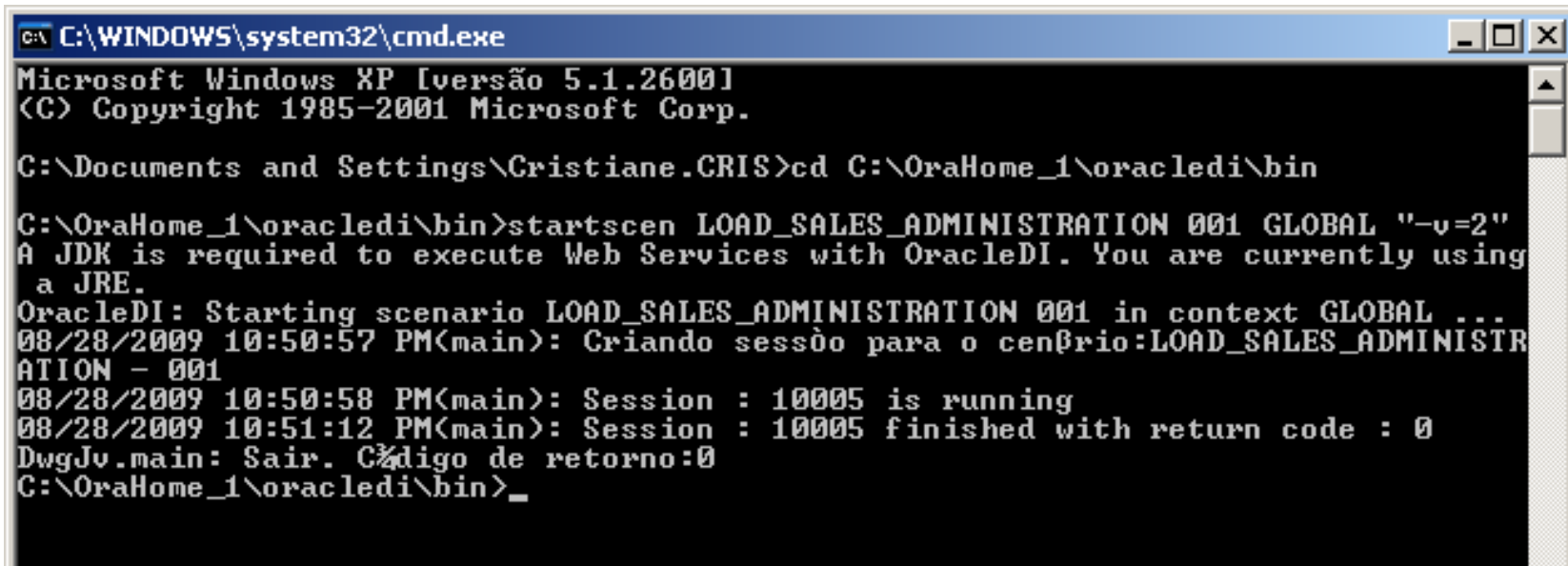
Automatização da Carga dos Dados no DW

- Executar a carga por meio de comando do Sistema Operacional
 - Criar cenário



Automatização da Carga dos Dados no DW

- Executar cenário do Sistema Operacional
 - Abrir prompt e ir para a pasta do Oracle Data Integrator, diretório bin
 - Digitar o comando:
 - startscen LOAD_SALES_ADMINISTRATION 001 GLOBAL "-v=2"



```
C:\WINDOWS\system32\cmd.exe
Microsoft Windows XP [versão 5.1.2600]
(C) Copyright 1985-2001 Microsoft Corp.

C:\Documents and Settings\Cristiane.CRIS>cd C:\OraHome_1\oracledi\bin

C:\OraHome_1\oracledi\bin>startscen LOAD_SALES_ADMINISTRATION 001 GLOBAL "-v=2"
A JDK is required to execute Web Services with OracleDI. You are currently using
a JRE.
OracleDI: Starting scenario LOAD_SALES_ADMINISTRATION 001 in context GLOBAL ...
08/28/2009 10:50:57 PM(main): Criando sessão para o cenário:LOAD_SALES_ADMINISTR
ATION - 001
08/28/2009 10:50:58 PM(main): Session : 10005 is running
08/28/2009 10:51:12 PM(main): Session : 10005 finished with return code : 0
DwgJv.main: Sair. Código de retorno:0
C:\OraHome_1\oracledi\bin>
```

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- **Conclusão**
- Referências

Conclusão

- Processo ETL no Oracle Data Integrator
 - Limpeza dos dados
 - Constraints
 - Extração e Transformação dos dados
 - Interfaces
 - Automatização da carga de dados
 - Packages e cenários

Agenda

- Oracle
- Oracle Data Integrator (ODI)
- Estudo de Caso
- Inicializando o ODI
- Limpeza dos Dados
- Transformação dos Dados
- Interface de Integração dos Dados
- Automatização
- Conclusão
- **Referências**

Referências

- [Oracle, 2008]: Oracle Data Integrator: Getting Started with an ETL Project. Disponível em <http://download.oracle.com/docs/cd/E15985_01/doc.10136/getstart/GSETL.pdf>. Acesso em 21 set. 2010.
- [Oracle, 2009]: Oracle Data Integrator User's Guide, 10g Release 3 (10.1.3). Disponível em <http://download.oracle.com/docs/cd/E15985_01/doc.10136/user/ODICU.pdf>. Acesso em 21 set. 2010.