IBM Optim Solutions with Designer Proof of Technology

Lab Exercises





Contents

Lab 1	DISCOVERING OPTIM ARCHIVE CAPABILITIES	1
Lab 2	DISCOVERING OPTIM DELETE AND RESTORE PROCESSING	25
Lab 3	ADDING OPTIM RELATIONSHIPS	35
Lab 4	DISCOVERING OPTIM TEST DATA MANAGEMENT CAPABILITIES	47
Lab 5	DISCOVERING OPTIM DATA PRIVACY CAPABILTIES	79
LAR 6	INFOSPHERE DISCOVERY FINDING RELATIONSHIPS AND COMPLETE BUSINESS OBJECTS	94

VMWare Image Userid & Password

User = Administrator Password = passw0rd

User = ISD_ASSETS Password = passw0rd

Lab 1 Discovering Optim Archive Capabilities

In this lab we will create an Optim archive from a set of tables that contain information about sales that have occurred within a company. We will be working with four tables in this lab: Orders, Details, Customer, and Items. These tables contain the following information for our archive scenario:

ISD ASSETS.OPTIM ORDERS – Contains information about orders placed by customer

ISD ASSETS.OPTIM DETAILS - Contains the order details associated with a customer order

ISD_ASSETS.OPTIM_CUSTOMERS - Contains customer information

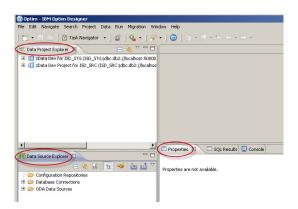
ISD ASSETS.OPTIM ITEMS - Contains information about items sold

These items have relationships defined within the database. In another lab we will show how you can define new relationships that are not defined in the database.

- __1. Open Optim Designer. This section of the lab will show you how to start Optim Designer and some basics about it.
 - __a. Double click on the **Optim Designer** on your desktop. This may take a minute or so to initially start, so please be patient.



b. View the Optim Designer Screen that displays



The upper left portion of the screen is called the **Data Project Explorer**. This contains all the Optim objects that you've created to perform various Optim tasks, such as an archive. As we create our archive process, you will see the items that you create within the folders in the Data Project Explorer.

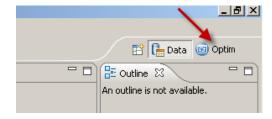
The bottom left portion of the screen is called the **Data Source Explorer**. This contains the database connections that are defined that will be used during our lab.

The top right portion of the screen is the editor screen and will vary in its content based on the type of item that you have opened.

The bottom right of the screen is the **Properties View**. It will contain all of the properties associated with a selected object.

_c. Be sure that you are in the Optim Perspective by clicking on the **Optim** icon in the upper

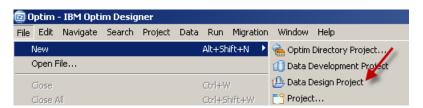
right part of your screen.



__2. Create a new Data Design Project.

We will now create a new Data Design Project. This project will contain all of the objects we create that are used to define our data archiving activities.

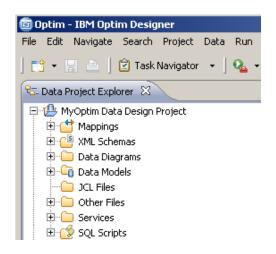
_a. Choose File → New → Data Design Project from the top task bar.



_b. You can give the project any name you'd like. For the purposes of this PoT call it "MyOptim Data Design Project". Click Finish.

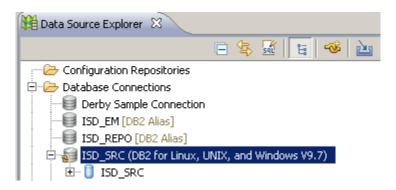


__c. You will now see the project that you just created listed in the Data Project Explorer part of your screen.

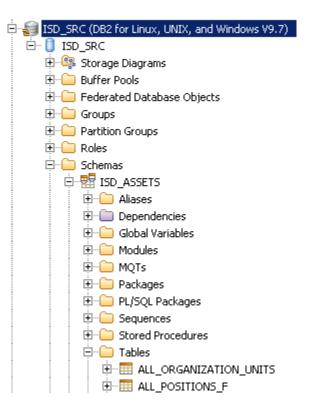


Page 2 Optim PoT Labs

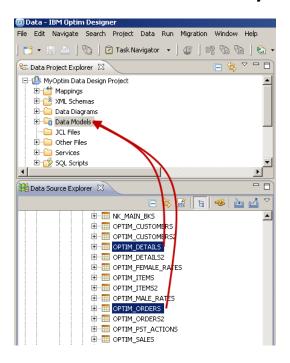
- 3. Create the Physical Data Model.
 - __a. The tables for our archive process reside in the database called ISD_SRC. This can be seen by expanding the **Database Connections** folder found in the **Data Source Explorer** at the bottom left of your screen. Double left click on the **ISD_SRC** icon to connect to the database.



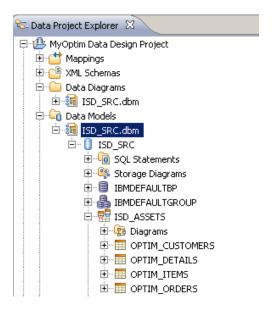
_b. Once connected, expand down to the **Schemas** folder to see the schemas of the tables within this database. The **ISD_ASSETS** schema contains the tables we will use for our archive process.



_c. We will now create our Physical Data Model. There are a few ways to do this. In this lab we will use the drag and drop method. Drill down into the ISD_ASSETS schema, press CTRL and select the OPTIM_ORDERS and OPTIM_DETAILS tables and drag and drop them onto the Data Models folder within the Data Project Explorer above.



__d. After completing this drag/drop step, you will see a model called ISD_SRC.dbm under the Data Models folder. You will see two additional tables, OPTIM_CUSTOMERS and OPTIM_ITEMS were added by Optim. These additional tables are the result of the RI that is defined in the database. (NOTE: In a subsequent lab you will see how Optim can reflect relationships, even if they are not defined in the database).

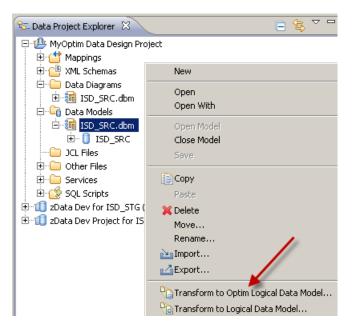


Page 4 Optim PoT Labs

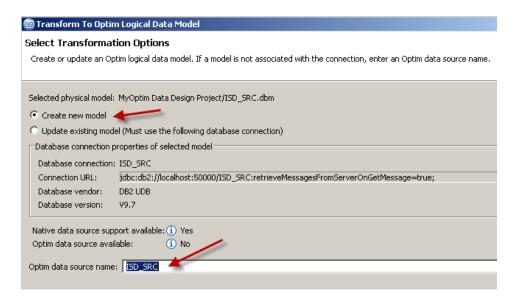
4. Create the Optim Logical Data Model from the Physical Data Model.

You now need to create what is called the Optim logical data model.

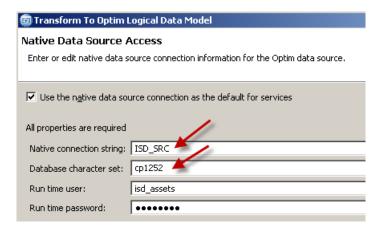
_a. Right click on the ISD_SRC.dbm model and choose Transform to Optim Logical Data Model.



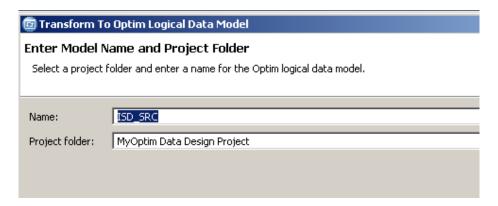
__b. When presented with the transformation options screen, be sure that **Create New Model** is selected and that your database name is **ISD_SRC**. Once you have verified the proper selections, click **Next**.



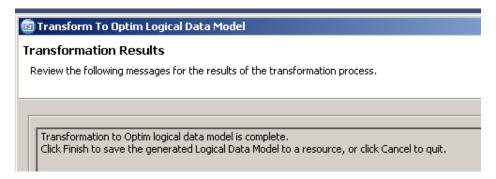
c. Enter the following options on the Native Data Source Access screen. Click Next.



d. Be sure that the following are selected for the Name and Project folder. Click **Next.**

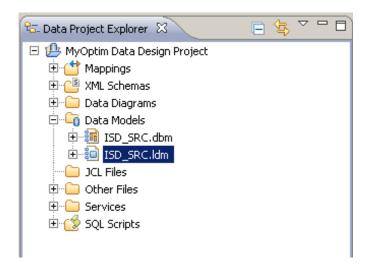


__e. The following screen should appear if everything has occurred successfully. Click **Finish**.

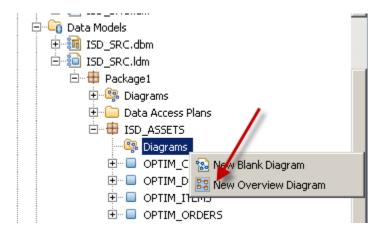


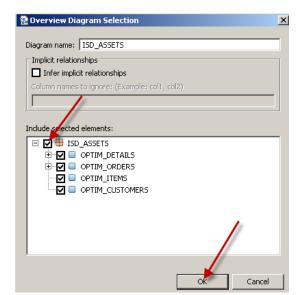
Page 6 Optim PoT Labs

__f. You should now see two models under your **Data Models** folder. We will be working in the **ISD_SRC.Idm** model in our next Step.



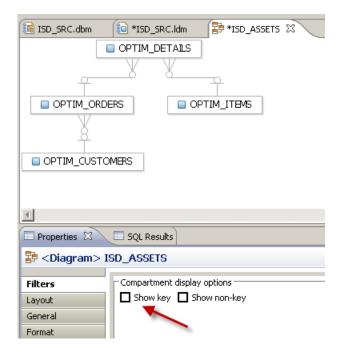
Drill into the ISD_SRD.Idm model by double-clicking on it, and then expand Package1 to see the Diagrams folder, Data Access Plans folder, and the ISD_ASSETS sub-package. Next expand the ISD_ASSETS package, and create an Overview Diagram by right clicking on Diagrams and selecting New Overview Diagram. Note, be sure to select the 2nd Diagrams icon.





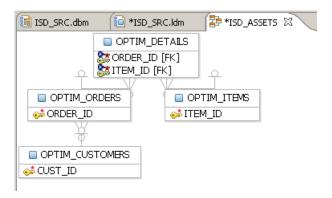
Be sure to **check all tables**. This will enable you to see the tables in a diagram. Click **OK**.

__g. The resulting diagram will only display entity names. This is because the Show key property is not chosen. That is selected in the **Properties** tab, in the bottom right part of your screen. Select the **Show key** check box to view a diagram that includes key attributes.



Page 8 Optim PoT Labs

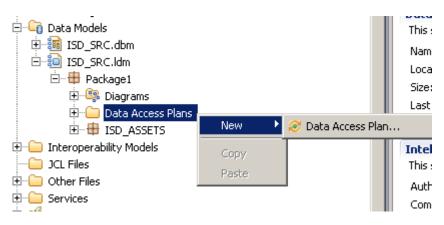
__h. The resulting diagram now includes the keys.



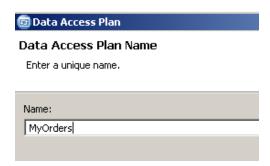
__5. Create a Data Access Plan (DAP).

A Data Access Plan allows you to select specific data you would like to archive. Within the Optim Logical Model that you just created you will find a folder for Data Access Plans.

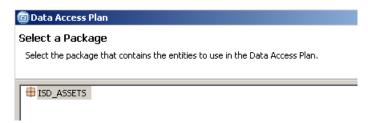
__a. Right click on the **Data Access Plans** folder and select **New** and **Data Access Plan**.



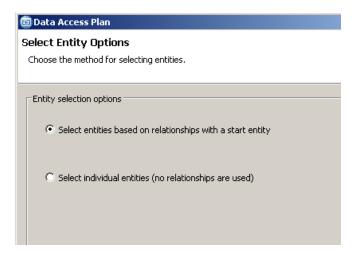
_b. Call the DAP "MyOrders" and click Next.



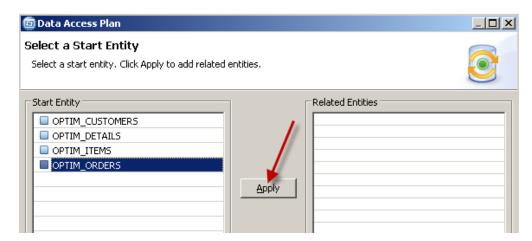
c. ISD_ASSETS should be your only package choice. Click Next.



_d. Choose the option to **select entities based on a start entity**. This allows Optim to pull data related to the chosen start entity. Click **Next**.

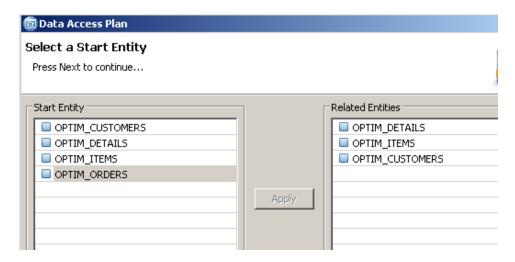


_e. Select OPTIM_ORDERS as the start table and push the Apply button. This will automatically pick up the entities related to OPTIM_ORDERS.

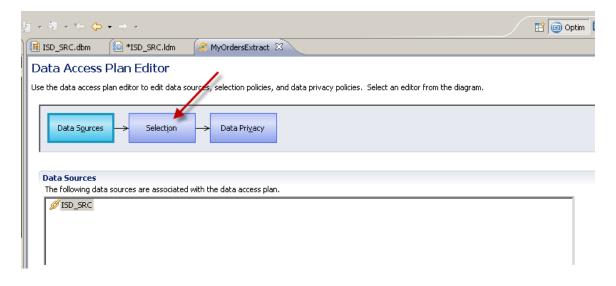


Page 10 Optim PoT Labs

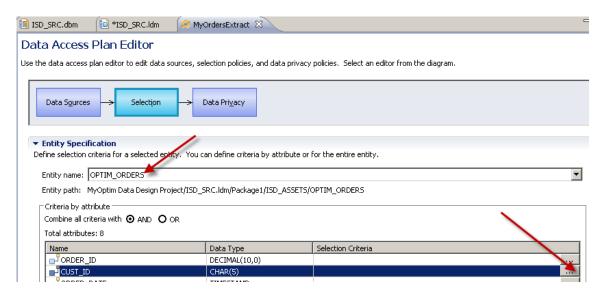
f. The next screen shows the related entities. Click **Finish**.



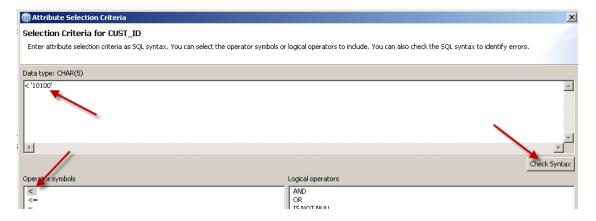
__g. You will now see the following screen. Select the **Selection** option from the Editor. This will allow you to subset the data that you want to archive.



__h. Scroll down until you see the **Entity Specification** section of the screen, choose the **OPTIM_ORDERS** entity and the **CUST_ID** attribute. Click on the ellipsis ("...") to the far right of the **CUST_ID** attribute to perform your selection criteria.



_i. Double click on the "<" symbol and type in '10100' with single quotes as shown below. We will be archiving all orders (and related data), for those CUST_IDs less than 10100. Click the **Check Syntax** button to be sure that your syntax is correct.



j. Click **OK** to add the selection criteria. From the keyboard, press "**CTRL-S**" to save it.

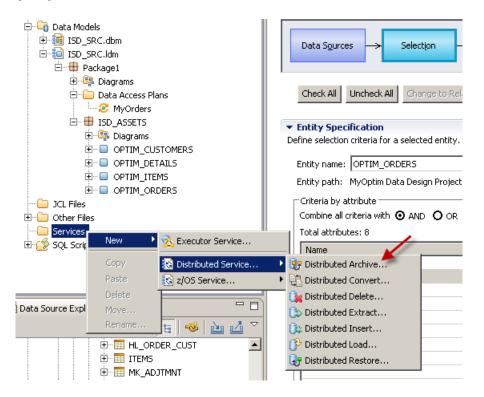
We have now created a logical Optim model and a Data Access Plan. We are now ready to create the Archive request. Note that you can have many different Data Access Plans associated with your Optim Logical Model, each one defining a different set of data.

Page 12 Optim PoT Labs

__6. Create Optim Service for archiving.

Now that you have created the DAP for the archive, you now need to create the service that will actually run on the Optim server. Optim comes with a component called the Optim Manager. Administrators can use this component to monitor and run services created from Optim Designer. Optim Manager can be accessed from Optim Designer or from a browser. We will use the browser interface in this lab. We'll first create the service from Optim Designer and then will run it from Optim Manager.

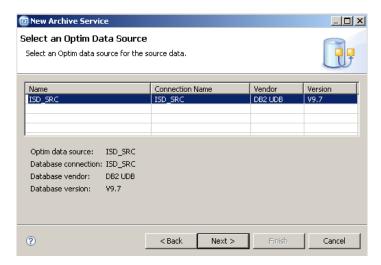
_a. From the **Data Project Explorer**, drill into the **MyOptim Data Design Project**, right click on the **Services** folder and choose **New** → **Distributed Services** → **Distributed Archive**.



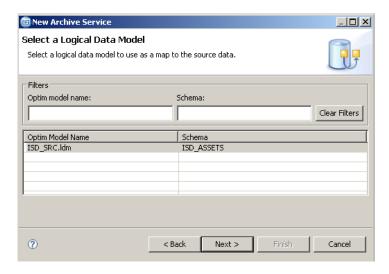
b. Give it a name of "MyArchive" and click Next.



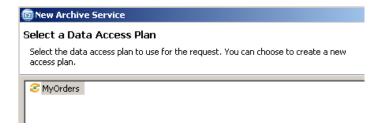
__c. Take the default for Select the Optim Data Source. Click **Next**.



__d. Take the default for Select a Logical Data Model. Click Next.

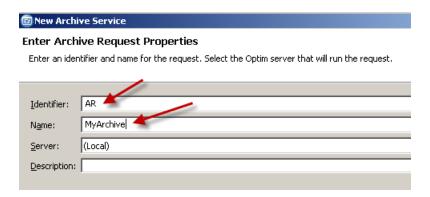


__e. Take the default of **MyOrders** for the Data Access Plan. (Remember, this DAP selects all orders and related tables for those CUST_IDs less than 10100). Click **Next.**



Page 14 Optim PoT Labs

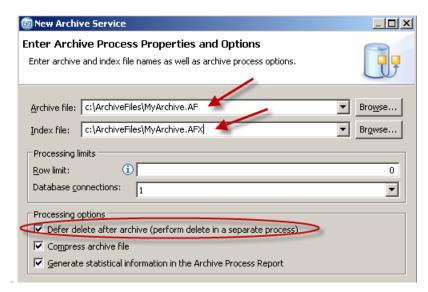
__f. Enter "AR" for the Identifier and "MyArchive" for the Name in the Archive Request Properties. These names are for the Optim Server that will execute the Archive process. Click Next.



_g. Enter "ISD_SRC" as the DB alias, this is the database and the ISD_ASSETS schema for the Access Definition Options. Click **Next**.

New Archive Service			
Enter Access Definition Properties			
Enter default DB alias and creator ID qualifiers for table names. The request will use a local access definition.			
Access definition:	(Local)		
<u>D</u> B alias:	ISD_SRC		
<u>C</u> reator ID:	ISD_ASSETS		

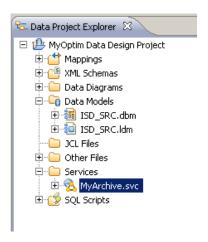
_h. Select the following for the Archive Process Properties. We are calling the archive file "MyArchive.AF" and the archive index file "MyArchive.AFX". You can use the Browse button to locate the directory and then type the file names. We are deferring our deletes to a separate delete process which will be defined in a separate lab. Click Next.



__i. Take the defaults on the Select Objects to Archive screen. Click **Finish**.



__j. You will now see the **MyArchive.svc** in your project.

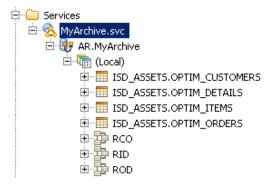


7. Select Tables for data deletion

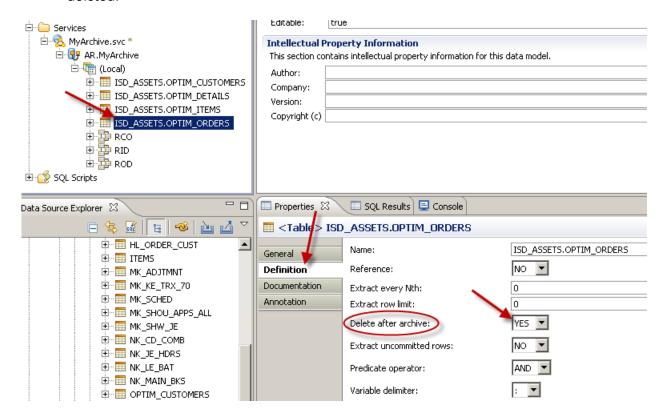
Just because you are archiving a piece of data doesn't mean you necessarily want to delete all the data from your source database. In our example, we are archiving all orders and related data for customer ID's less than 10100. We only want to delete the Orders and Details information, but not the Customers or Items information. We will go into our Archive Request service that we just created and specify which tables should have data deleted when we create our delete processes in our next lab.

Page 16 Optim PoT Labs

__a. Double-click and expand on MyArchive.svc service that you just created so that you can see the four tables involved in our archive process. We will be setting the parameter for the "OPTIM_ORDERS" and "OPTIM_DETAILS" tables to allow archive data to be deleted in our delete process (this will defined in a subsequent lab)



__b. Highlight the OPTIM_ORDERS table and look at its properties in the lower right part of your screen. Within the properties table/view select **Definition** and change the setting for **Delete after archive** to **Yes**. This will allow for the orders data that is archived to be deleted.



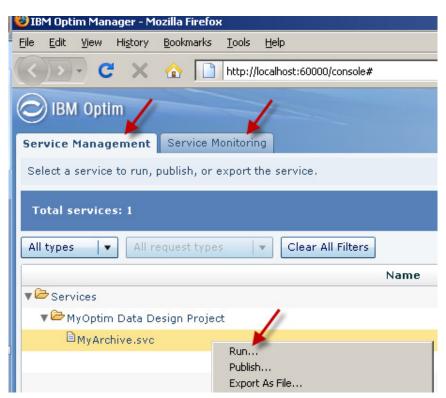
- __c. Complete the same steps for the **OPTIM_DETAILS** table. Note that we will not delete data from the OPTIM_ITEMS and OPTIM_CUSTOMERS tables. We want to archive the customer and item information about our archived orders, but we do not want to actually delete customer and item data from our source database.
- __d. Save your MyArchive.svc model via the **save icon** or via the "CTRL-S" keys.

8. Run the Archive Service

__a. We will now open Optim Manager from a web browser so that we can monitor and run the newly created service. Open **Mozilla Firefox** and you will open Optim Manager.

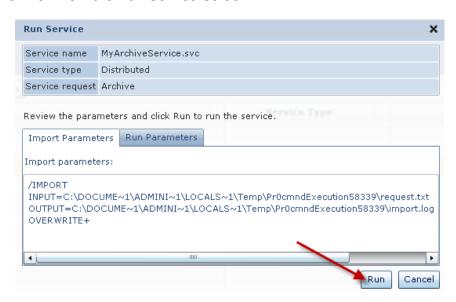


- __b. You will be presented with a screen with two tabs, one for service management and the other for service monitoring. You will see your **MyArchive.svc** listed under the **Service Management** tab.
- __c. Right click on the **MyArchive.svc** service and select **Run**.



Page 18 Optim PoT Labs

d. Click **Run** from the Run Service Screen.



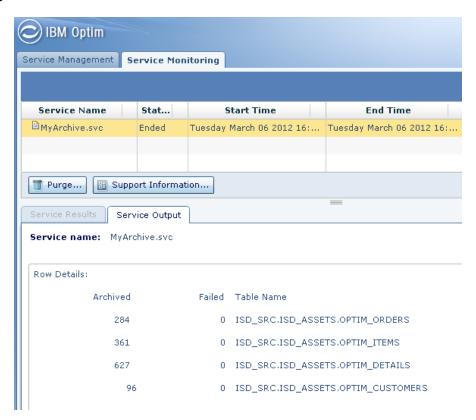
__e. You should receive the following confirmation box, stating that your job has been submitted. Click Close.



__f. Select the **Service Monitoring** Tab and you will eventually see that your service completed successfully. Use the **refresh** button (two green arrows on the right side of the screen) to refresh the screen.

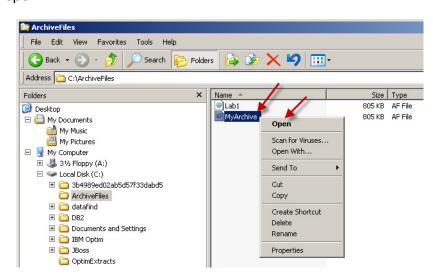


__g. Highlight the **MyArchive** service and select the **Service Output** tab at the bottom of the screen. **Scroll down** and you will see the number of rows successfully archived into your archive file.



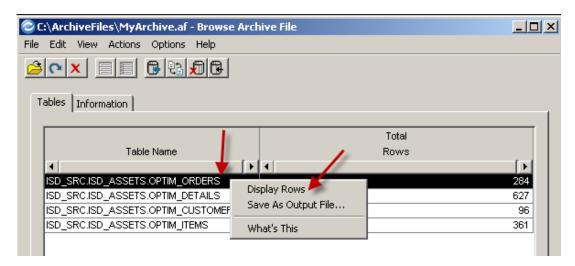
- 9. Browse the archive file
 - __a. Open **Windows Explorer** and navigate to the "c:\ArchiveFiles" directory. Double-click on the **MyArchive** file or right-click on the file and select **Open**. It may take a few seconds to open.



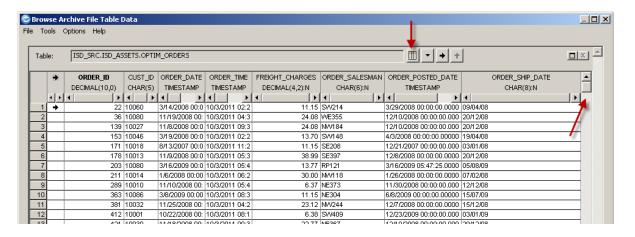


Page 20 Optim PoT Labs

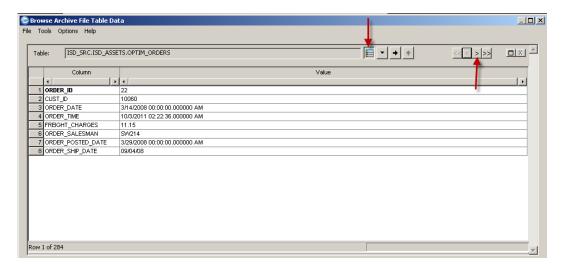
__b. The Optim file browser will display the tables that are in the file. Double-click the OPTIM_ORDERS table or right-click and select Display Rows.



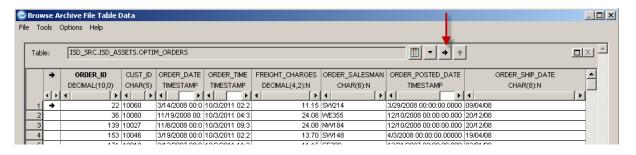
__c. Review the data using the scroll bar on the right, notice all CUST_ID's are less than 10100. Toggle the view from a column display to a side label display by **clicking on the icon**.



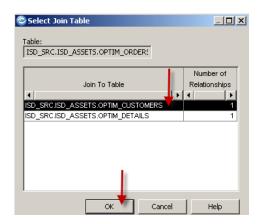
__d. Each row is displayed individually. Click on the ">" to view the next row. Click the icon to toggle back to the column display.



__e. The Optim file browser maintains relationships between tables. Click on the **right arrow** icon to join additional tables.



__f. Select the **OPTIM_CUSTOMERS** table. Click **OK**.

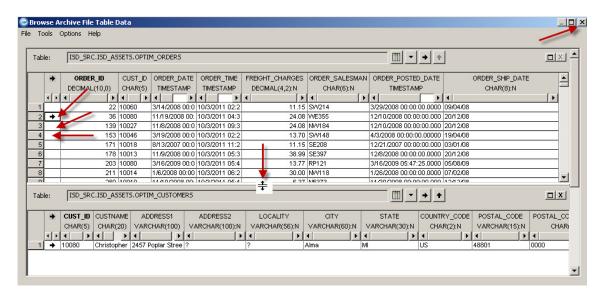


Page 22 Optim PoT Labs

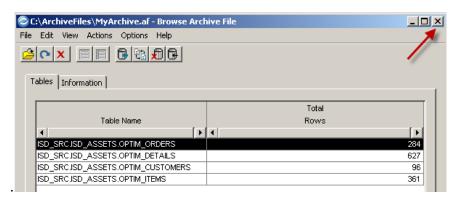
__g. Resize the window to see both tables easier. Move the cursor to the top of the OPTIM_CUSTOMER table until the cursor changes to the following shape. Left-click and hold as you move the lower window down and then release.

Click on each row in the OPTIM_ORDERS table to see the corresponding related data in the OPTIM CUSTOMER table.

Close the window.

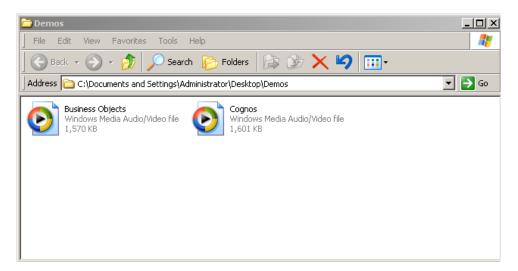


__h. Close the last window.



- ___10. View the archive file with Business Objects and Cognos.
 - _a. The **Demos folder** on the desktop contains several video files showing examples of various BI tools accessing the archive file. Double-click on the files to see how reports can easily be created.

If you have existing BI reports today that are reporting against the data sources that you plan to archive or decommission, often those same reports can be redirected at the archive files just by changing the connection.



CONGRATULATIONS, you have successfully finished the lab.

Page 24 Optim PoT Labs

Lab 2 Discovering Optim Delete and Restore Processing

In the previous lab we created and ran our data archiving process. We archived all orders and orders related data for all customers who have an ID less than 10100. By using Optim, we were able to archive data related to our orders table. This related data included order details, customers, and items. Though we archived data from these four related tables, we specified that we only wanted to delete data from the orders and details tables.

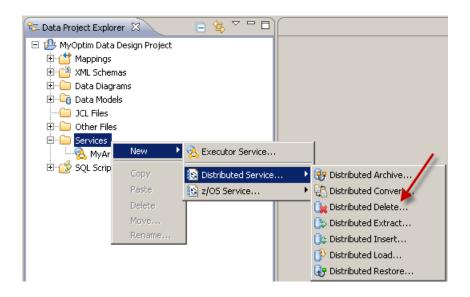
In this lab you will define the request to actually delete the data from the source that we archived from. The process will be very similar to what we did during the archive lab. We will be defining both delete and restore processes in this lab. By doing this we will show how to delete archived data from a source database, as well as how to restore that deleted data (if the need arises to restore data that was previously deleted from the source database).

LAB STEPS

- Start Optim Designer
 - a. Double click on the **Optim Designer** icon on your desktop.



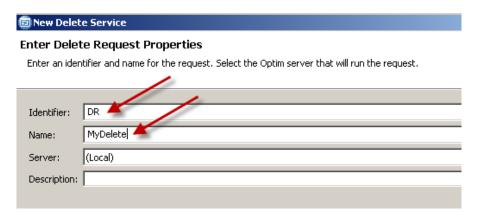
- ___2. Create Optim Delete Service for the delete process.
 - __a. From the **Data Project Explorer**, drill into the **MyOptim Data Design Project**, right click on the **Services** folder and choose **New > Distributed Service > Distributed Delete**.



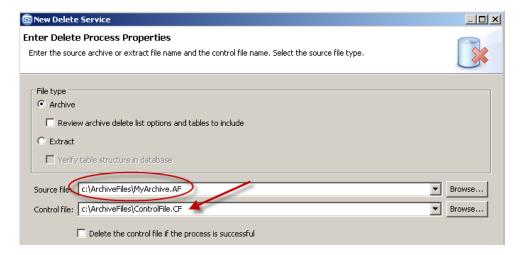
__b. Name the Delete Request Model "MyDelete". Click Next.



__c. Enter "DR" for the Identifier and "MyDelete" for the Name as shown below. Click Next.



__d. Be sure the archive file is the **MyArchive.AF** that you created in the Archive Lab. Enter "C:\OptimExtracts\ControlFile.CF" for the Control file name. Click **Next**.

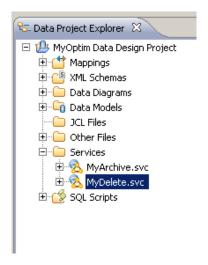


Page 26 Optim PoT Labs

📵 New Delete Service **Enter Delete Processing Options and Limits** Enter options and limits for the delete process. Processing options Lock tables until all rows are processed Generate statistical information in the Delete Process Report Compare row content to only delete rows that exactly match rows in the source file ☐ Include LOBs in row content comparison Processing limits Commit frequency: 200 Discard row limit: 0 Database connections: ▾

e. Check the Generate statistical information ... box as shown below. Click Finish.

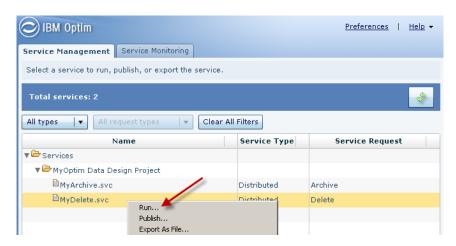
You will now see the MyDelete.svc in your project.



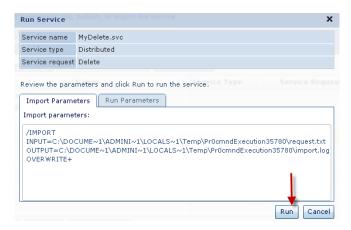
- __3. Run the Optim service to run the delete request.
 - __a. We will now open Optim Manager from a web browser so that we can monitor and run the newly created service. Open **Mozilla Firefox** and you will open Optim Manager.



__b. You will be presented with screen with two tabs, one for service management and the other for service monitoring. You will see your MyDelete.svc Service listed under the Service Management tab. Right click on MyDelete.svc and select Run.



__c. Click on **Run** from the Run Service Screen.

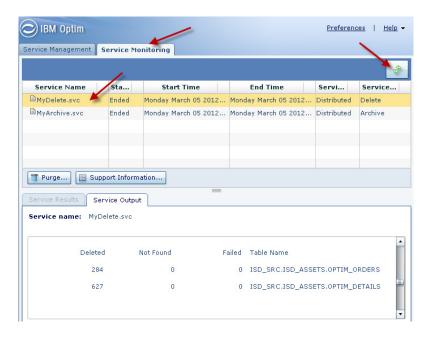


__d. You should receive the following confirmation box, stating that your job has been submitted. Click **Close**.



Page 28 Optim PoT Labs

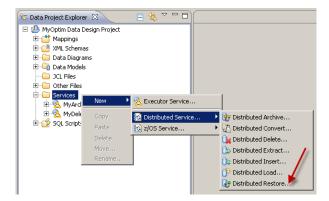
_e. Select the Service Monitoring tab and you will eventually see that your service completed successfully. Use the refresh button (two green arrows on the right side of the screen) to refresh the screen. Look within the Service Output tab and you will see the number of rows deleted from the OPTIM ORDERS and OPTIM DETAILS tables.



__4. Create Optim service for the restore process.

The need may arise to restore all or part of the contents of data that was previously deleted from your database Optim delete processing. The deleted data that resides in your archive file can always be restored to a database. We will restore the data deleted in the previous step via Optim restore processing.

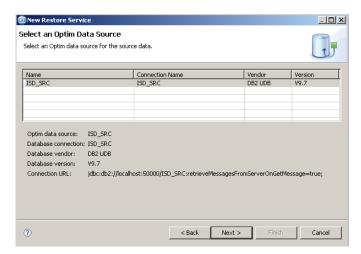
__a. From the **Data Project Explorer**, drill into the **MyOptim Data Design Project**, right click on the **Services** folder and choose **New** → **Distributed Service** → **Distributed Restore**.



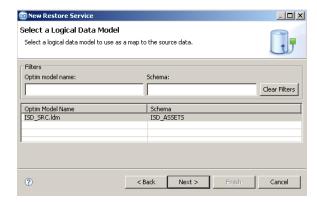
__b. Name the Restore Request "MyRestore". Click Next.



__c. Take the default ISD_SRC as the Optim Data Source. Click **Next**.

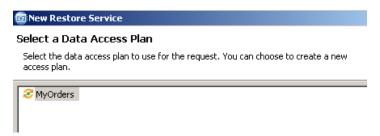


__d. Take the default ISD_SRC.ldm as the Logical Data Model. Click Next.

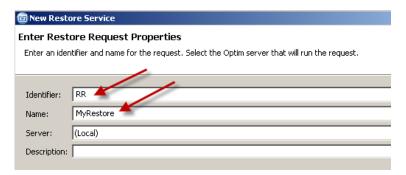


Page 30 Optim PoT Labs

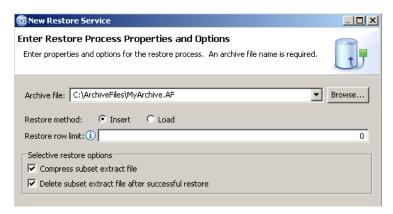
__e. Select the **MyOrders** DAP that was used by the archive process. We will also use it for the restore process. Click **Next**.



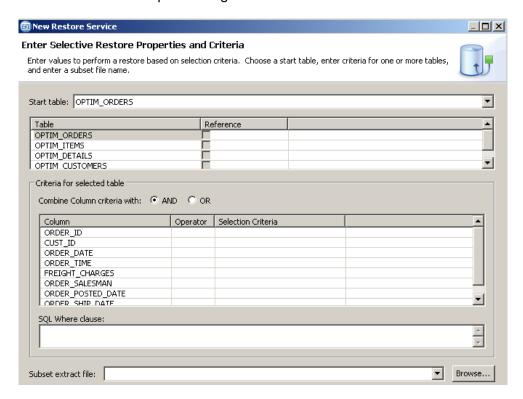
__f. Enter "RR" as the Identifier and "MyRestore" as the Name for the Restore Request Properties. Click Next.



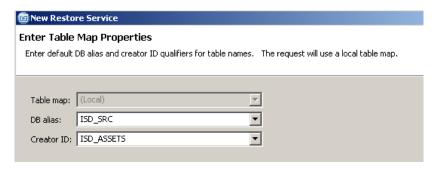
__g. Take the defaults in the Restore Properties and Options. Click **Next**.



__h. Be sure that **OPTIM_ORDERS** is the start table option. This was the same start table we used for our archive processing. Click **Next**.

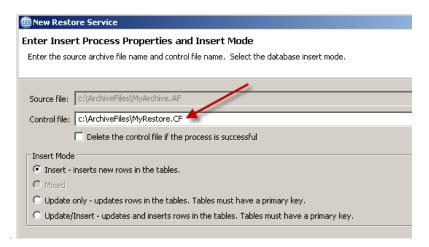


__i. We want to use the **ISD_SRC** database to restore into and a schema of **ISD_ASSETS**. Click **Next**.

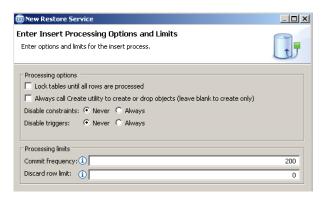


Page 32 Optim PoT Labs

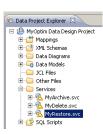
__j. Enter "c:\ArchiveFiles\MyRestore.CF" for the Control File. Click Next.



__k. Take the defaults for the Insert Processing Options and Limits. Click **Finish**.



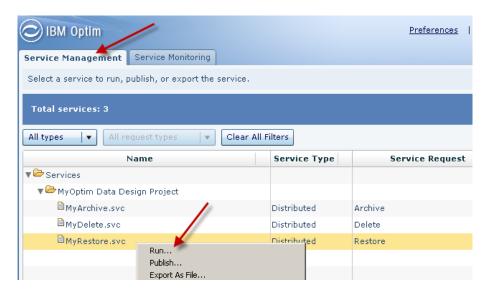
I. You will see the **MyRestore.svc** service.



- __5. Create the Optim service to run the restore request.
 - __a. We will now open Optim Manager from a web browser so that we can monitor and run the newly created service. Open **Mozilla Firefox**.



__b. You will be presented with screen with two tabs, one for service management and the other for service monitoring. You will see your MyRestore.svc service listed under the **Service Management** tab. Right click on **MyRestore.svc** and select **Run**.



- c. Click the **Run** button from the Run Service screen and click **Close** on the confirmation.
- _d. Select the **Service Monitoring** tab and you will eventually see that your service completed successfully. Use the **refresh** button (two green arrows on the right side of the screen) to refresh the screen. Look within the service output tab and you will see the number of rows inserted into the **OPTIM_ORDERS** and **OPTIM_DETAILS** tables. Note that the inserts for the **OPTIM_CUSTOMERS** and **OPTIM_ITEMS** tables failed, this is because you did not delete data from these two tables and duplicate rows were not inserted during the restore.



CONGRATULATIONS, you have successfully finished the lab.

Page 34 Optim PoT Labs

Lab 3 Adding Optim Relationships

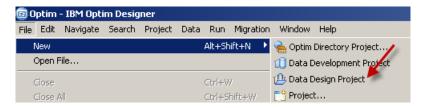
In the Archive lab that you finished earlier, Optim relationships were defined automatically. Optim used the relationships already defined at the database level to archive related data. What if the database that you are archiving from has no relationship defined in the database? What if you need to add additional relationships to what are already defined at the database level? Not a problem for Optim. You can easily define new and additional relationships within Optim. The relationships that you define to Optim allow you to archive related data, even if the relationships are not part of the relational database. In this lab we will show you how to add additional relationships within Optim.

LAB STEPS

- __1. Open Optim Designer.
 - a. Double click on the Optim Designer icon on your desktop.



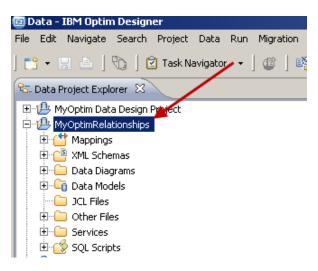
- ___2. Create a new Data Design Project.
 - _a. Choose **File** → **New** → **Data Design Project** from the top task bar.



__b. You can give the project any name you like. For the purposes of this PoT call it "MyOptim Relationships". Click Finish.



__c. You will now see the project that you just created listed in the Data Project Explorer part of your screen.

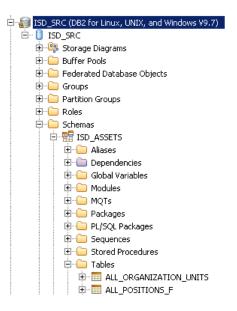


- Create the Physical Data Model.
 - _a. The tables for our archive process reside in the database called ISD_SRC. This can be seen by expanding the **Database Connections** folder found in the **Data Source Explorer** in the bottom left of you screen. Double click on the **ISD_SRC** icon to connect to the database.

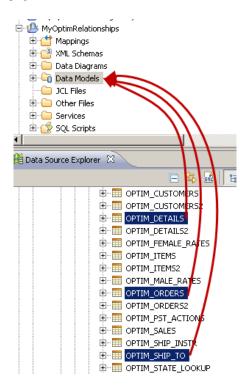


Page 36 Optim PoT Labs

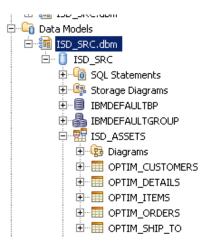
__b. Once connected, expand down to the **Schemas** folder to see the schemas of the tables within this database. The **ISD_ASSETS** schema contains the tables we will use for our archive process.



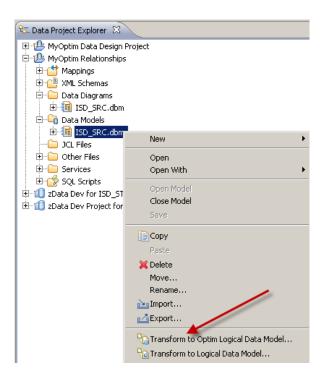
_c. We will now create our Physical Data Model. There are a few ways to do this. In this lab we will use the drag and drop method. Drill down into the ISD_ASSETS schema, hold the CTRL key and select the OPTIM_ORDERS, OPTIM_DETAILS, and OPTIM_SHIP_TO tables and drag and drop them onto the Data Models folder within the Data Project Explorer.



__d. After completing this drag/drop step, you will see a model called ISD_SRC.dbm under the Data Models folder. You will see that two additional tables, OPTIM_CUSTOMERS and OPTIM_ITEMS were added by Optim. These additional tables are the result of the RI that is defined in the database.

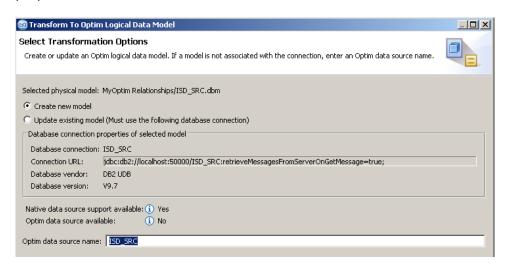


- ___4. Create the Optim Logical Data Model from the Physical Data Model.
 - _a. Right click on the ISD_SRC.dbm model and choose Transform to Optim Logical Data Model.

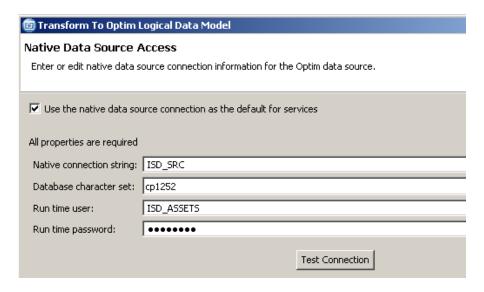


Page 38 Optim PoT Labs

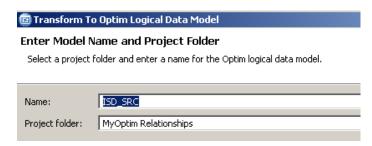
__b. When presented with the transformation options screen be sure that **Create New Model** is selected and that your database name is **ISD_SRC**. Once you have verified the proper selections, click **Next**.



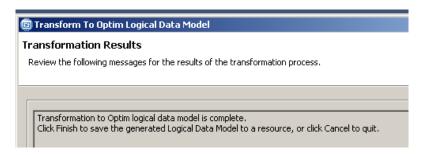
__c. Enter the following options on the Native Data Source Access screen. Click Next.



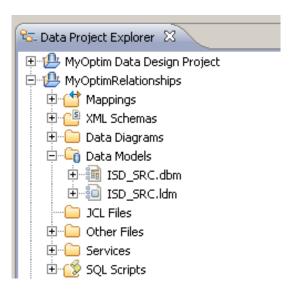
__d. Be sure that the following are selected for the Name and Project folder. Click **Next**.



__e. The following screen should appear if everything was successful. Click **Finish**.

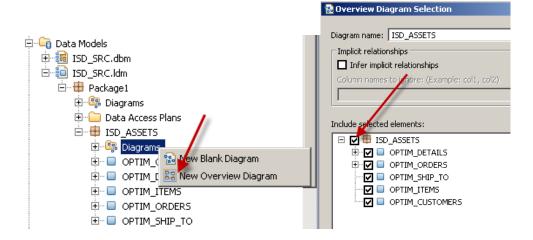


__f. You should now see two models under your **Data Models** folder, **ISD_SRC.dbm** and **ISD_SRC.ldm**. We will be working in the ISD_SRC.ldm model.

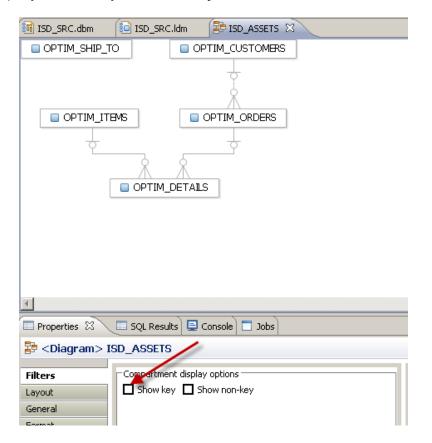


Page 40 Optim PoT Labs

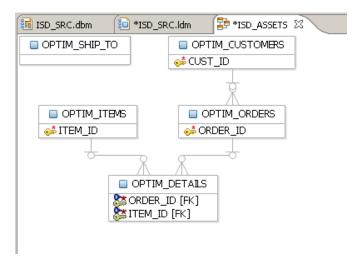
__g. Drill into the ISD_SRC.Idm model and create an Overview Diagram by right clicking on Diagrams under ISD_ASSETS and selecting New Overview Diagram. Be sure to check all tables. This will enable you to see the tables in a diagram. Click OK.



_h. The resulting diagram will only display entity names. This is because the Show key filter property is not chosen. Select the **Properties** tab, in the bottom right part of your screen. Select the **Show key** check box to view a diagram that includes key attributes. (Optional) Try different layouts in the **Layout** tab.



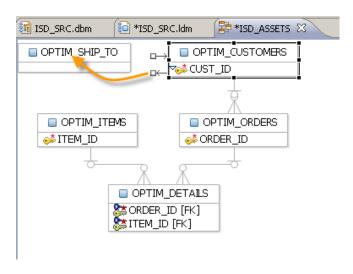
__i. The resulting diagram now includes the keys.



__5. Draw new Optim relationship from the OPTIM_CUSTOMER entity to the OPTIM_SHIP_TO_ENTITY.

We want to be able to have a relationship between our table containing customer data with our entity containing OPTIM_SHIP_TO information. This will allow us to pull ship to address information about customers.

_a. Highlight the OPTIM_CUSTOMERS entity and hold the cursor inside the entity until the relationship pop-ups appear. Drag and drop the one-to-many pop-up from OPTIM_CUSTOMERS to OPTIM_SHIP_TO.

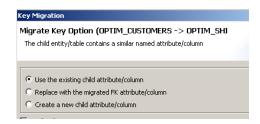


Page 42 Optim PoT Labs

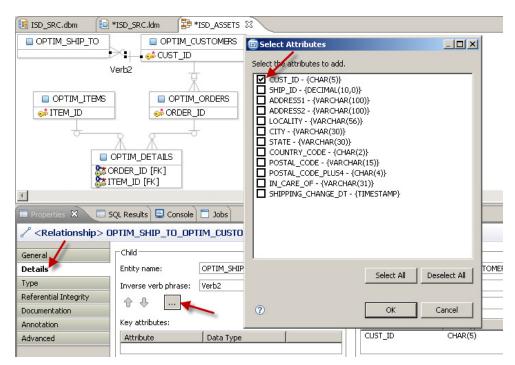
__b. You will be presented with an option to choose the type of relationship you want. For this relationship choose Create Non-Identifying Mandatory Foreign Key Relationship.



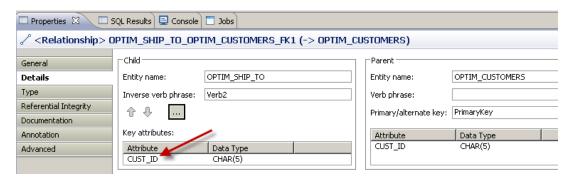
__c. Click **OK** to use the existing child attribute/column.



__d. A relationship line will now be drawn between the two entities. Click on the relationship to highlight it. In the Properties window and from the Details section click on the ellipsis (...) within the Child entity section and then check the CUST_ID attribute as the foreign key choice. Click OK.

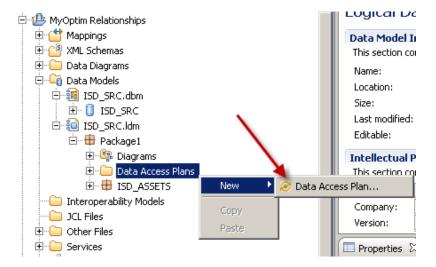


_e. After making the selection in the above step, you will see that **CUST_ID** is now filled in as your attribute.



You have now defined a new relationship to Optim. Even though this relationship does not exist on our source database, as a result of this new relationship within Optim, you can use this relationship to archive or extract related data that involves customer and ship to information. To prove that this newly created Optim relationship is correct, you will create a Data Access Plan that automatically uses this relationship in the next step.

- __f. Save diagram via the **save icon** lor via the "CTRL-S" keys.
- 6. Create Data Access Plan (DAP) that uses the new Optim relationship created in Step 5.
 - __a. Drill down into the ISD_SRC.Idm model. Right click on the Data Access Plans folder and choose New → Data Access Plan



Page 44 Optim PoT Labs

__b. Name the DAP "Customer Ship To". Click Next.



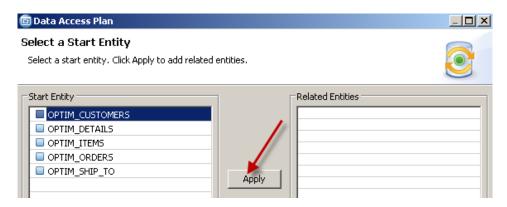
__c. ISD_ASSETS will be your only package choice. Click Next.



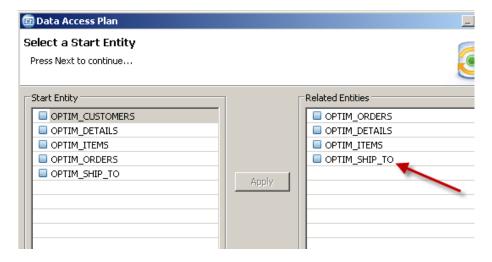
__d. Select the button **Select entities based on relationships with a start entity**. Click **Next**.



__e. Select **OPTIM_CUSTOMERS** as the start table. Click **Apply**. This will automatically pick up the entities related to OPTIM_CUSTOMERS.



__f. Note that Optim picked up the OPTIM_SHIP_TO entity as a related entity because of you adding this relationship within Optim. Click **Finish**.



This Data Access Plan (DAP) can now be used in any Optim process (archive, extract, etc...) and will pull all related OPTIM_SHIP_TO information for all OPTIM_CUSTOMERS data.

CONGRATULATIONS, you have successfully finished the lab.

Page 46 Optim PoT Labs

Lab 4 Discovering Optim Test Data Management Capabilities

Optim Test Data Management (TDM) allows customers to move referentially intact subsets of complex relational data. In this lab we will use IBM's Optim solution to create subsets of data for our test environment. We will be working with four tables in this lab: Orders, Details, Customer, and Items. These tables contain the following information for our TDM scenario:

ISD ASSETS.OPTIM ORDERS - Contains information about orders placed by customer

ISD ASSETS.OPTIM DETAILS - Contains the order details associated with a customer order

ISD_ASSETS.OPTIM_CUSTOMERS - Contains Customer information

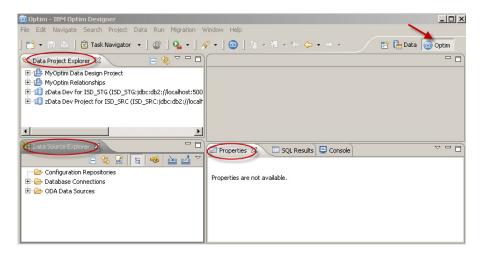
ISD ASSETS.OPTIM ITEMS - Contains information about items sold

Lab Steps

- __1. Open Optim Designer.
 - __a. Double click on the **Optim Designer** icon on your desktop. It may take a minute or so to initially start Optim Designer, so please be patient.

Optim

_b. View the Optim Designer screen that displays.



The upper left portion of the screen is called the **Data Project Explorer**. This contains all the Optim objects that you've created to perform various Optim tasks, such as an archive. As we create our archive process, you will see the items that you create within the folders in the Data Project Explorer.

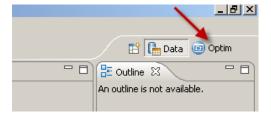
The bottom left portion of the screen is called the **Data Source Explorer**. This contains the database connections that are defined that will be used during our lab.

The top right portion of the screen is the editor screen and will vary in its content based on the type of item that you have opened.

The bottom right of the screen is the **Properties View**. It will contain all of the properties associated with a selected object.

_c. Be sure that you are in the Optim Perspective by clicking on the **Optim** icon in the upper

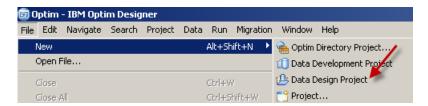
right part of your screen.



__2. Create a Data Design Project.

We will now create a new Data Design Project. This project will contain all of the objects we create that are used to define our data subsetting activities.

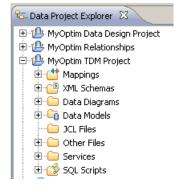
a. Choose File → New → Data Design Project from the top task bar.



__b. You can give the project any name you'd like. For the purposes of this PoT call it "MyOptim TDM Project". Click Finish.

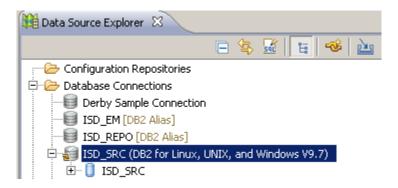


__c. You will now see the project that you just created listed in the Project Explorer part of your screen.

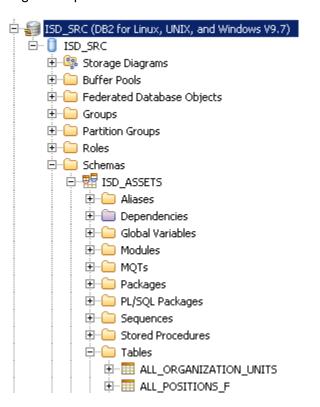


Page 48 Optim PoT Labs

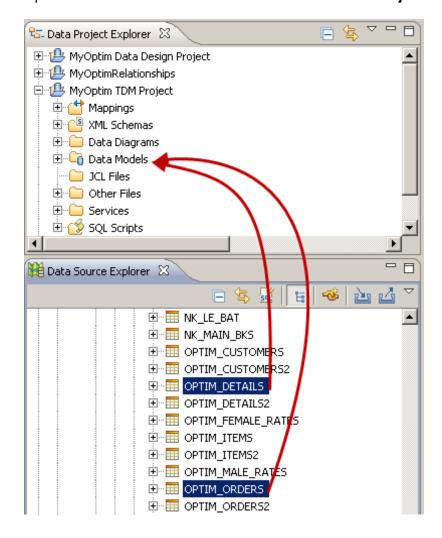
- 3. Create the Physical Data Model for the source Production system, ISD SRC.
 - _a. The source production tables for our test data management process reside in the database called ISD_SRC. This can be seen by expanding the **Database Connections** folder found in the **Data Source Explorer** at the bottom left of your screen. Double click on the **ISD SRC** icon to connect to the database.



__b. Once connected, expand down to the **Schemas** folder to see the schemas of the tables within this database. The **ISD_ASSETS** schema contains the tables we will use for our test data management processes.

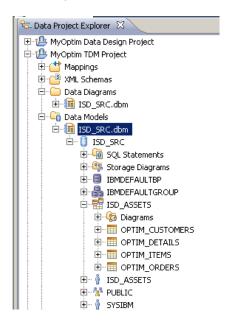


_c. We will now create our Physical Data Model. There are a few ways to do this. In this lab we will use the drag and drop method. Drill down into the ISD_ASSETS schema, hold the CTRL key and select the OPTIM_ORDERS, and OPTIM_DETAILS tables and drag and drop them onto the Data Models folder within the Data Project Explorer.



Page 50 Optim PoT Labs

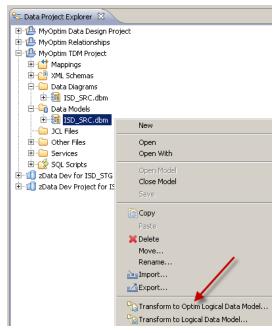
__d. After completing this drag/drop step, you will see a model called ISD_SRC.dbm under the Data Models folder. You will see that two additional tables, OPTIM_CUSTOMERS and OPTIM_ITEMS were added by Optim. These additional tables are the result of the RI that is defined in the database. (NOTE: In the previous lab you saw how Optim can add relationships, even if they are not defined in the database).



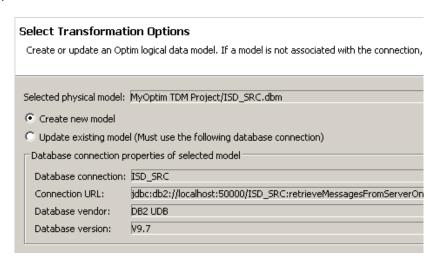
___4. Create the Optim Logical Data Model from the Physical Data Model.

You now need to create what is called the Optim logical data model from the ISD_SRC.dbm physical data model.

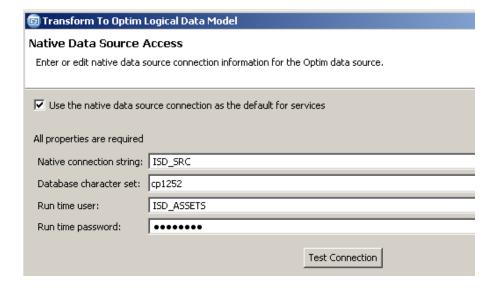
__a. Right click on the ISD_SRC.dbm model and choose Transform to Optim Logical Data Model.



__b. When presented with the transformation options screen, be sure that **Create New Model** is selected and that your database name is **ISD_SRC**. Once you have verified the proper selections, click **Next**.

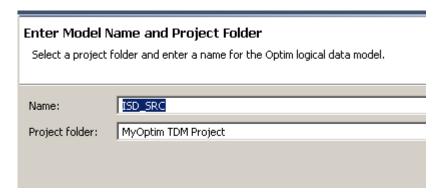


__c. Enter the following options on the Native Data Source Access screen. Click Next.

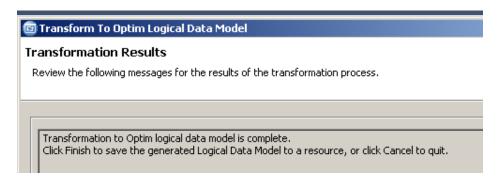


Page 52 Optim PoT Labs

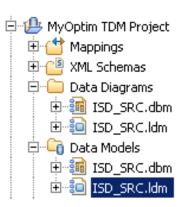
d. Be sure that the following are selected for the Name and Project folder. Click Next.



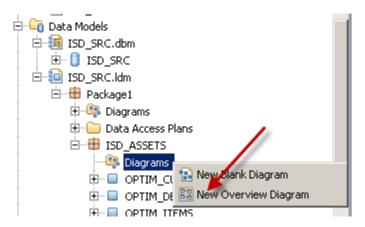
__e. You will see this message after the transformation is complete. Click **Finish**.



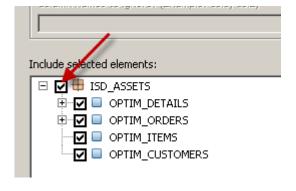
__f. You will now see the newly created **ISD_SRC.ldm** Optim logical model in your **Data Project Explorer**.



__g. To create a diagram to show the entities involved with your Optim logical model, right-click on the **Diagram** folder underneath the **ISD_ASSETS** Package and select **New Overview Diagram**. Note, be sure to select the 2nd Diagrams icon.

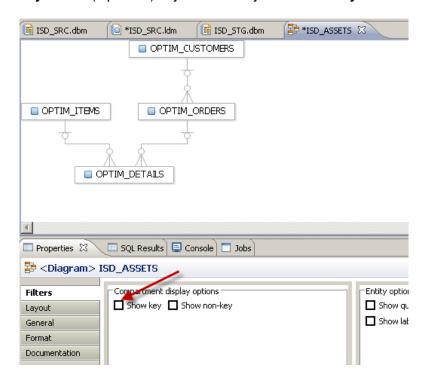


__h. Check all the ISD_ASSETS entities. Click OK.

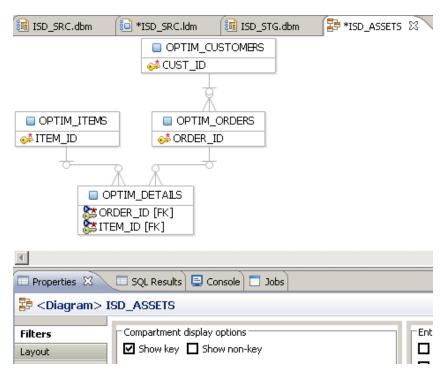


Page 54 Optim PoT Labs

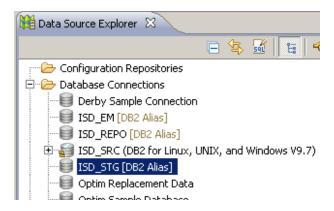
__i. The resulting diagram will have no attributes shown. We would like to see key attributes. Select the **Properties** tab shown underneath the diagram and choose **Filters**. Check the **Show key** box. . (Optional) Try different layouts in the **Layout** tab.



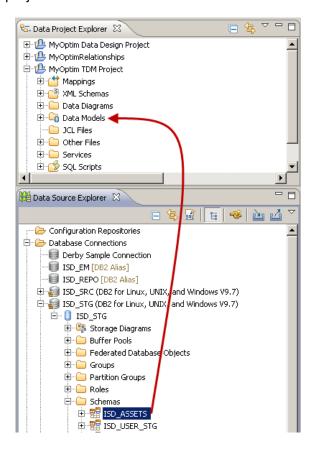
__j. You will now see keys in your diagram.



- __5. Create the Physical Data Model for the target/test database
 - _a. Now that we have defined the source data, we need to identify the target database for the test data by creating physical and logical data models for the target environment. In this lab we will be loading our subset of test data to a non-production environment called ISD_STG. Connect to the ISD_STG target test database by double clicking on the ISD_STG database from the Data Source Explorer.

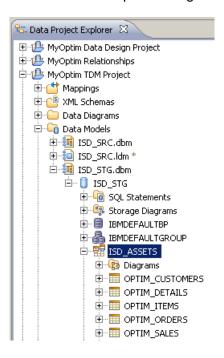


_b. Once connected, expand down to the Schemas folder to see the schemas of the tables within our target test database. The ISD_ASSETS schema contains the tables we will be populating with data. Drag and Drop the ISD_ASSETS schema onto the Data Models Folder in our project.

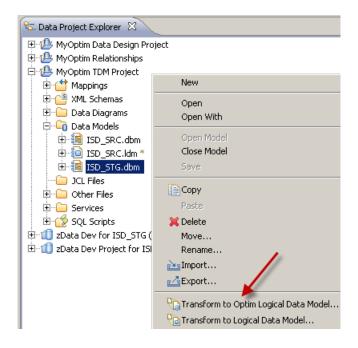


Page 56 Optim PoT Labs

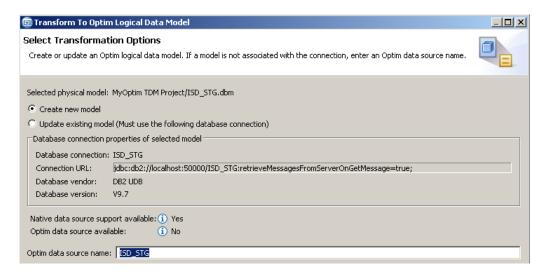
_c. You will now see the **ISD_STG.dbm** physical model in your project. Notice that we have 5 tables that can be used for our TDM processing. In this lab we will only use 4.



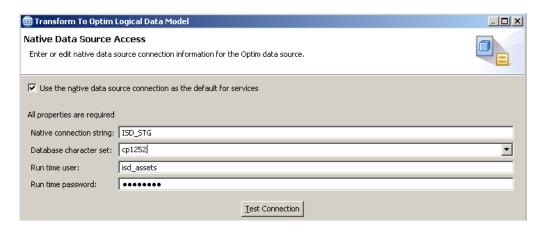
- __6. Create the Optim Logical Data Model from the physical model just created.
 - __a. Right click on the ISD_STG.dbm model and choose Transform to Optim Logical Data Model.



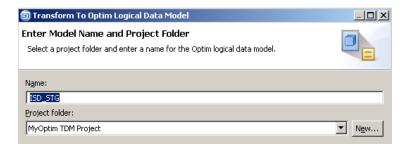
__b. When presented with the Transformation Options screen, be sure that **Create New Model** is selected and that your Optim data source name is **ISD_STG**. Once you have verified the proper selections, click **Next**.



_c. Enter "ISD_STG" as the native connection string and "cp1252" as the database character set. Press Next.

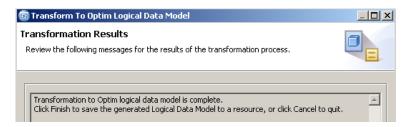


__d. Be sure the model and project fields are filled as shown below. Press **Next**.



Page 58 Optim PoT Labs

__e. From the Transformation Results screen, click **Finish**.



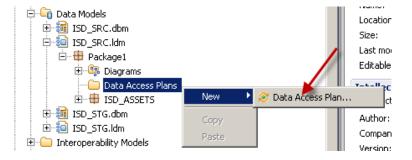
__f. **(Optional)** Create the "Overview Diagram" as you did for the earlier lab (see 4g through 4j for additional detail). The diagrams are not required for processing.

We now have the needed physical and Optim logical models for the rest of our lab.

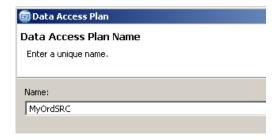
__7. Create the Data Access Plan for the Source production system (ISD_SRC).

A Data Access Plan contains policies for the selection and transformation of data. These policies are an important element for a Test Data Management solution, where we need to consistently extract data from our production systems that represent different use cases and test cases, and mask or privatize sensitive or personally identifiable data as we load it into our non-production systems. The Data Access Plan allows you to specify what data you would like to select from the source database system for inclusion in your target database system. Within the Optim Logical Model that you just created you will find a folder for Data Access Plans.

__a. Drill into the **ISD_SRC.Idm** Optim logical model. Right click on the **Data Access Plans** folder to create a **new Data Access Plan** (DAP).



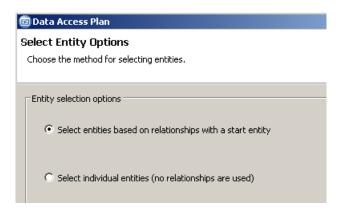
__b. Name the DAP "MyOrdSRC". Click Next.



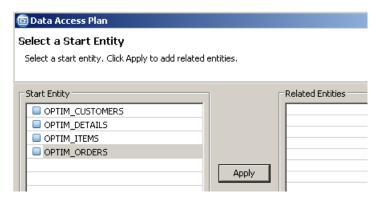
__c. Choose ISD_ASSETS as your package choice. Click Next.



__d. For the Entity Option, choose **Select entities based on relationships**. Click **Next**.

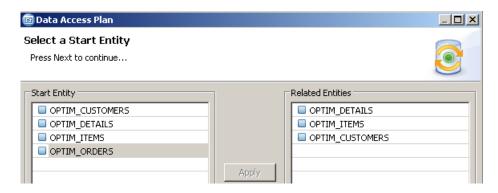


__e. Select OPTIM_ORDERS as the start table. Click Apply. This will automatically pick up the entities related to OPTIM_ORDERS.



Page 60 Optim PoT Labs

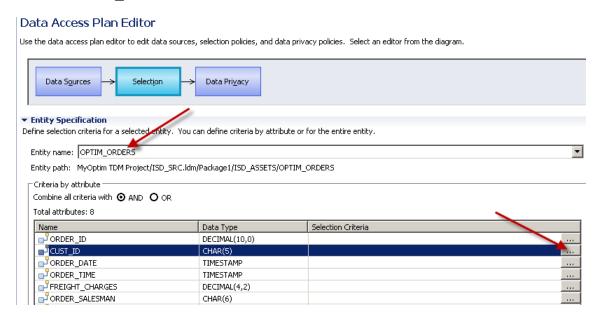
__f. The resulting screen will show the related entities. The start entity and the related entities identify all of the tables that will be included in the extract from the production system. Click **Finish**.



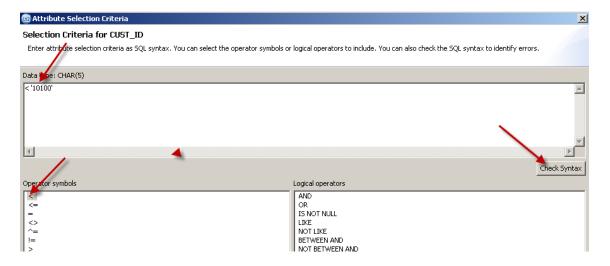
_g. Click the **Selection** option from the Editor. This will allow you to identify the subsetting criteria from the collection of tables that will be extracted from the production system.



_h. Scroll down to the Entity Specification section of the screen. Click on the arrow to expand this section if you can't see it. Be sure to select the OPTIM_ORDERS entity, and then select the ellipsis (...) to the far right of the CUST_ID attribute to define the selection criteria for CUST_ID.



_i. Double click on the "<" symbol and type in '10100' as shown below. We will be extracting all orders (and related data), for CUST_ID less than 10100. Click the **Check Syntax button** to be sure that your syntax is correct.

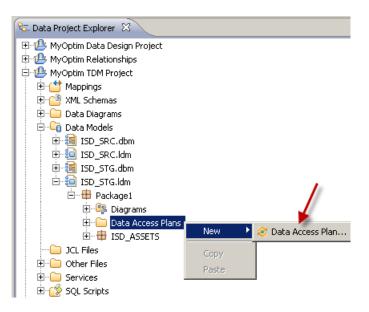


__j. Click OK. Save via the Save icon or CTRL-S keys.

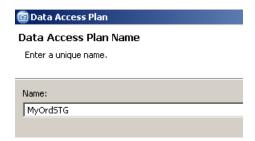
You have now created a Data Access Plan to select all orders for those customers with a customer ID less than 10100.

Page 62 Optim PoT Labs

- __8. Create the DAP for the target test system (ISD_STG).
 - __a. Drill into the **ISD_STG.Idm** Optim logical model. Right click on the **Data Access Plans** folder to create a **new Data Access Plan** (DAP).



__b. Call the DAP "MyOrdSTG". Click Next.



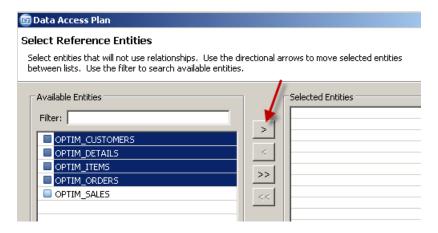
_c. ISD_ASSETS should be your only package choice. Click Next.



__d. Unlike our source production system, the tables in our target test system do not have any RI defined against them. Choose **Select individual entities**. Click **Next**.



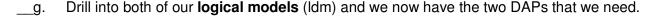
_e. Select the OPTIM_CUSTOMERS, OPTIM_DETAILS, OPTIM_ITEMS, and OPTIM_ORDERS entities by highlighting them with the "CTRL" key pressed and selecting the ">" button.

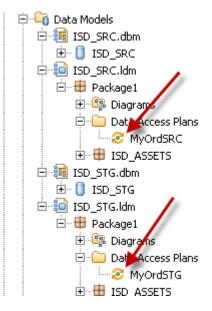


__f. Click **Finish**. Save your work via the **Save icon** or **CTRL-S** keys.



Page 64 Optim PoT Labs

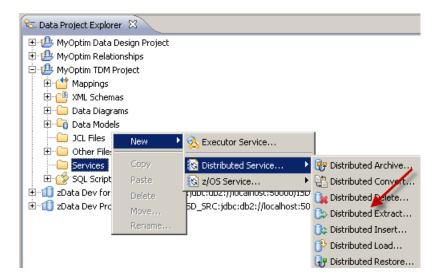




Create the Optim service to run the extract request.

We are now ready to create our extract request. We do this by creating what Optim calls an Extract Service that will actually run on the Optim server. Optim comes with a component called the Optim Manager. Administrators can use this component to monitor and run services created from Optim Designer. Optim Manager can be accessed from Optim Designer or from a browser. We will use the browser interface in this lab. We'll first create the service from Designer and then will run it from Optim Manager.

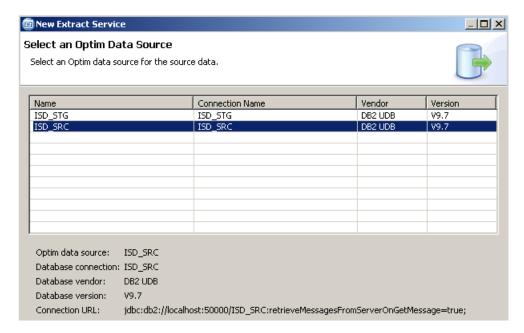
__a. From the **Data Project Explorer**, drill into the **MyOptim TDM Project**. Right click on the **Services** folder and choose **New** → **Distributed Service** → **Distributed Extract**.



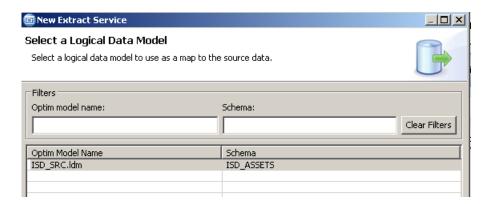
b. Name the Extract Service "MyOrdExt". Click Next.



__c. Select the **ISD_SRC** model as we are extracting from the source production DB. Click **Next**.

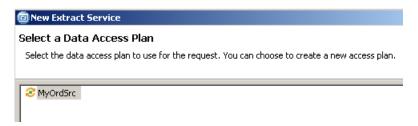


d. Take the default. Click **Next**.

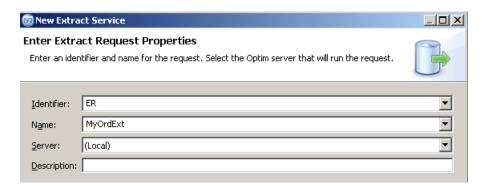


Page 66 Optim PoT Labs

__e. Select the MyOrdSRC Data Access Plan that you created earlier. (Remember, this DAP selects all orders and related tables for CUST_ID less than 10100). Click Next.



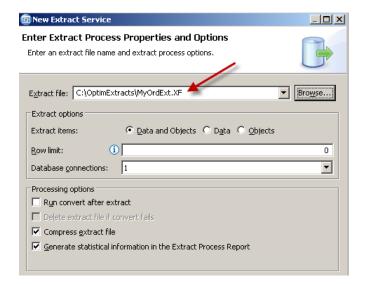
__f. Enter "ER" as the identifier and "MyOrdExt" as the name in the Extract Request Properties. Click Next.



__g. Enter ISD_SRC as the DB alias and ISD_ASSETS for the creator id. Click Next.



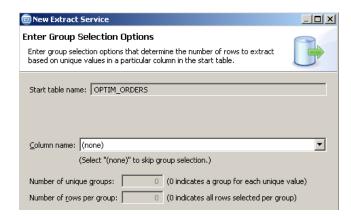
h. Name the extract file "C:\OptimExtracts\MyOrdExt.XF". Click Next.



i. Take the defaults. Click Next.

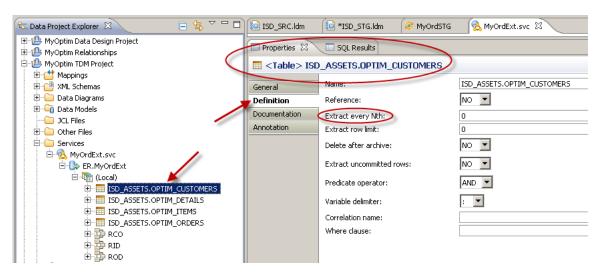


j. Take the defaults. Click **Finish**.

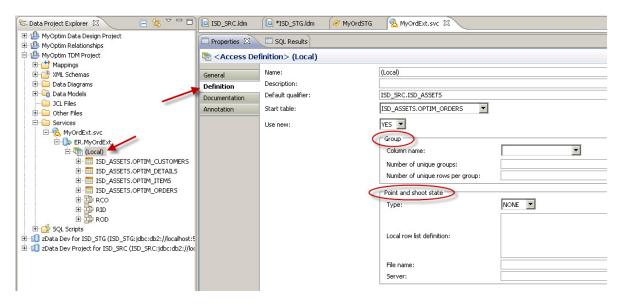


Page 68 Optim PoT Labs

- __k. You will now see your **MyOrdExt.svc** service under the **Services** folder.
- __I. To see some additional options for the extract criteria, double-click the **MyOrdExt.svc** service, **expand the tree, highlight a table**, and look at the **Definition** tab of the **Properties view**, as shown below. Notice there are data sampling options (extract every Nth row), and options for extract row limits. To include these options in your extract model, simply enter a value and save the service.



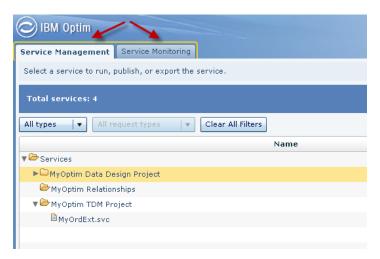
_m. There are also additional options available such as Grouping (for example, extract 20 customers from each of the 50 states) and Point and Shoot. The Point and Shoot facility allows you to manually select specific rows from the start table. This can be helpful when you require specific rows to be included in the extract to facilitate Use Case testing.



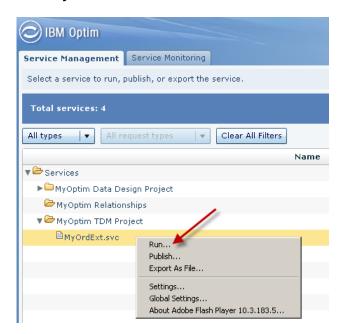
__n. We will now open Optim Manager from a web browser so that we can run and monitor the newly created service. Open Mozilla Firefox.



__o. You will be presented with screen with two tabs, one for service management and the other for service monitoring. You will see your MyOrdExt.svc listed under the Service Management tab.

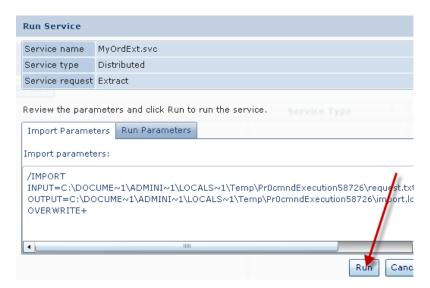


__p. Right click on the **MyOrdExt.svc** service. Select **Run**.



Page 70 Optim PoT Labs

g. Click **Run** from the Run Service Screen.



__r. You should receive the following confirmation box, stating that your job has been submitted. Click Close.



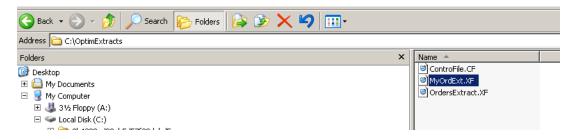
_s. Select the **Service Monitoring** Tab and you will eventually see that your service completed successfully. Use the **refresh** button (two green arrows on the right side of the screen) to refresh the screen.



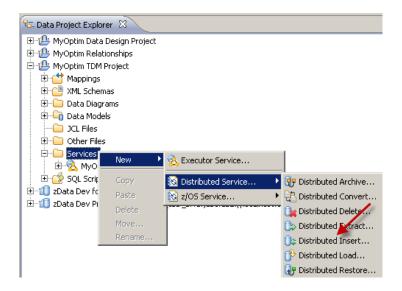
__t. Highlight the MyOrdEx.svc service and select the Service Output tab at the bottom of the screen. Scroll down and you will see the number of rows successfully extracted into your extract file.



__u. You may also view the existence of the extract file from Windows Explorer.



- __10. Now that we have our extract file, we are ready to load the data into the target database. To do this we need to create an Optim Insert Service.
 - __a. From the **Data Project Explorer**, drill into the **MyOptim TDM Project**, right click on the **Services** folder and choose **New** \rightarrow **Distributed Service** \rightarrow **Distributed Insert**.

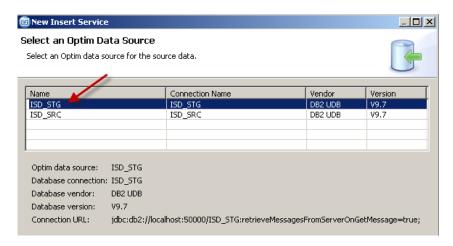


Page 72 Optim PoT Labs

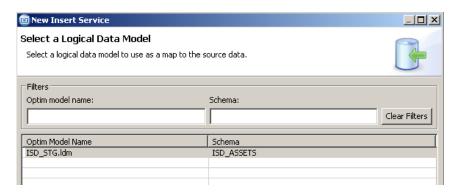
__b. Name the Insert Service **MyOrdIns**. Click **Next**.



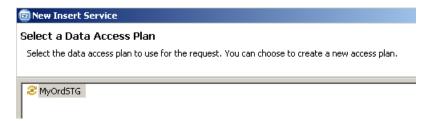
__c. Select ISD_STG for the Optim Data Source. Click Next.



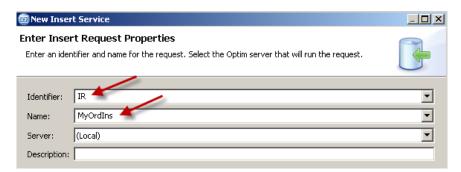
__d. Take the default. Click Next.



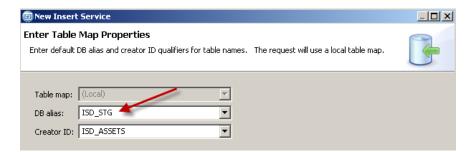
__e. Select the MyOrdSTG Data Access Plan that you created earlier. Click Next.



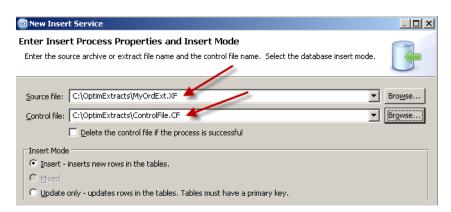
__f. Enter "IR" for the identifier and "MyOrdIns" for the name. Click Next.



g. Enter ISD_STG for DB alias and ISD_ASSETS for Creator ID. Click Next.

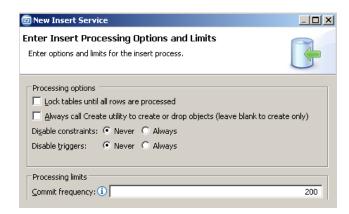


_h. The source file for the insert request will be the "C:\OptimExtracts\MyOrdExt.XF" extract file that we created earlier. Browse can be used to find the file. Use "C:\OptimExtracts\ControlFile.CF" as the control file. Click Next.

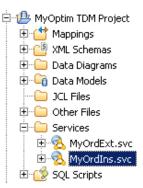


Page 74 Optim PoT Labs

__i. Take the defaults. Click Finish.



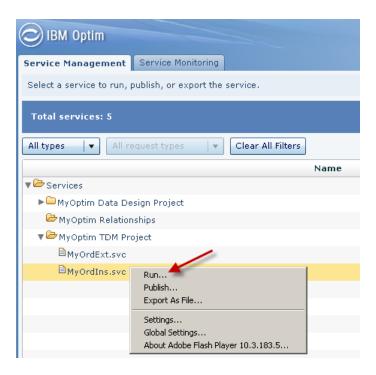
__j. You will now see your **MyOrdIns** service under the **Services** folder



__k. We will now open Optim Manager from a web browser so that we can run and monitor the newly created service. Open **Mozilla Firefox**.



__l. As we did before with our Extract service, you will see the **MyOrdIns.svc** service listed under the **Services Management** tab. Right click on the **MyOrdIns.svc** service and select **Run**.

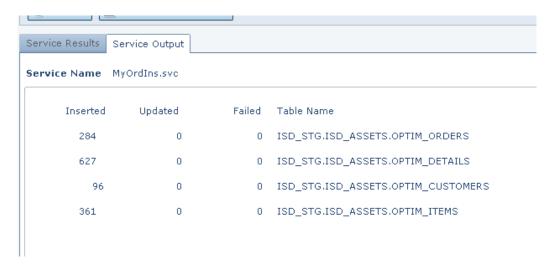


- m. Click the **Run** button from the Run Service screen and click **Close** on the confirmation.
- __n. Select the Service Monitoring Tab and you will eventually see that your service completed successfully. Use the refresh button (two green arrows on the right side of the screen) to refresh the screen.

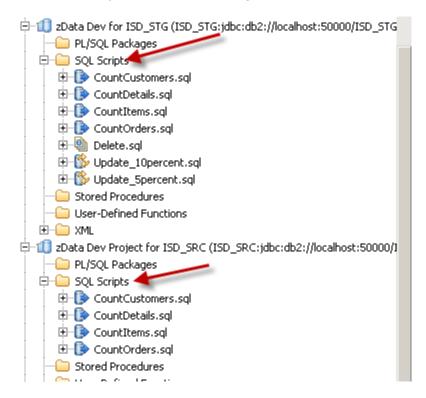


Page 76 Optim PoT Labs

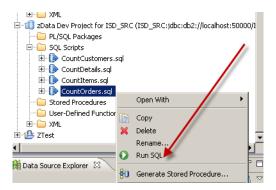
_o. Highlight the MyOrdIns.svc service and select the Service Output tab at the bottom of the screen. Scroll down and you will see the number of rows successfully inserted into your non-production database tables. Notice that they match the counts from your extract results.



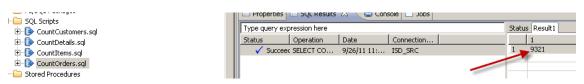
- __11. Verify row counts to prove ISD_STG (our test database) contains a subset of what is in ISD SRC (our production database).
 - __a. Open the two **zData Development** projects and drill down to the **SQL Scripts** folders to see a list of SQL scripts available to run against the **ISD_SRC** and **ISD_STG** databases.



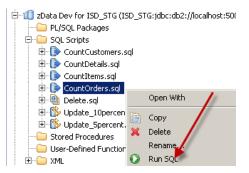
__b. Run the **CountOrders.sql** from the **ISD_SRC** database. Do this by right clicking on the script and selecting **Run SQL**.



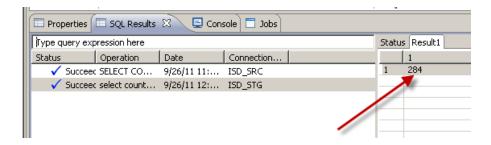
c. You will see the output of the query in the bottom right part of your screen.



__d. Run the CountOrders.sql from the ISD_STG database. Do this by right clicking on the script and selecting Run SQL.



e. You will see the output of the query in the bottom right part of your screen.



You should see that the ISD_STG test database has far fewer rows than the ISD_SRC productions database.

CONGRATULATIONS, you have successfully finished the lab.

Page 78 Optim PoT Labs

Lab 5 Discovering Optim Data Privacy Capabilities

In the previous lab we created an extract file which was a subset of related data from a production system. We extracted all orders for customers having a customer ID less than 10100. The related data that was included in the extract file consisted of order details, and item and customer information. Together this information constitutes the business object for Orders. After the data was extracted from the production system, it was then loaded into a non-production environment represented by the staging database.

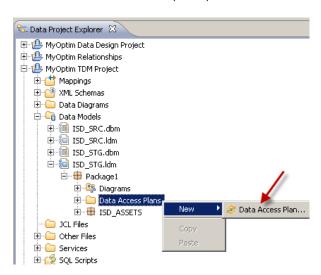
In this lab we will define a request to mask the data. We will then use Optim's Compare function to compare the extract file from Lab 4 with the masked extract file from Lab 5, to validate our privacy policies.

Lab Steps

- __1. Open Optim Designer.
 - __a. Double click on the **Optim Designer** icon on your desktop.



- __b. We will continue to work in the TDM project that we created in Lab 4.
- __2. Create a new Data Access Plan for the target non-production system ISD_STG. We will specify the privacy policies in this DAP.
 - __a. Drill into the **ISD_STG.Idm** Optim logical model. Right click on the **Data Access Plans** folder to create a **New Data Access Plan** (DAP).



__b. Name the DAP **MyOrdSTGP**. Click **Next**.



__c. Choose ISD_ASSETS as your package choice. Click Next.

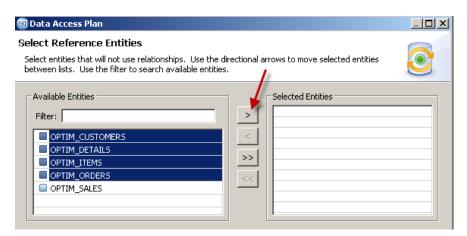


__d. Unlike our source production system, the tables in our target test system do not have any RI defined against them. Choose **Select individual entities**. Click **Next**.

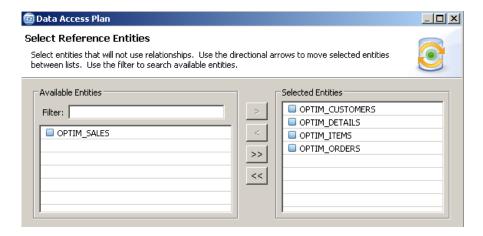


Page 80 Optim PoT Labs

__e. Select the OPTIM_CUSTOMERS, OPTIM_DETAILS, OPTIM_ITEMS, and OPTIM_ORDERS entities by highlighting them and selecting the ">" button individually or hold the SHIFT key and highlight them all together before selecting the ">" button.

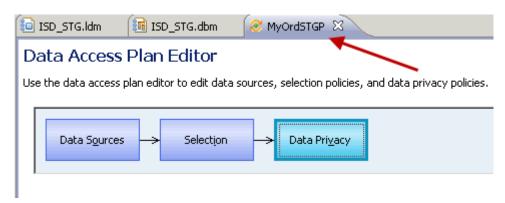


__f. The resulting screen should look like the following screen. Click **Finish**.

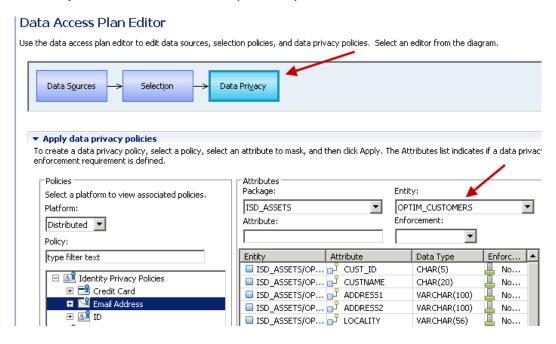


Now we will define the data privacy rules for this Data Access Plan, so that the sensitive data in our non-production environment is masked. The next few steps will be done in the Data Access Plan Editor.

__g. Double-click on the **MyOrdSTG** Tab for this DAP to enable full-screen mode.

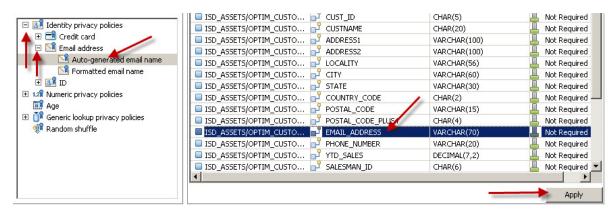


_h. To define the privacy policy, click on the **Data Privacy** box and choose the **OPTIM_CUSTOMER** entity. We will mask the email address in this table. Make sure that you have the **Distributed** platform policies selected.

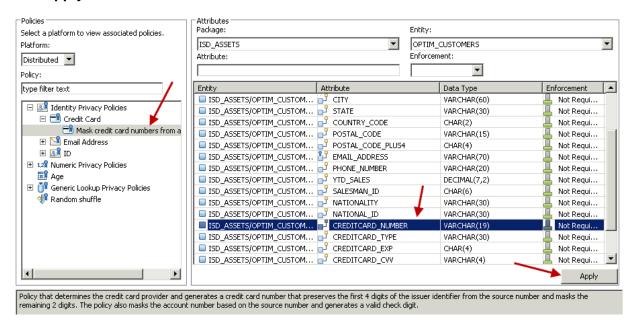


Page 82 Optim PoT Labs

__i. Expand the **Identity privacy policies** and select **Auto-generated email name**. Select the **EMAIL_ADDRESS**" column in the list of attributes. Click **Apply**.



- __j. Click **OK** to confirm the privacy policy has been applied.
- _k. The OPTIM_CUSTOMERS also has credit card information stored in it, so this data should also be masked in the non-production environments. Select Mask credit card numbers from all providers. Select the CREDITCARD_NUMBER attribute. Click Apply.

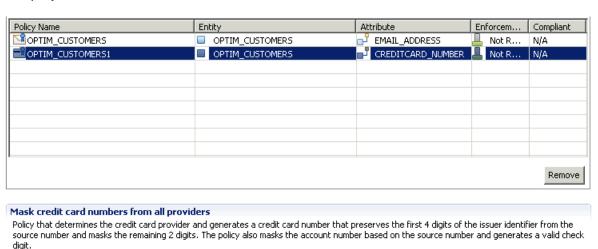


__l. Click **OK** to confirm the privacy policy has been applied.

__m. To check the properties of the masking algorithms that we have selected, **scroll down** within the Data Access Plan Editor to the **Data privacy policies in use** section, as shown below.

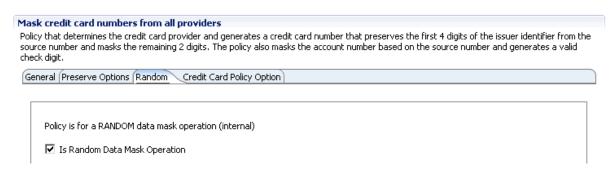
▼ Data privacy policies in use Data privacy polices applied to the data access plan are listed below. Select a policy to view associated properties. You can remove a selected policy from the plan. The list indicates if a policy complies with a data privacy enforcement requirement. Filters Error status: Entity: Attribute: All policy names 🔻 All statuses ▾ Compliance status: Clear All All compliance statuses Enforcem... Compliant Policy Name Entity Attribute ☑OPTIM_CUSTOMERS EMAIL_ADDRESS OPTIM_CUSTOMERS Not R... N/A ➡OPTIM_CUSTOMERS1 OPTIM_CUSTOMERS CREDITCARD_NUMBER Not R... N/A

__n. Select OPTIM_CUSTOMERS1 policy name, which is the credit card policy. Scroll down further to see the details about the selected privacy policy, which are shown in a tabbed display.

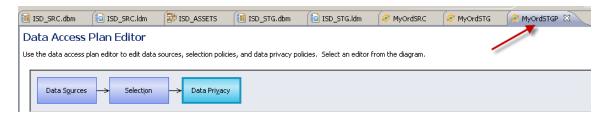


Page 84 Optim PoT Labs

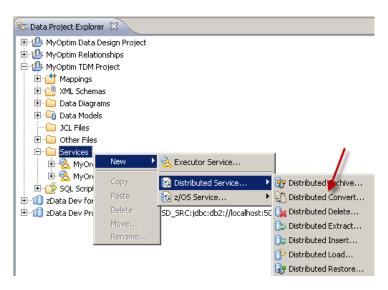
__o. Click on the **Random** tab. Check the option to enable random values.



- __p. Take a few minutes to browse some of the other properties available to the data privacy policies.
- __q. Save this Data Access Plan by either clicking on the **Save icon**, or using **CTRL-S** keys. This completes our work with Data Access Plans.
- r. Double click the **MyOrdSTGP** tab to take the editor out of full screen.



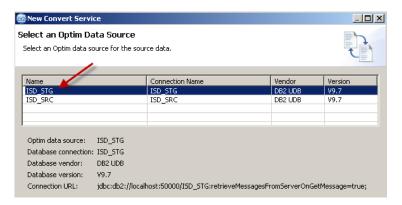
- __3. Optim Convert Service.
 - __a. From the **Data Project Explorer**, drill into the **MyOptim TDM Project**. Right click on the **Services** folder and choose **New** \rightarrow **Distributed Service** \rightarrow **Distributed Convert**.



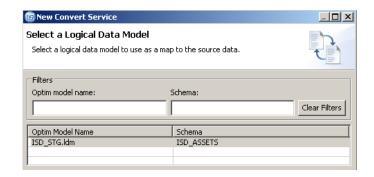
__b. Name the service MyOrdCvt. Click Next.



__c. Select **ISD_STG** for the Optim Data Source. Click **Next**.



d. Take the default. Click Next.

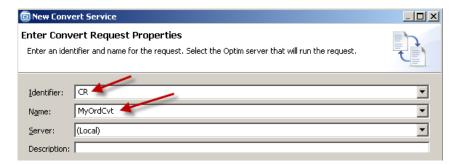


__e. Choose the MyOrdSTGP DAP. Click Next.

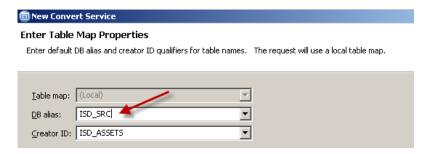


Page 86 Optim PoT Labs

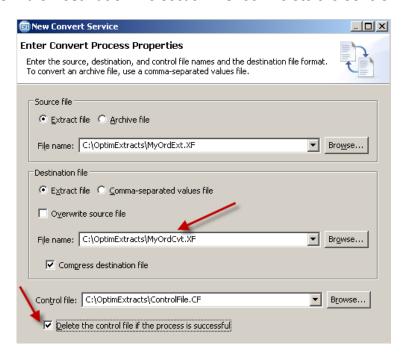
__f. Enter "CR" for the identifier and MyOrdCvt for the name. Click Next.



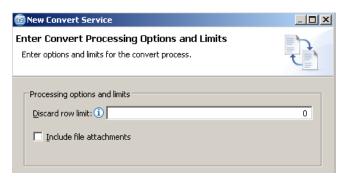
__g. Enter the Table Map properties, specifying "ISD_SRC" and "ISD_ASSETS". Click Next.



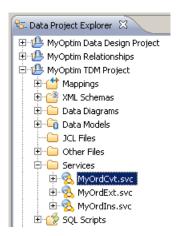
__h. Enter the Convert Request properties. Note that the source file should be the extract file that we created in the previous lab. Click the **Browse** button to locate the extract file. We want to create a new file called "C:\OptimExtract\MyOrdCvt.XF" with our masked values in the **Destination file** section. Check **Delete the control file**. Click **Next.**



__i. Leave the default values on the Convert processing options screen. Click **Finish**.



__j. You will now see your **MyOrdCvt** service under the **Services** folder.

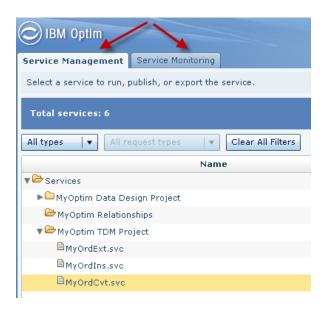


- __4. We will now open Optim Manager from a web browser so that we can run and monitor the newly created service.
 - _a. Open Mozilla Firefox.

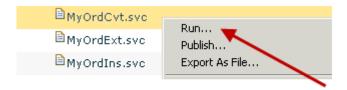


Page 88 Optim PoT Labs

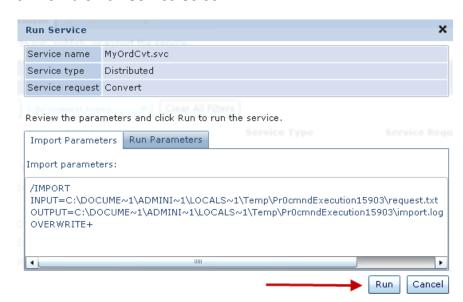
__b. You will be presented with screen with two tabs, one for Service Management and the other for Service Monitoring. You will see your **MyOrdCvt.svc** listed under the **Service Management** tab.



__c. Right click on the **MyOrdCvt.svc** service and select **Run**.



d. Click **Run** from the Run Service Screen.



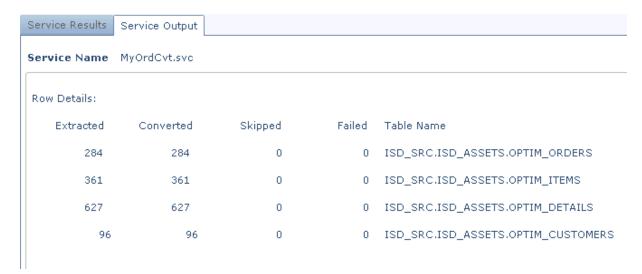
__e. You should receive the following confirmation box, stating that your job has been submitted. Click Close.



__f. Select the **Service Monitoring** tab and you will eventually see that your service completed successfully. Use the **refresh** button (two green arrows on the right side of the screen) to refresh the screen.

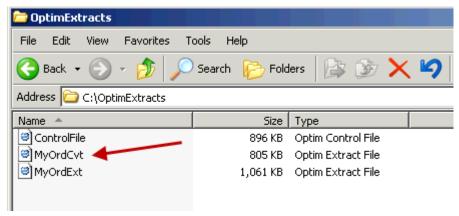


_g. Select the MyOrdCvt.svc service and select the Service Output tab at the bottom of the screen. Scroll down and you will see the number of rows successfully converted into your new extract file.



Page 90 Optim PoT Labs

_h. You may also view the existence of the converted extract file from **Windows Explorer**.



__5. Compare and Browse

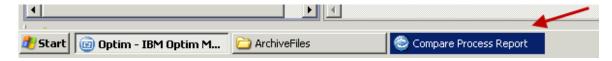
We now have 2 extract files; the first one represents a sub-set of data from the production environment and contains sensitive data. The second extract file contains the same data sub-set, but the sensitive data has been masked. Before loading the masked data into the target database, we need to validate our masking policies. We will do this by using the Compare and Browse utilities of Optim.

_a. Launch the Compare utility by clicking on Run -> External Tools -> 2. Execute Compare.



__b. When the Compare process completes, it will generate a report, which will be flashing at the bottom of your screen, in the Windows Task Bar. You may review the report.

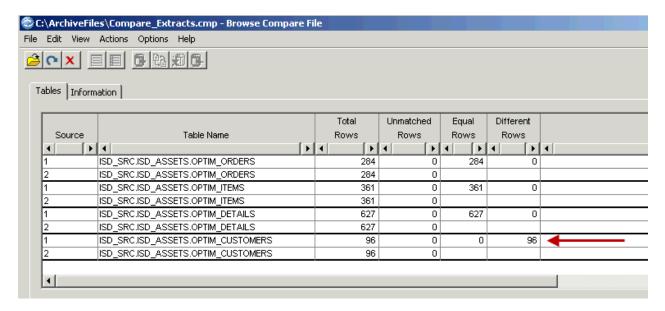
Close the Compare Process Report.



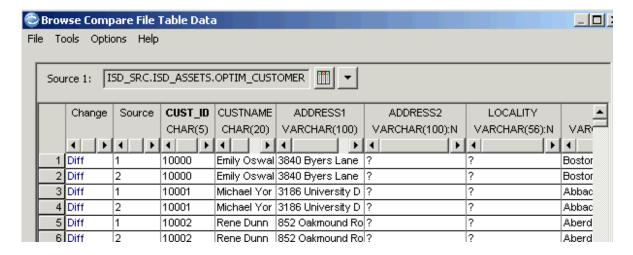
__c. Launch the Browse by clicking on Run -> External Tools -> 1. Browse Compare File.



Notice that the customer table has differences between the two extract files.

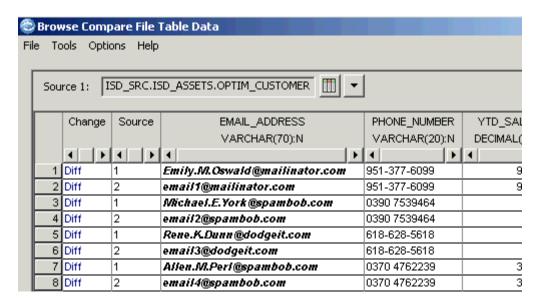


__d. Double-click on the bottom two rows (**OPTIM_CUSTOMERS**) to view the differences between the two Customer tables.

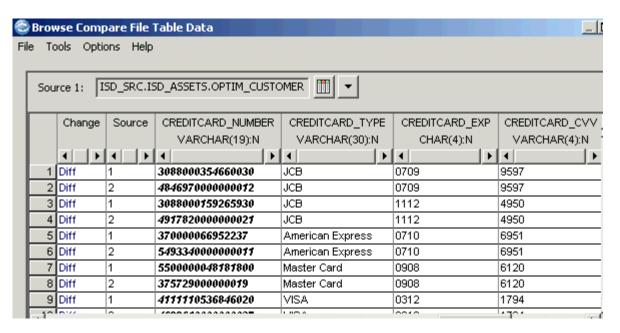


Page 92 Optim PoT Labs

_e. Scroll to the right until you find the Email Address column. Compare the email addresses to confirm that that the data masking policy was properly applied, as shown below.



__f. Continue scrolling until you find the **credit card** columns. Verify that the credit cards were masked.



We now have subsetted, masked data that is ready to be loaded into the non-production environment. You may refer to the Test Data Management lab for a review of the steps involved in loading the masked data into the target database.

CONGRATULATIONS, you have successfully finished the lab.

Lab 6 InfoSphere Discovery Finding Relationships and Complete Business Object

Introduction

In order to deploy an IBM Optim solution successfully, it is critical that we archive business objects; tables that are related to each other. To identify business objects correctly is often a complex task. A typical data set has a large number of tables and may not have well declared or documented keys, foreign or primary. When we work with such a data set, establishing the boundary of tables to be archived together, with the correct relationships between them becomes a challenge. Today we meet this challenge by talking to people, chasing down any documentation we can find. When enterprise memory fails us, we open SQL tools to start querying the data to understand possible relationships. When we do this, it often takes longer than expected before we can start creating Access Definitions and other Optim objects correctly, causing project delays.

Objectives

InfoSphere Discovery will automatically discover implicit relationships in a large schema and also cluster tables into "business objects". In this lab we will only be working with a small set of tables to show how Discovery finds relationships. In this example we will be looking at orders data. We will accomplish the following tasks in this lab to demonstrate how easily Discovery can be used to rapidly create Business Objects.

- 1. Start InfoSphere Discovery, or Discovery, for short.
- 2. Create data sets containing tables from which we wish to create Business Objects.
- 3. Use Discovery to find and review foreign keys.
- 4. Use Discovery to find and review Business Objects, called "Data Objects" within Discovery.

Lab Prep

- 1. Discovery requires a fair amount of memory. Close all the windows and applications from previous labs.
- 2. Start Discovery Server and Engine Services
 - a. Double-click and open the **Services** desktop icon.



Page 94 Optim PoT Labs

_b. Select the IBM InfoSphere Discovery Engine Service. Right-click and Start the service. Repeat this for the IBM InfoSphere Discovery Server Service. Close the services windows.

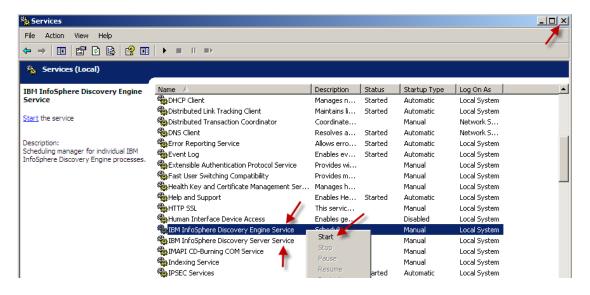
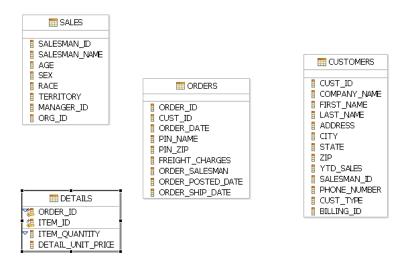


Table Structure

Below is a diagram of the tables that exist in our database. The tables contain order and customer related data but does not have any relationships defined within the database. We will use Discovery to find the relationships.



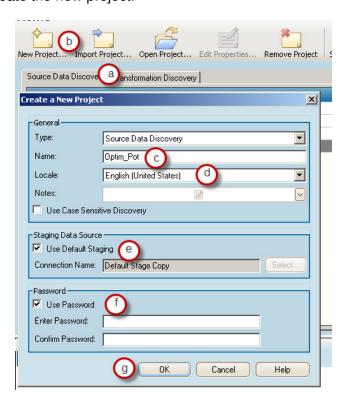
Lab Steps

- __1. On your desktop, double-click the **Discovery Studio** icon. If you get a connection error, wait a minute and try again.
- Discovery Studio

- 2. **Maximize the window** to full screen.
- __3. Create the Discovery project.

You will see 2 tabs, representing 2 types of discovery projects: **Transformation Discovery** and **Source Data Discovery**. In this lab, we will be doing a Source Data Discovery project. Create project Optim Pot as shown below.

- __a. Click the **Source Data Discovery** tab.
- __b. Click **New Project** to create a new project.
- __c. Name the project "Optim_Pot".
- __d. Leave the Locale drop down as **English (United States)**.
- e. Leave the **Use Default Staging** checkbox selected to use the Default Stage Copy.
- f. Uncheck the **Use Password** box to not require a password for this project.
- __g. Click **OK** to create the new project.



Page 96 Optim PoT Labs

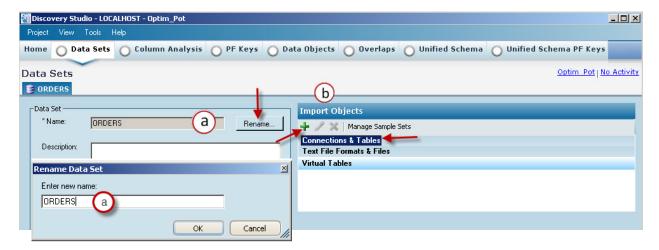
4. Create a connection to data source.

Upon creating a new project you will be taken to the Data Set tab. You will now define the data to be used for your project. If you are opening an existing project, you only need to click the Data Sets tab.

Once at this window you will create a data set.

- a. Click the **Rename** button and rename this Data Set "**ORDERS**". Click **OK**.
- __b. Under Import Objects, click on the **Connections & Tables** folder to highlight it. **Click on the icon** to open a window titled **Create Connection**.

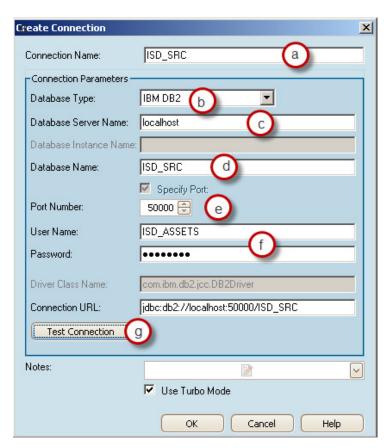
Now you can connect to one or more relational databases or add text files into the ORDERS data set. In this lab, the Orders information exists in a DB2 database. The next lab step will guide you through the entries required to connect to a database called ISD_SRC.



__5. Connect by providing Database Connectivity information to Discovery.

Enter the following information to create the Connection:

- __a. Connection Name. This is free text. Call this connection "ISD_SRC" (this is also the DB2 database name).
- __b. Select the Database Type = **IBM DB2**.
- c. Database Server Name = "localhost".
- __d. Database Name = "ISD_SRC".
- e. Port Number = "**50000**".
- __f. Username and Password = "ISD_ASSETS/passw0rd" (note that we use a zero in place of the 'o' in passw0rd).
- __g. Click the **Test Connection** button. Click **OK** after it is successful.
- h. Once the connect works click **OK** to save the connection.



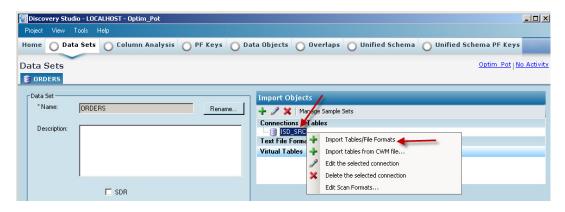
Note: By creating more than one connection it is possible to discover relationships between schemas or even between different databases.

Page 98 Optim PoT Labs

__6. Import ISD_SRC tables for analysis.

Once you successfully connected, you need to "import" tables. You will have the option of importing all tables from this connection, or selectively importing the tables you need. We will only be selecting four tables for this lab.

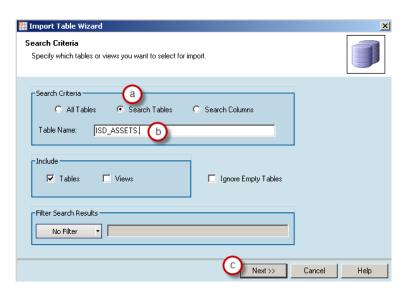
- __a. Right click the connection ISD_SRC.
- __b. Click Import Tables/File Formats.



__7. Search for tables owned by ISD_ASSETS.

If you know something is common for all the relevant tables, e.g. common prefix, common user name, etc. you can search for these tables to import them. In our case, we search for all tables owned by ISD_ASSETS.

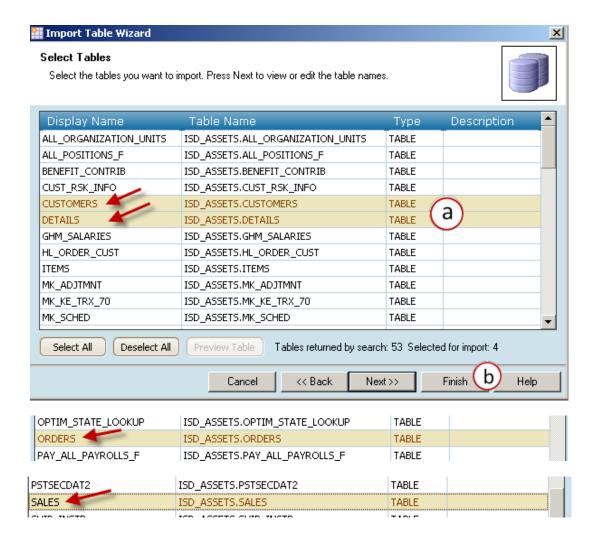
- a. Click the **Search Tables** radio button.
- b. Enter "ISD ASSETS." (NOTE: don't forget to put a period at the end).
- c. Click the **Next** button.



__8. Import the ORDERS, CUSTOMERS, DETAILS, and SALES tables into Discovery.

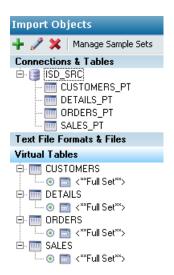
We can selectively import all tables into Discovery.

- __a. Use the **Control** key to hold and left click on the **CUSTOMERS** and **DETAILS** tables. **Scroll down** and click on the **ORDERS** and **SALES** tables.
- b. Click Finish.



Page 100 Optim PoT Labs

Note: Virtual tables vs. Physical Tables. After import, you will see that for each table imported, Discovery displays both a physical table and a virtual table, as shown above.. The virtual tables are identical to the physical tables. The rest of the analysis is performed on virtual tables. In Discovery product training, you will learn how to use virtual tables to perform more advanced analysis. For the purpose of this lab, do not make changes to virtual tables; just remember they are exactly the same as the actual tables from the source database.



__9. Run the Discovery analysis.

We are now ready to analyze these tables.

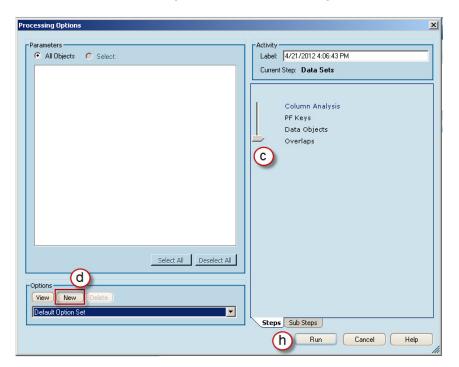
__a. If you are not in the Data Sets tab, click the **Data Sets** tab to return to the first step in the Discovery analysis methodology.



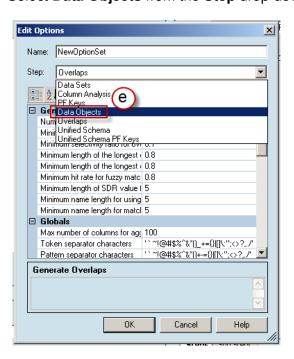
b. Click the **Run Next Steps** button found on the bottom right of the window.



- __c. Click and drag the slider bar down to Overlaps.
- __d. We will make some changes to the default settings. Click the button labeled **New**.

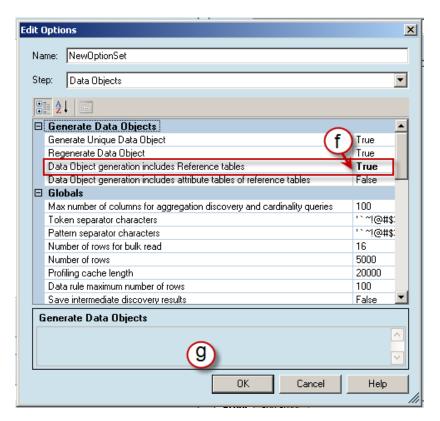


_e. Select **Data Objects** from the **Step** drop-down list



Page 102 Optim PoT Labs

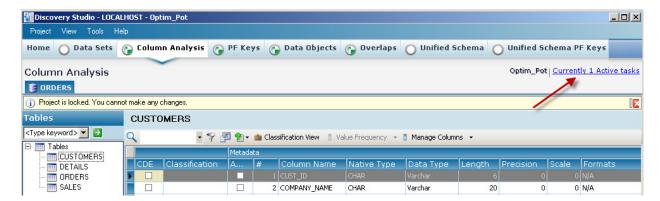
- __f. Change Data Object generation includes reference tables from False to True.
- __g. Click OK.



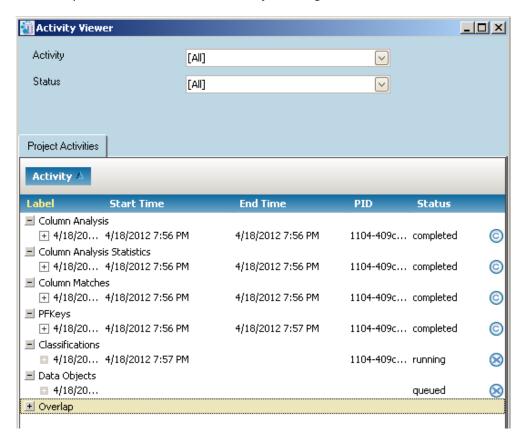
__h. Click **Run** (see Figure below)

_10. Discover relationships and data objects.

After the task is submitted, you can use the Activity Viewer to monitor its progress. When the task indicator in the upper right corner of the Studio shows "No Activity", it means processing is completed. While the processing is running, the project will be locked; you can still browse, but you won't be able to modify anything.



If you click the Activity Viewer, the following window will appear. This will allow you to see which steps have completed and which are currently running.

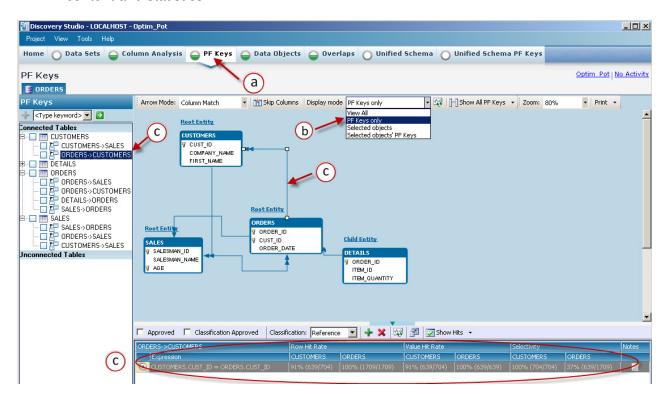


Page 104 Optim PoT Labs

11. Review discovered relationships.

With a large schema, the relationships could be numerous and complex. It is important for an analyst to focus on the purpose for analysis, in this case we've only done a simple analysis to show that Discovery finds valid relationships. IBM does offer a session that focuses solely on Discovery. We are only showing a fraction of its function in this lab. We are only using Discovery to find relationships on relational tables.

- __a. Select the **PF Keys** tab.
- __b. Select the **PF Keys Display mode**.
- _c. Review all relationships around the Orders table by selecting that relationship either from the diagram or from the tree view. Once a relationship is selected, you can see its content and statistics.



Note: Examining a discovered PF key. Each discovered PF key is presented on the graphical user interface with statistical properties as well as hit/miss data views. In this lab, we are looking for perfect keys, so these facilities are not as useful as when we examine an imperfect key and try to determine whether it is a real relationship or an accidental one. Discovery can discover perfect keys as well as "almost keys". Ask your instructor about an "under the hood" session for working with "almost keys".

12. Review data objects.

Review the Data Objects for completeness. It is possible to add or change the data object as outlined later in this lab:

- __a. Select the **Data Objects** tab.
- b. Select the Data Object **DO ORDERS**.
- __c. Review the selected relationship.

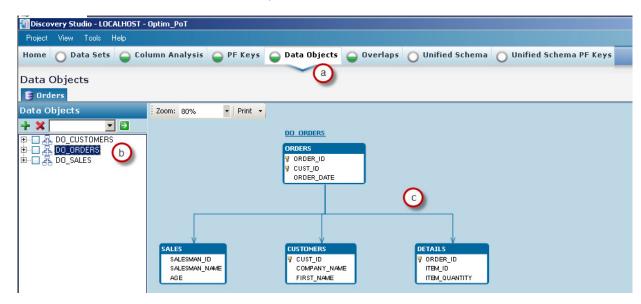


Figure 11. Review discovered business objects.

Note: How we find these Data Objects in Discovery? The tool discovers business objects by analyzing overall relationships. Depending on how deep and wide an Access Definition should be, sometimes you may need to expand or shrink the boundaries of the business objects. Even in those cases, Discovery provides a critical starting point for you to work with business objects. While we cannot go to that depth during a PoT, we cover these techniques more extensively in our product training.

13. Export artifacts consumable by Optim.

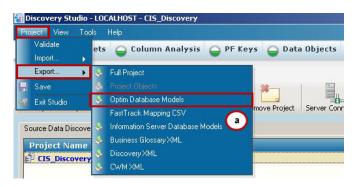
After reviewing relevant data objects, in this case the Orders data object, you have all you need to generate archiving or test data management "code", the access definitions and associated objects. This export will be consumable by Optim Designer. The question is, how do we get these data objects to Optim Designer so it can do the code generation for me? It works as follows:

Page 106 Optim PoT Labs

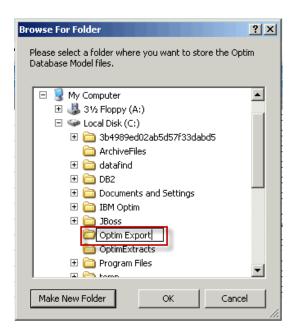
From Discovery, an analyst uses the Export to Optim feature to generate a set of artifacts as an XML file. In particular, Discovery generates one Physical Data Model (PDM) file for each Discovery data set, and one Logical Data Model (LDM) file for each Data Object from Discovery.

Optim Designer will be able to read the generated files and turn them into access definitions. We will not walk through the Optim import in this lab.

__a. Click on Project > Export > Optim Database Models.



- b. Click on Make New Folder.
- __c. Name it 'Optim Export'.
- d. Click OK.



__e. The files generated are seen in the Discovery Workspace Directory. Go to that directory to examine the XML files.

Congratulations, you have completed this lab.



© Copyright IBM Corporation 2012.

The information contained in these materials is provided for informational purposes only, and is provided AS IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, these materials. Nothing contained in these materials is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software. References in these materials to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates. This information is based on current IBM product plans and strategy, which are subject to change by IBM without notice. Product release dates and/or capabilities referenced in these materials may change at any time at IBM's sole discretion based on market opportunities or other factors, and are not intended to be a commitment to future product or feature availability in any way.

IBM, the IBM logo and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or TM), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at ibm.com/legal/copytrade.shtml

Other company, product and service names may be trademarks or service marks of others.



Please Recycle