In ETL/Data Warehouse, you will encounter different sources and targets.

**Source Definition in Informatica**

A **Source Definition** in Informatica is an entity from where you pull the records, and then you store these records in temporary tables (staging tables) or Informatica transformation caches. On the basis of mapping design document/business requirement you make a change in these data records (transform the data) and then you load the transformed data in another tables structures, called target tables.

In every Informatica mapping, there will always be a source and a target. To manage different sources and targets in Informatica, you have to use source analyzer and target designer. These tools are integrated into Powercenter designer tool and can be launched from there itself.

**Source Analyzer in Informatica**

**Source Analyzer** in Informatica is a tool that allows you to create and modify source definitions. With the help of Source analyzer, you can create or import different types of sources in Informatica like flat files, relational databases, Excel worksheets, XML definitions, etc.

Similarly by using target designer, you can create or import different types of targets. Informatica provides you the feature to either create a source/target from scratch by designing its structure, or you can create a source/target by importing the definitions. When you Import Source from the [Database](https://www.guru99.com/introduction-to-database-sql.html), the following metadata gets imported:

* Source (Table) name
* Database location
* Name of Columns
* Columns Data types
* Constraints

You can also define the key relationships in the tables, which is applicable only at Informatica level and is stored in the repository.

The source or target created/imported in Informatica can be reused any no of times in different mappings. Every mapping must have at least on loadable target. Otherwise mapping will be invalid.

**Note**– When we create source/target in source analyzer/target designer, structures are only created in Informatica. At the database level, there is no object created. So, you have to create Database objects having the same structure as you have created in Informatica.

You can import following type of sources using source analyzer

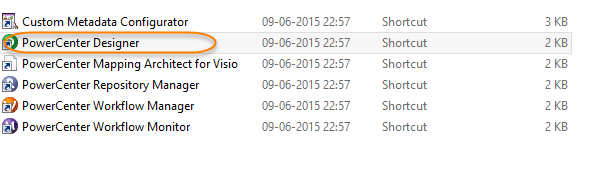
* Relation tables ( database tables), views and synonyms
* Flat files
* Cobol files
* XML files

**Performance tip** – To improve the performance of Relational Source tables, use indexes on the source database tables. On the target, tables disable or remove constraints and indexes for performance.

**How to Open Source Analyzer in Informatica**

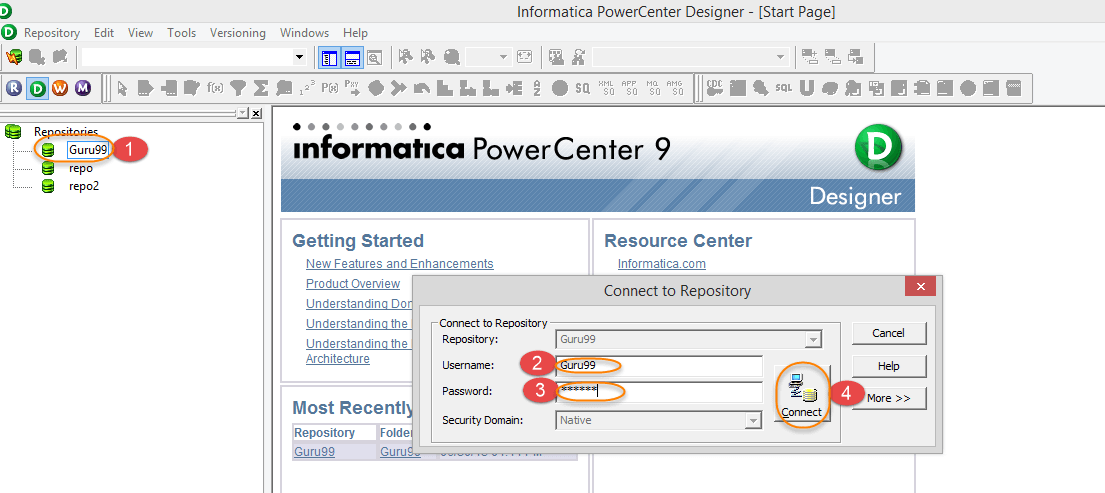
Below is a step by step process for opening Source analyzer in Informatica:

**Step 1)** Open Informatica PowerCenter designer tool



**Step 2)** In next screen

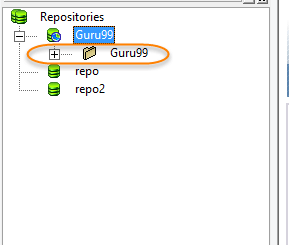
1. Double click on the repository which you want to connect
2. Enter username
3. Enter Password
4. Click on Connect button



After Successful login, the folders of the user will be listed under the repository name

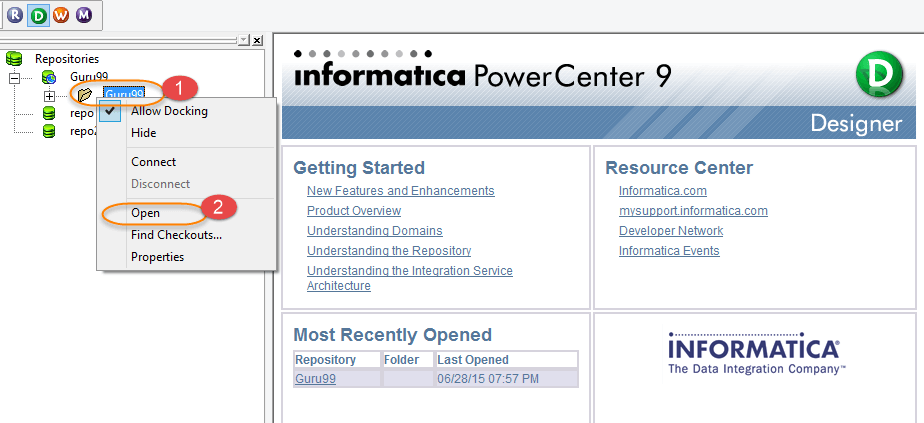
**Note –**

1. If login is not successful, check if the user has privileges of connecting to repository
2. If no folder is visible under repository name, check if the folder is created.
3. If a folder was created recently then disconnect from the repository and reconnect.

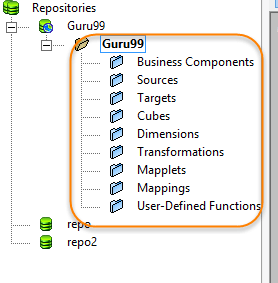


**Step 3)** In the next step

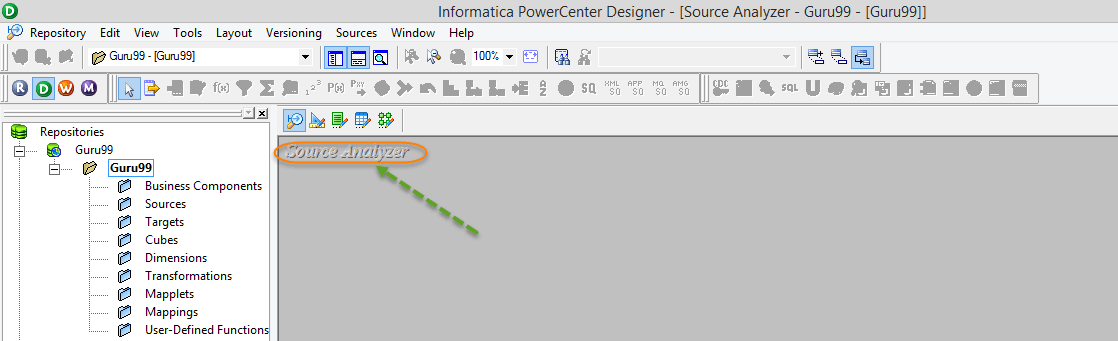
1. Right click on the folder
2. Select open option



When a folder is open, subfolders of that folder will be listed.



**Step 4)** Click on the Source analyzer menu, as shown in the figure.



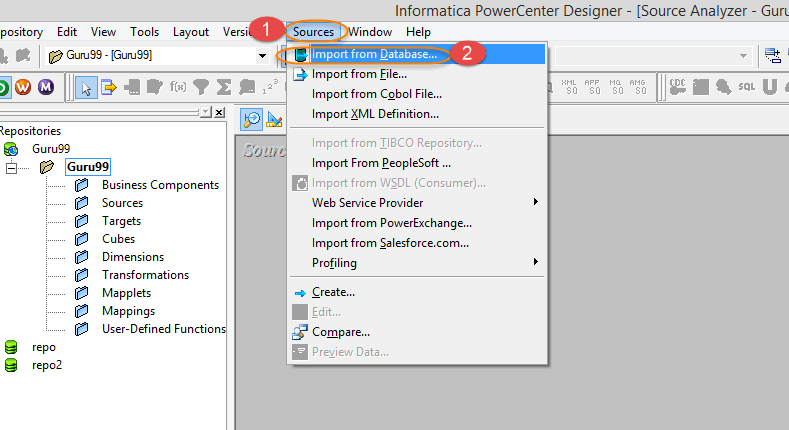
**How to Import Source Table in Source Analyzer**

Following are the steps to import source table in Informatica Source Analyzer:

**Step 1) Go to “Sources” option**

In source analyzer

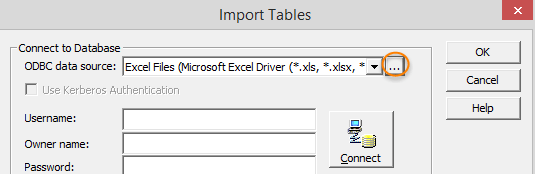
1. Click on tab “Sources” from the main menu
2. Select import from database option, after this ODBC Connection box will open.



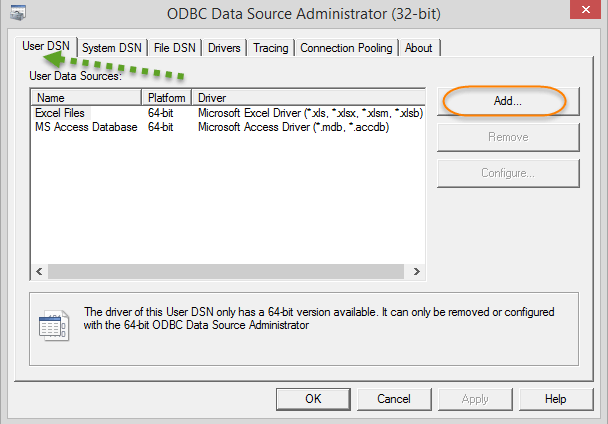
**Step 2) Create ODBC connection**

We will now create ODBC connection (If you already have ODBC connection, move to step 3)

1. Click on the button next to ODBC data Source.

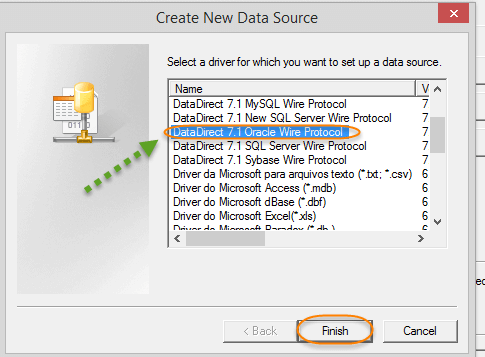


1. On the next page, Select user DSN tab and click Add button.



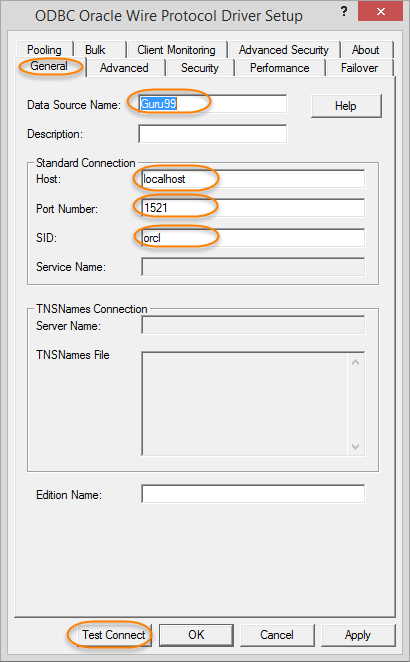
A DSN (Data Source Name) is basically a data structure, and it contains information about a specific details (database type, location, user details, etc.). This information is required by ODBC driver so that it can connect to that database.

1. When you click on the add button, you will see a list of drivers for various databases (Oracle, SQL Server, Sybase, Microsoft Access, Excel, etc.) The driver which you select depends on what database you want to connect. Select the driver for the data source. Here in guru99 you will be using oracle database tables. Select Oracle wire protocol driver.

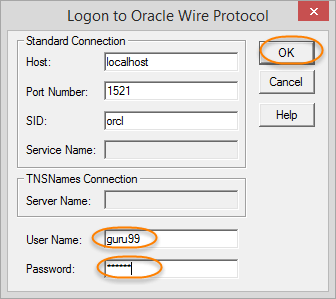


1. On the next page, select the general tab and enter database details. Then Click on test connect.

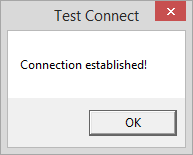
**Note** – Database details are specific to your installation depending on hostname, port and SID you selected during database installation. You can get this details by opening **tnsnames.ora file** located in your oracle installation folder.



1. Once you made a Test Connect, enter database username, password and then select “OK” button



1. If the connection is successful, it will show a message box.



1. Select OK for test connect window and ODBC Driver setup window. The Created ODBC Data Source will be added to user data sources. Now we are set with ODBC Data source.

**Step 3) Create Schema tables**

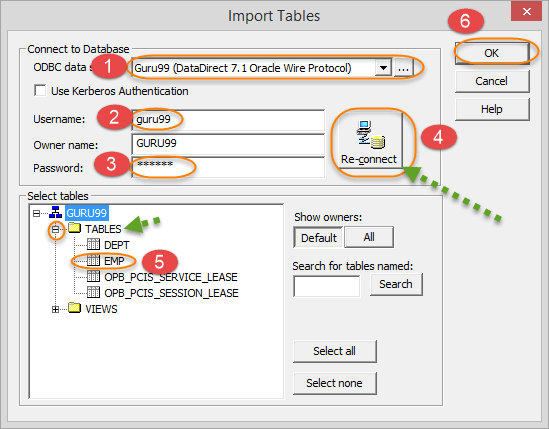
In guru99, you will be using oracle’s Scott/Tiger schema tables. If you do not have these tables in the database, create those using this script

[**Download the above Scott.sql File**](https://drive.google.com/uc?export=download&id=0ByI5-ZLwpo25TDBzTExTdEQzSzQ)

**Step 4) Enter the Database details**

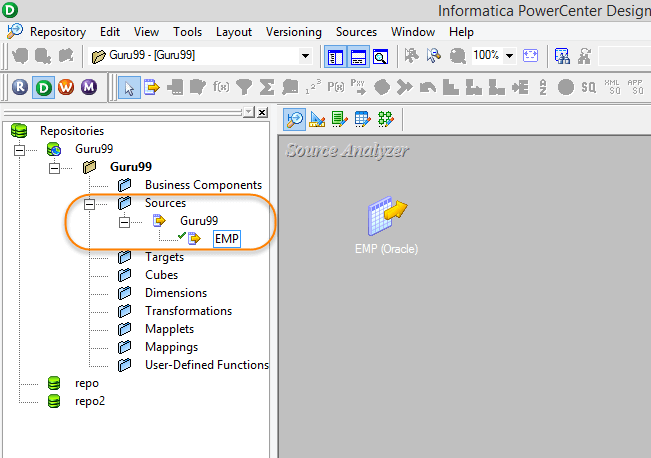
In the import tables window:

1. Select ODBC data source for the oracle database.
2. Enter database username
3. Enter database password
4. Click on connect/reconnect button. This will show tables for the database user.
5. Expand the tree under tables folder and select EMP table
6. Select OK button.



**Step 5) Verify and save the imported table**

The table will be imported in the Informatica Source Analyzer. Use “Ctrl+S” keys to save changes to the repository



Similarly, you can import any other database tables in source analyzer.

**How to Import Target in Informatica Target Designer**

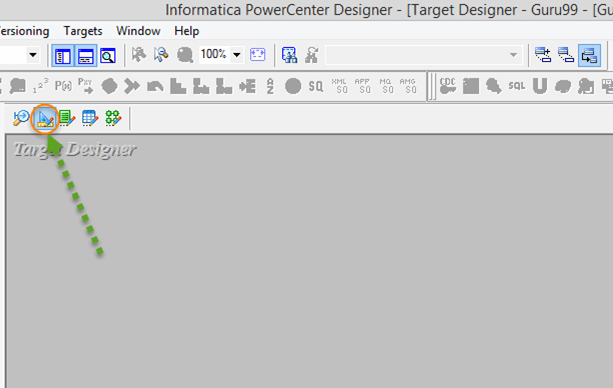
Here is a process to import target in Informatica Target Designer:

In this tutorial, you will create target table emp\_target, which will be having the identical structure of the source EMP table. To import target table, the target table must be present in the database. To create target table use the below script. Types of targets available in Informatica are relational, XML and flat files.

[**Download the above emp\_target.sql File**](https://drive.google.com/uc?export=download&id=0ByI5-ZLwpo25RFZFaVpsUGM1M1U)

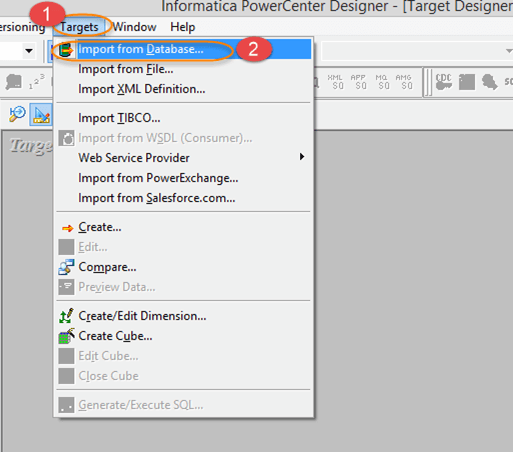
**Note** – here in this tutorial, both source and target table are present in the same database schema (guru99). However, you can also create a target in a different schema.

**Step 1)** In the Informatica Designer, click target designer icon to launch target, designer.



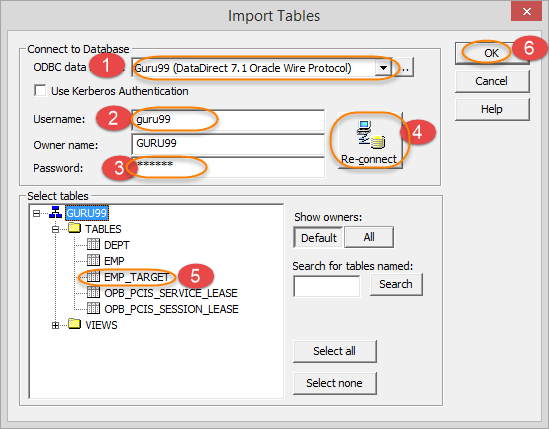
**Step 2)** In next step

1. Select option “Targets” from the main menu
2. Select “import” from database option

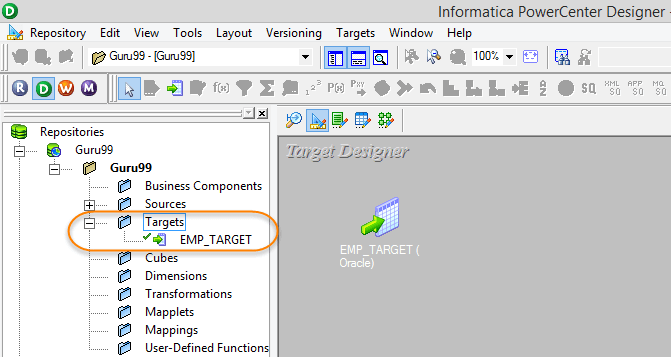


**Step 3)** In the import table window.

1. Select ODBC data source for the oracle database.
2. Enter database username
3. Enter database password
4. Click on connect/reconnect button. This will show tables for the database user.
5. Expand the tree under tables folder and select EMP\_TARGET table
6. Select OK button



This will import target in Target Designer of Informatica Powercenter. Use “ctrl+s” key to save the changes.



In this tutorial, you have imported sources and targets. Now, you are all set to create your first mapping.

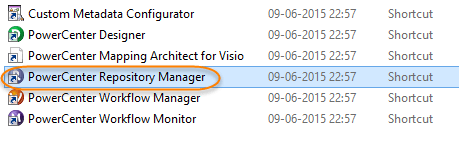
**How to Create a Folder in Informatica**

Below are the steps on how to create a folder in Informatica:

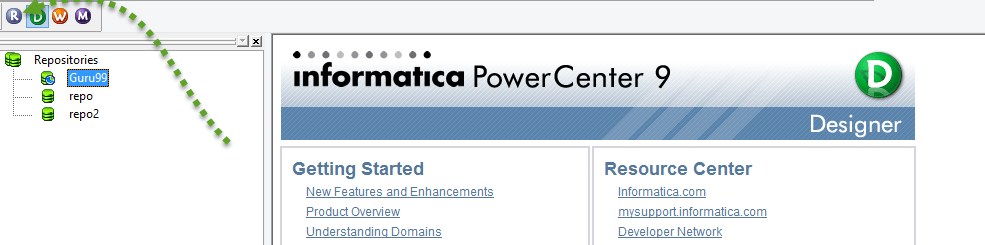
Informatica folders are logical containers, which keeps the objects that you create in Informatica designer/ Workflow manager (mappings, mapplets, sources, targets, etc.)

To create a folder:

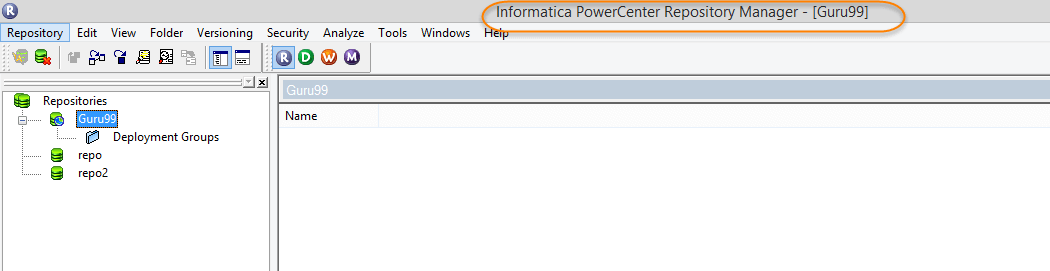
**Step 1)** Open Repository Manager



**Note** – If Informatica Designer is already open, then repository manager can be opened using the shortcuts available in the toolbox.

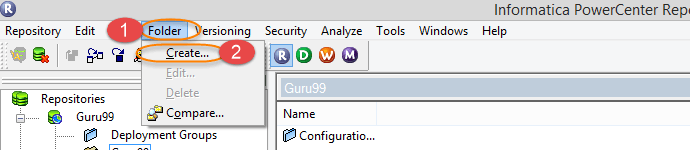


This will open Informatica Repository Manger in a separate window.



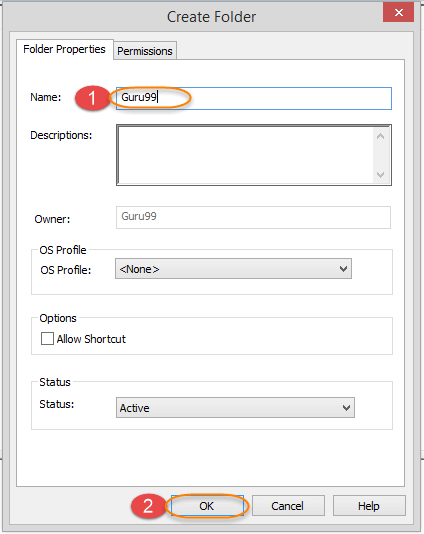
**Step 2)** Follow these steps

1. Click on folder menu
2. Select Create option

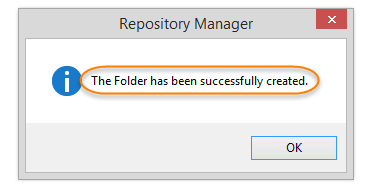


**Step 3)** In next step

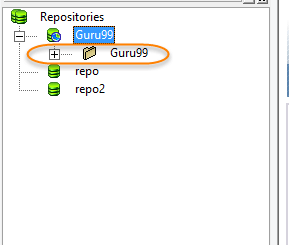
1. Enter Desired folder name
2. Select ok



This will create a new folder and a message will appear saying “The folder has been successfully created.”



The newly created folder will be listed under the respective repository.



This completes the tutorial. See you in the next one!

**What is a Mapping?**

Mapping is a collection of source and target objects linked together by a set of transformations. These transformations consist of a set of rules, which define the data flow and how the data is loaded into the targets.

A mapping consists of following set of objects

* **Source Definition** – Source definition defines the structure and characteristic of the source, its underlying data types, type of the data source, etc.
* **Transformation** – Transformation objects define how the source data is transformed, and various functions can be applied during the process.
* **Target Definition** – Target definition defines the final target where the data will be loaded.
* **Links** – Links connect the source definition to different transformations and target tables. It defines how the data flows from source to target and the transformations.

**Why do you need Mapping?**

Mapping is an **object** in **Informatica** with the help of which you can define how the source data is modified before it reaches the destination or target object. Like if you have employee name as “Bill Clinton” in your source system and in the target system the requirement is to have employee name in the format as “Clinton Bill”, such operations can be designed at the mapping level. In basic terms, what you do with the source data is defined at the mapping level.

Mapping is the basic Informatica object with the help of which we can define the data transformation details and source/target object characteristics. Mappings help us to define the data transformation at the individual column levels for each row. Even in a single mapping you can handle multiple sources and targets.

**Components of Mapping**

Basic components of a mapping are

* **Source tables**
* **Mapping parameters and variables**
* **Target objects**
* **Mapping transformations**

There are various objects that constitute a mapping. A mapping can consist **of sources, targets, mapping parameter and variables, mapplets, various transformations, and user-defined functions**.

* **Mapping Source**: Mapping sources are the objects from where you fetch the source data. It can be a database table, flat file, XML source or COBOL file source
* **Mapping target:** Mapping target is our destination objects where final processed data gets loaded. Mapping target can be a relational table of a database, a flat file or XML file. Sources and targets are mandatory in any mapping, their type can differ
* **Mapping Parameters and Variables**: Mapping parameters and variables helps you to create temporary variable objects which will help you to define and store temporary values while mapping data processing. Mapping parameters and variables are optional users defined data types, which can be created for a mapping and can be referenced and updated for a specific requirement. We will learn more about mapping parameters and variables in this section
* **Mapplets**: They are objects which consist of a set of transformation, source or targets. Mapplets are generally created to reuse the existing functionality of a set of transformations. It can be used in any no of mappings.

**What is Stage Mapping?**

A stage mapping is a mapping in where we create the replica of the source table. For Example, in a production system if you have an “employee” table then you can create an identical table “employee\_stage” in ETL schema.

Having a local stage table offers various advantages, like production downtime, won’t affect your ETL system because you have your own “employee\_stage” table, instead of referring to production “employee” table. In a Production system, there can be other operations and processes which affect the performance. However, when you have replica staging table, only ETL processes will access it. This offers performance benefits.

In Stage Mappings,

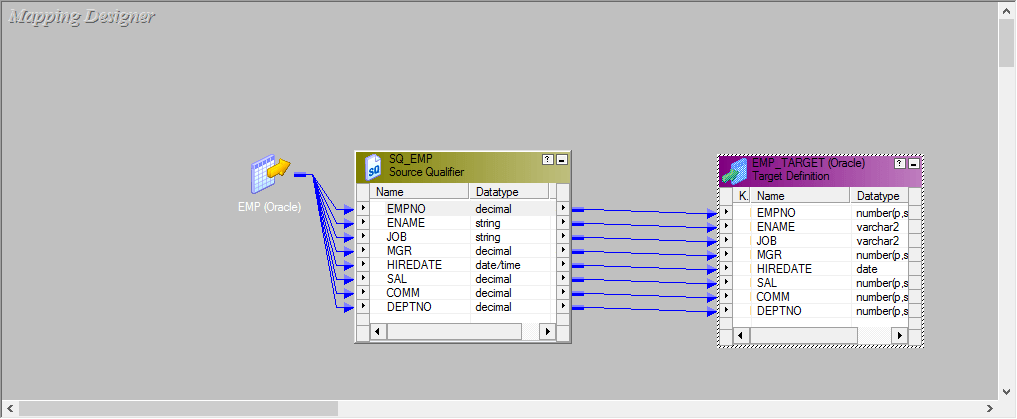
* Source and Target tables have identical structures
* The data in the target table is a replica of source table data or
* Data in stage (target) table is a subset of source data.

For example, if your source table contains employee details of deptno 10, 20, 30, and 40. The staging table can be a table having employee records of deptno 10 & 30 only.

The purpose of creating stage tables in Data warehouse is to make the process of data transformation efficient by fetching only those data which is relevant to us and also to minimize the dependency of ETL/Data Warehouse from the real-time operational system.

**How to Create a Mapping**

In this exercise, we will create a stage mapping, in which source will be “emp table” and the target will be “emptgt”.

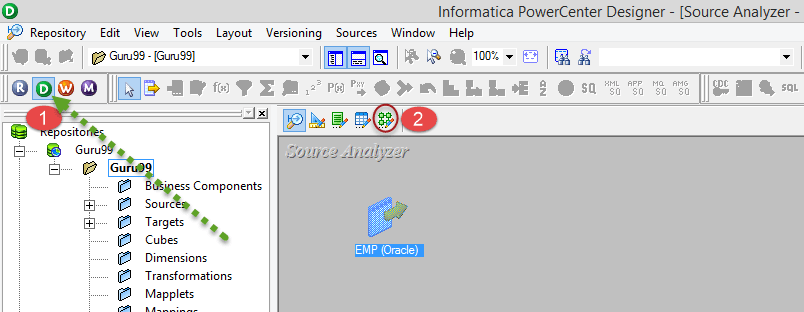


**Naming Convention** – mapping names are prefixed with ‘m\_’ followed by source and target table names separated by underscore sign.

Example – if you are loading emp\_target table from the emp table, then mapping name can be ‘m\_emp\_emp\_target’.

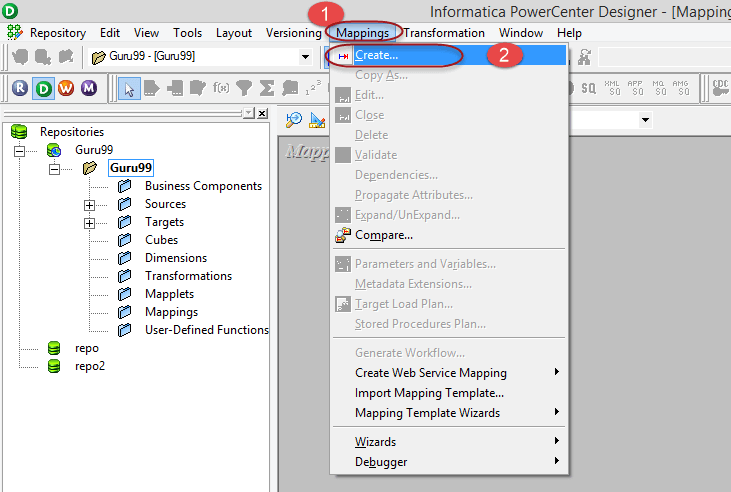
**Step 1** – Launching Mapping Designer

1. Open Informatica Designer Tool
2. Click on Mapping Designer Icon to launch Mapping Designer

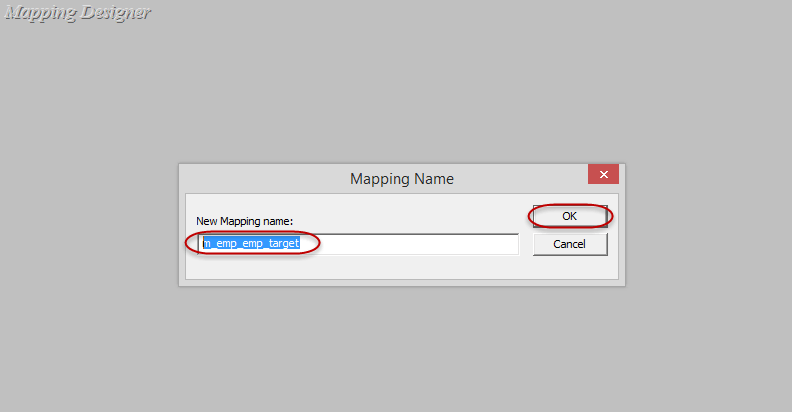


**Step 2** – In Mapping Designer

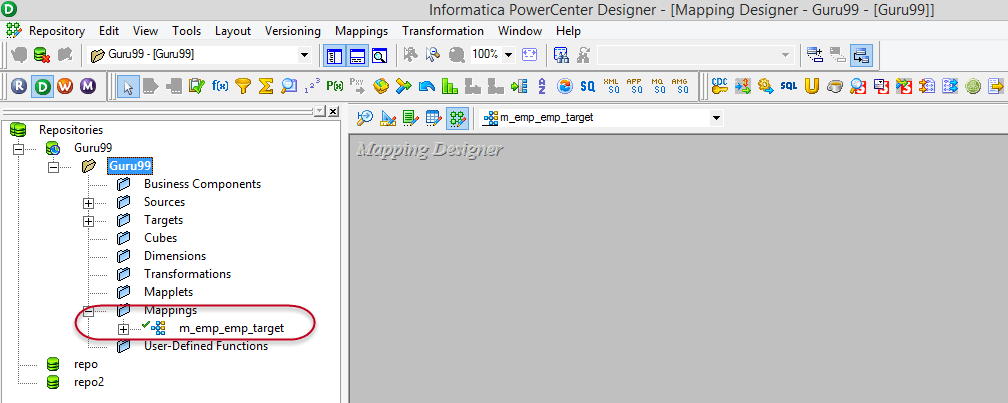
1. Click on Mappings Menu
2. Select Create Option



**Step 3** – Enter Mapping name as ‘m\_emp\_emp\_target’ and select OK Button.



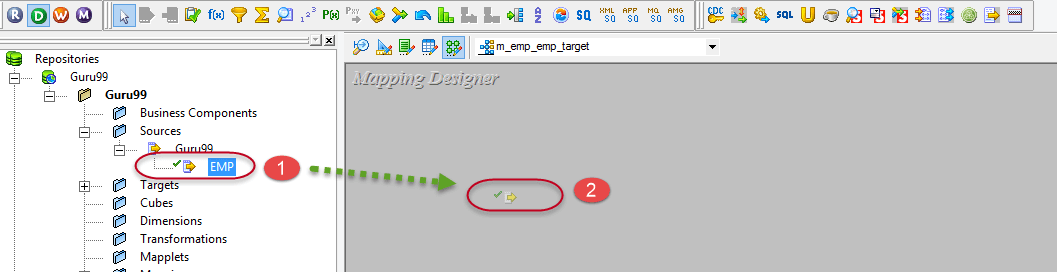
Mapping will be created and listed under mappings folder.



A Mapping must have at least a **source and a target**, you will add sources and targets to the mapping.

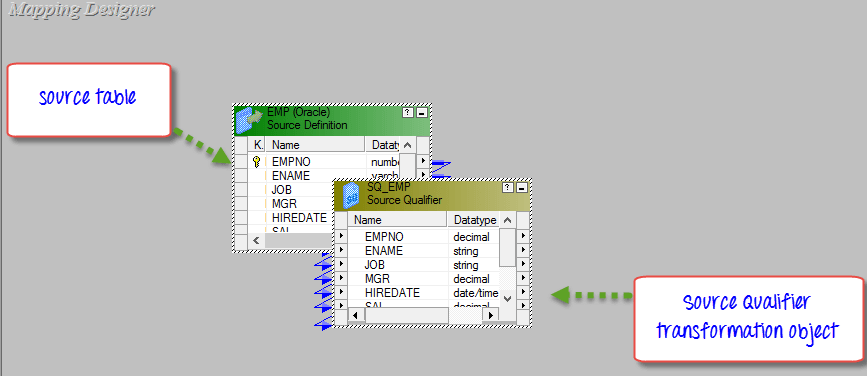
**Step 4** – In this step we will,

1. Select “emp” source table under sources folder.
2. Drag and drop “emp” table into mapping designer.



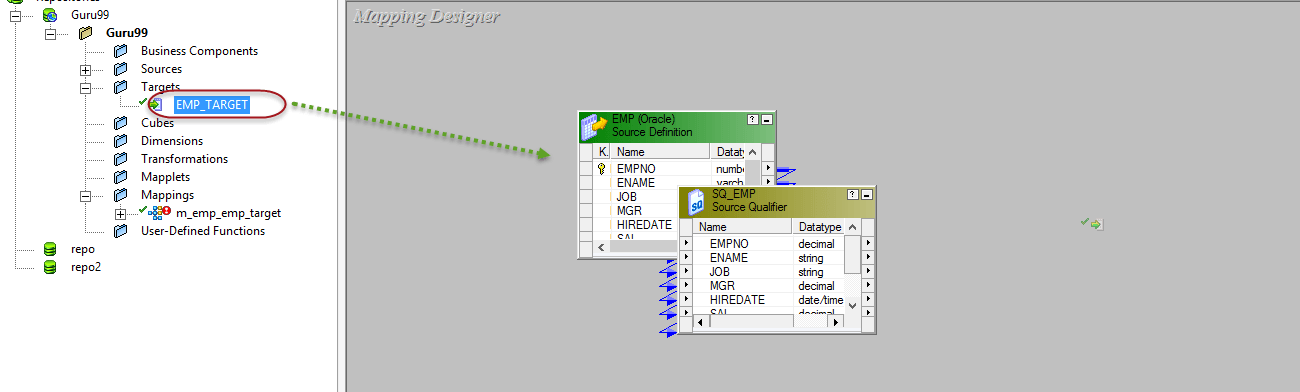
In mapping designer, imported source table will be shown.

**Note** – When you import any relational (database) table in a mapping, an additional object of source qualifier type will also be created. This source qualifier transformation is necessary and helps Informatica integration service to identify the source database table and its properties. Whenever you import a source table, source qualifier transformation will also be created. You should never delete a source qualifier object in a mapping.

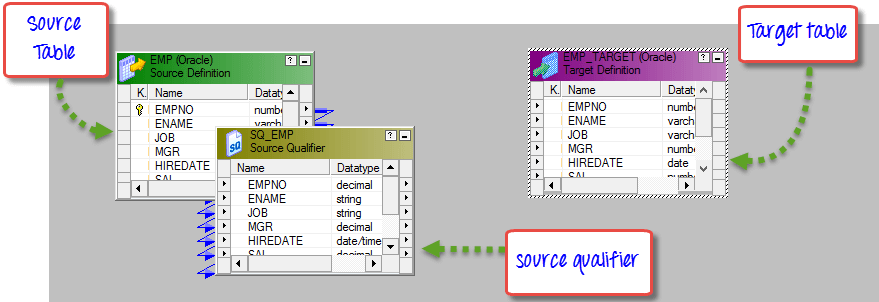


**Step 5** – In this step we will,

1. Select “emp\_target” source table under Targets folder.
2. Drag and drop “emp\_target” table into mapping designer

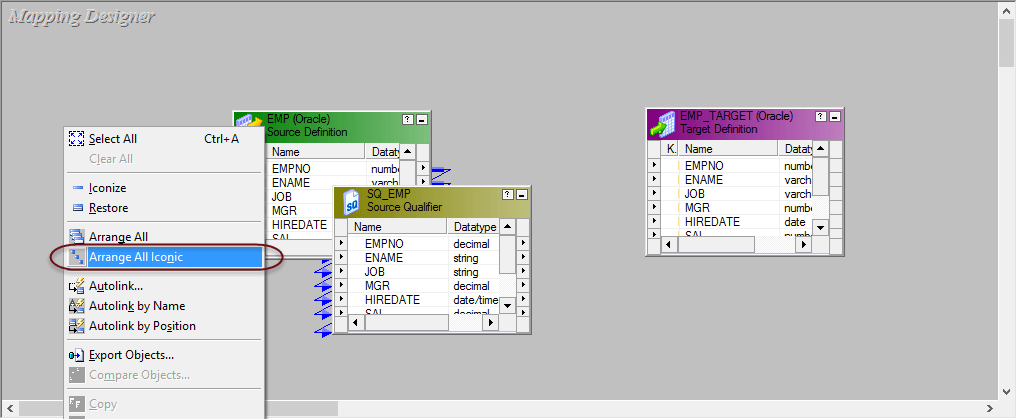


In mapping designer, “target table” will be imported and shown.

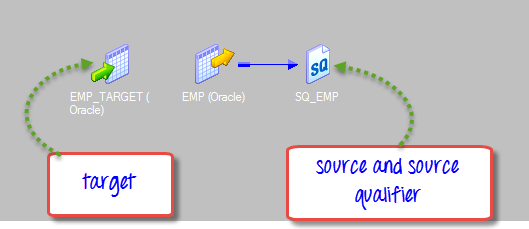


To manage the view space, you can iconize these objects in the mapping.

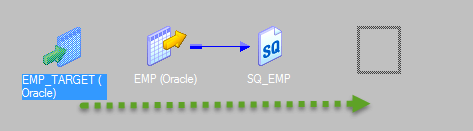
**Step 6 –**Right click anywhere in the mapping designer empty workspace and select option – Arrange all iconic.



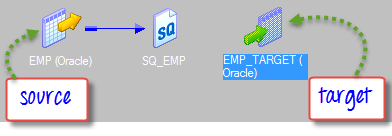
After selecting the option “Arrange all Iconic”, the workspace will look like this.



**Step 7** – In Informatica, We design with the flow from left to right. So, source tables should be at the left side, and target tables should be at right. To arrange tables in our workspace, Select the “emp\_target” table then drag and drop it to the right side of the emp table.



After this rearrangement, the workspace will look like this.

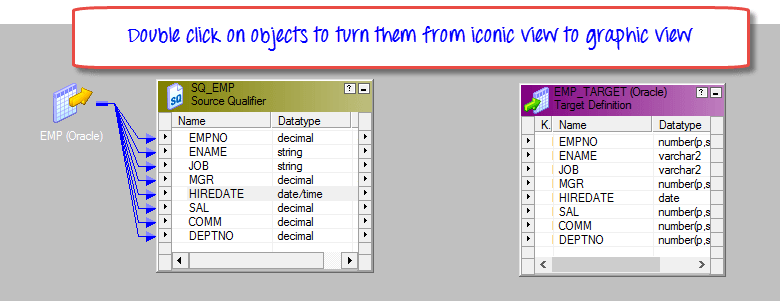


**Note** – Periodically use “ctrl+s” shortcut to save changes to the repository.

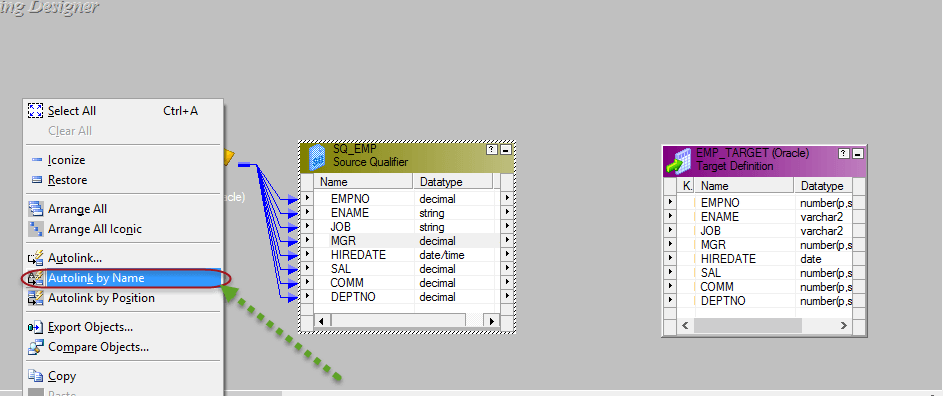
**Step 8 –**Now you have source and target tables in your mapping, but the mapping is not yet complete. The source and target tables should be linked to complete a mapping.

To Link source and targets

Double click on SQ\_EMP table and EMP\_TARGET table to change the view of them from iconic to the graphic.

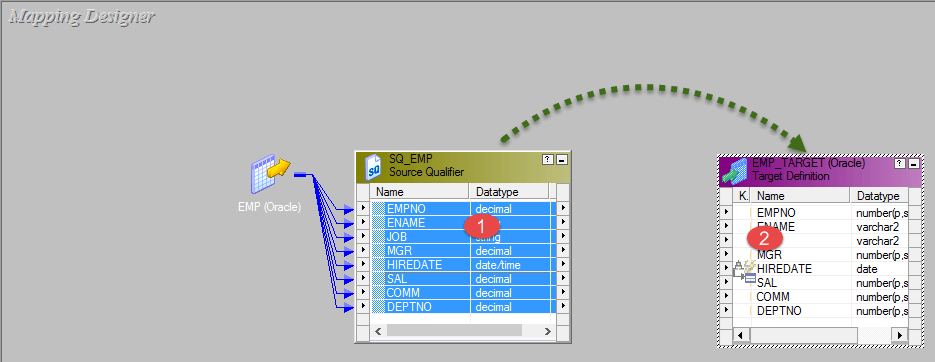


**Step 9 –**Right Click on mapping designer workspace and select “Autolink” by name option.

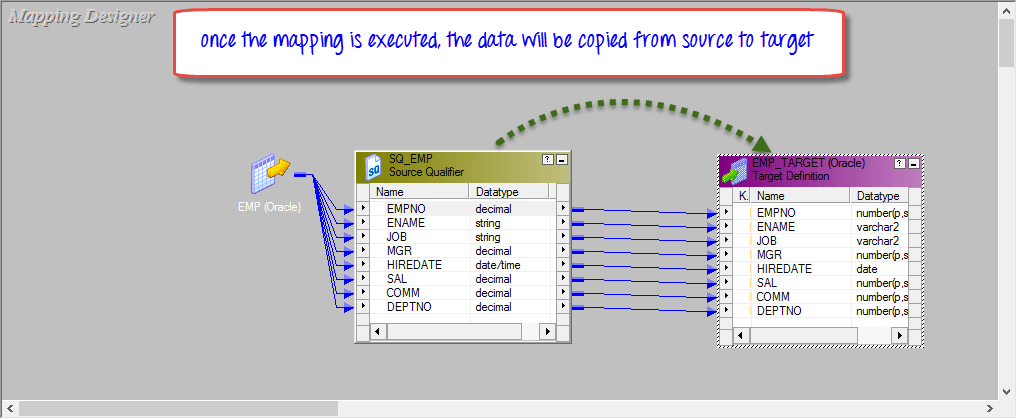


**Step 10 –**To link source with target table

1. Select Source table columns.
2. Drag and drop columns to the target table.

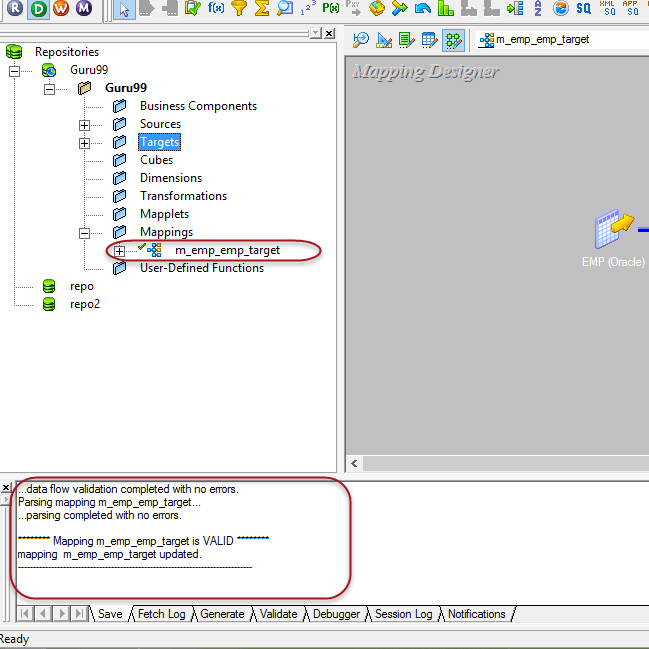


The Source and the Target tables will be linked, and connecting arrows will appear from source to the target table.



**Note** – Here you have linked all source columns to the respective target table column. It means, for every source record fetched, all the columns of the target will get loaded. If you want to exclude any specific column from getting loaded, click on that column link and press delete key from the keyboard. The link will get removed, and the target column will not get loaded.

**Step 11** – Use shortcut “ctrl+s” to save changes to your mapping. On the output window, you can see the message of mapping validation/parsing. It confirms that your mapping is valid. Also, there will be a temporary green tick mark next to the mapping name in mapping folder tree indicating mapping is done successfully.



In mappings there can be a requirement, where we need to pass variable to the mapping or there can be a scenario where we need to calculate temporary variables and further required to store them for the next session run of the mapping. For these purposes, we create mapping parameters and variables.

**Mapping Parameters and Variables**

Like every programming language, Informatica has its own way of defining parameters and variables. But unlike other programming languages, Informatica isn’t a code based language. To create parameters and variables in Informatica, you have to follow the predefined syntax and navigation.

Difference between parameters and variables –

| **Mapping Parameters** | **Mapping Variables** |
| --- | --- |
| Mapping parameters are those data types whose value once assigned remains constant throughout the mapping run. Like if you have created a mapping parameter deptno=20, then the value 20 will be constant for the whole mapping run. The parameter wherever it will be referenced will always return value 20 for that instance of mapping run. For a new mapping instance, the parameter value can be redefined. | Mapping variables are objects which can be referenced throughout the mapping run (to access their values) and their values can be reassigned. For example, a mapping variable of total\_salary can be used in a mapping, and its value can be updated based on salaries. |

The mapping parameters and variables are specific to a mapping only. They cannot be referenced in another mapping.

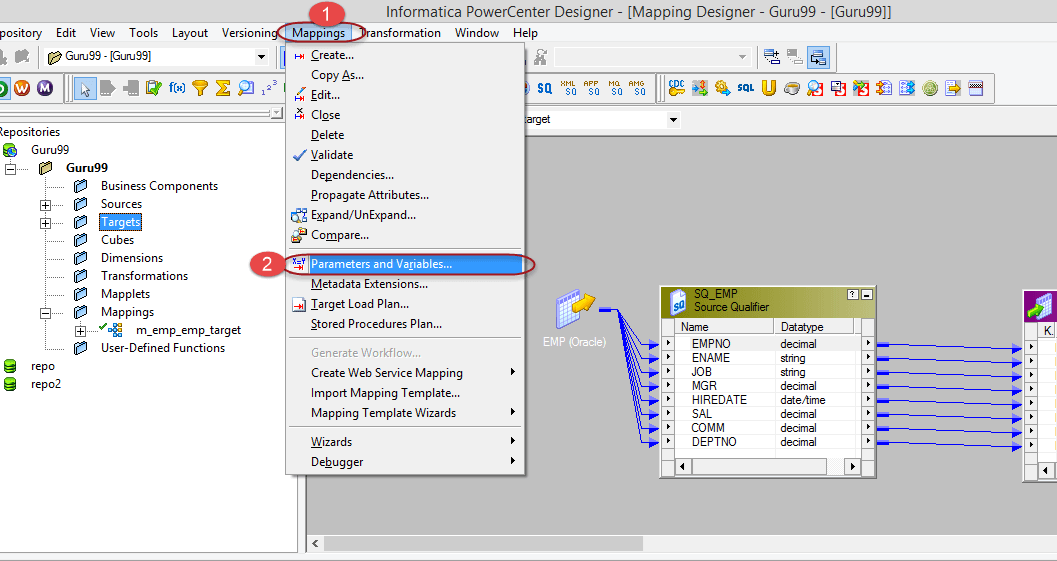
**How to Create Mapping Parameter**

When you create a mapping parameter, during execution of mapping Integration service looks for its assigned value. This values can be assigned to following places.

* Inside parameter file
* In pre-session variable assignment
* Initial value in repository
* Default value assigned during variable creation

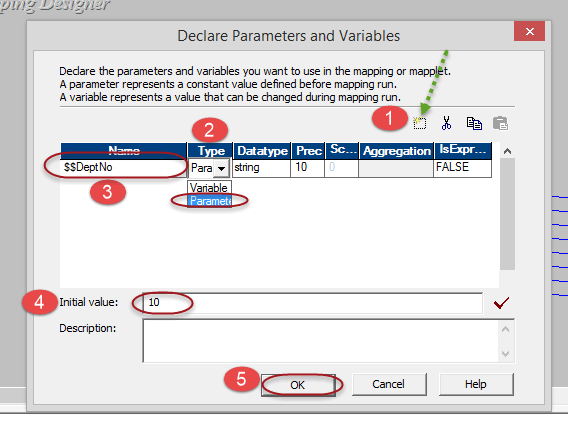
**Step 1 –**To Create mapping parameter **–**In mapping designer,

1. Select mappings menu
2. Select parameters and variables menu



**Step 2**– In next screen,

1. Click on add new variable menu
2. From drop down, select type as parameter
3. Enter parameter name as $$Deptno
4. Enter an initial value of 10
5. Select OK button

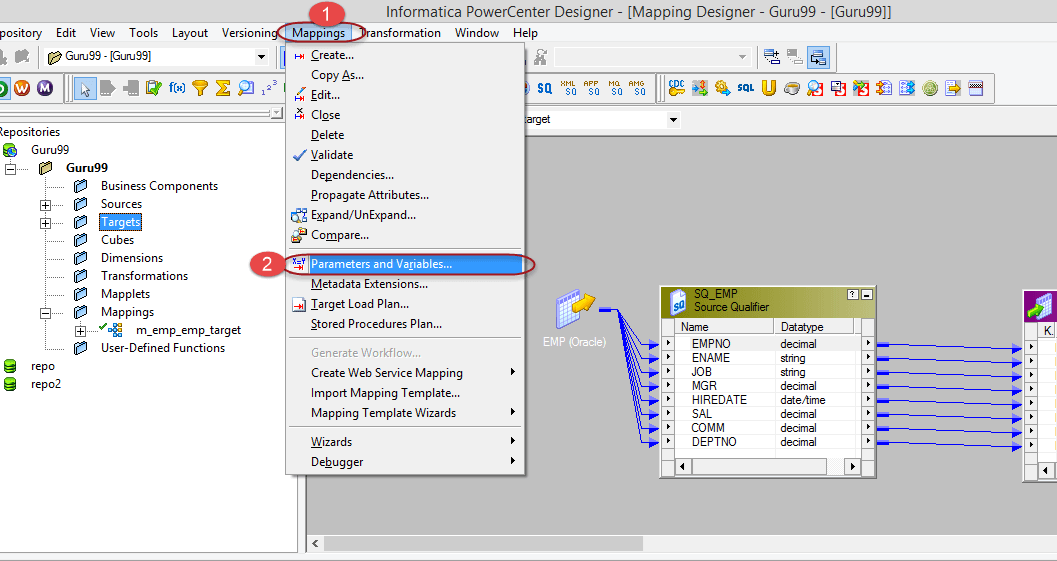


Now, you have created a mapping parameter deptno, with initially assigned value of 10, and this parameter can be referenced inside the mapping.

**How to Create Mapping Variable**

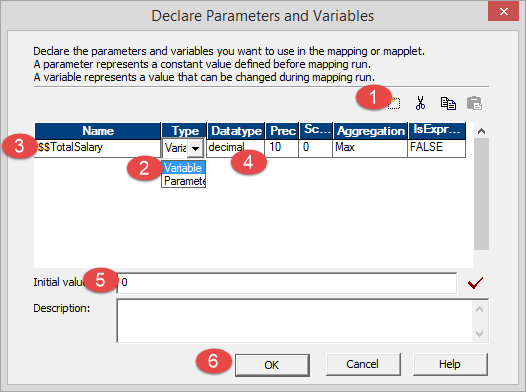
**Step 1 –**In mapping designer

1. Select mappings menu
2. Select parameters and variables menu



**Step 2**– On the next screen

1. Click on add new variable menu
2. From drop down, select type as variable
3. Enter variable name as $$TotalSalary
4. Select DataType as decimal
5. Enter an initial value of 0
6. Select OK button



This will create a mapping variable.

**Note**– mapping parameter and variable names always begin with $$.

**Summary**

Mappings are important in Informatica to match source data with target as per project requirements. We have discussed stage mappings and concept of Mapping Variables and Parameters.

**What is Workflow?**

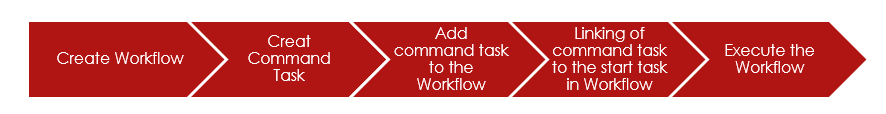
Workflow is a group of instructions/commands to the integrations service in Informatica. The integration service is an entity which reads workflow information from the repository, fetches data from sources and after performing transformation loads it into the target.

Workflow – It defines how to run tasks like **session task, command task, email task**, etc.

To create a workflow

1. You first need to create tasks
2. And then add those tasks to the workflow.

A Workflow is like an empty container, which has the capacity to store an object you want to execute. You add tasks to the workflow that you want to execute. In this tutorial, we are going to do following things in workflow.



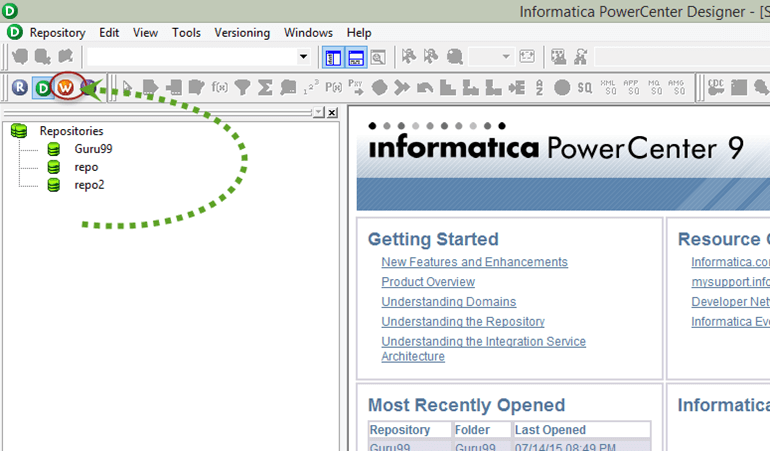
Workflow execution can be done in two ways

* **Sequence** : Tasks execute in the order in which they are defined
* **Event based :**Tasks gets executed based on the event conditions.

**What You Will Learn**

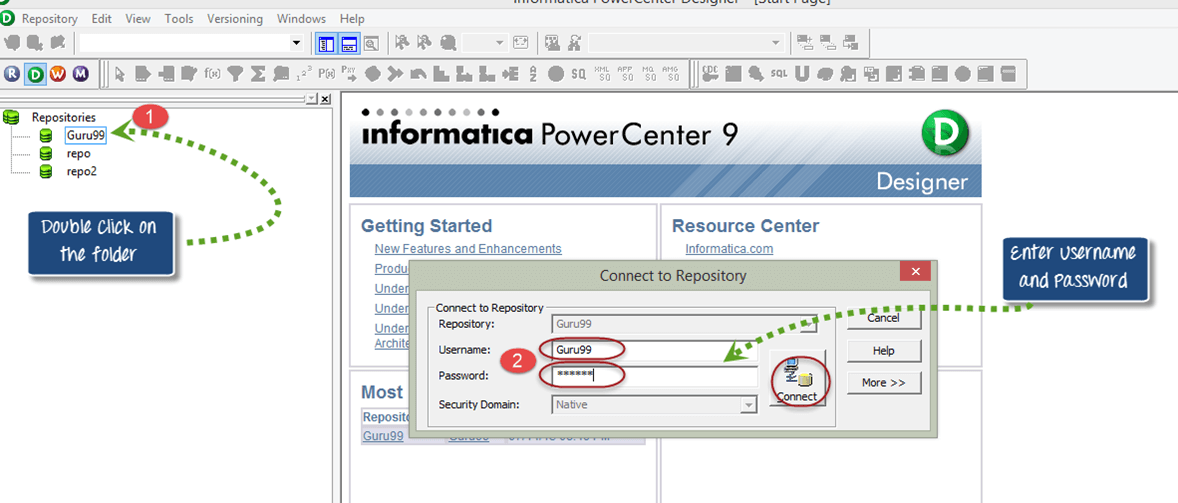
**How to open Workflow Manager**

**Step1** – In the Informatica Designer, Click on the Workflow manager icon



**Step 2** – This will open a window of Workflow Manager. Then, in the workflow Manager.

1. We are going to connect to repository “guru99”, so double click on the folder to connect.
2. Enter user name and password then select “Connect Button”.



**Step 3**– In the workflow manager.

1. Right click on the folder
2. In the pop up menu, select open option

This will open up the workspace of Workflow manager.

**How to Create Connections for Workflow Manager**

To execute any task in workflow manager, you need to create **connections**. By using these connections, Integration Service connects to different objects.

For Example, in your mapping if you have source table in oracle database, then you will need oracle connection so that integration service can connect to the oracle database to fetch the source data.

Following type of connections can be created in workflow manager.

* Relational Connection
* Ftp Connection
* Queue
* Application

The choice of connection you will create, will depend on the type of source and target systems you want to connect. More often, you would be using **relational connections.**

**To Create a Relational Connection**

**Step 1** – In Workflow Manager

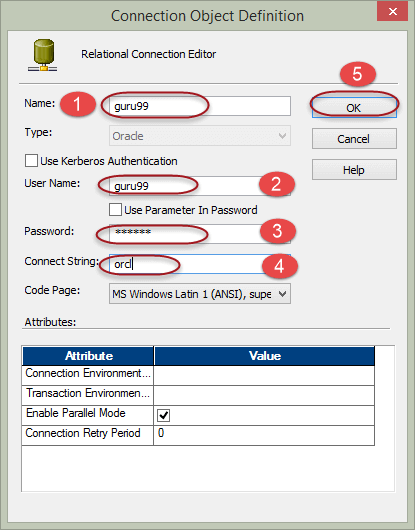
1. Click on the Connection menu
2. Select Relational Option

**Step 2** – In the pop up window

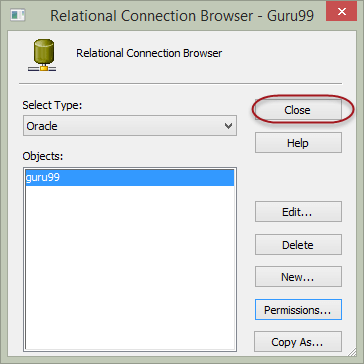
1. Select Oracle in type
2. Click on the new button

**Step 3** – In the new window of connection object definition

1. Enter Connection Name (New Name-guru99)
2. Enter username
3. Enter password
4. Enter connection string
5. Leave other settings as default and Select OK button



**Step 4** – You will return on the previous window. Click on the close button.

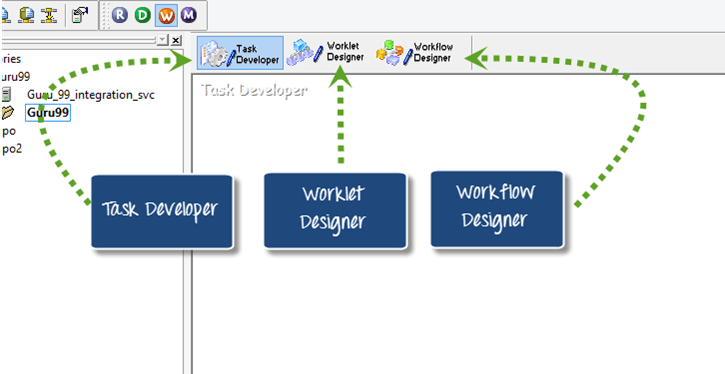


Now you are set with the relational connection in workflow manager.

**Components of Workflow manager**

There are three component tools of workflow manager that helps in creating various objects in workflow manager. These tools are

* Task Developer
* Worklet Designer
* Workflow Designer



**Task Developer**– Task developer is a tool with the help of which you can create reusable objects. Reusable object in workflow manager are objects which can be reused in multiple workflows. For Example, if you have created a command task in task developer, then you can reuse this task in any number of workflows.

The role of Workflow designer is to execute the tasks those are added in it. You can add any no of tasks in a workflow.

You can create three types of reusable tasks in task developer.

* Command task
* Session task
* Email task

**Command task –**A command task is used to execute different windows/unix commands during the execution of the workflow. You can create command task to execute various command based tasks. With help of this task you can execute commands **to create files/folders, to delete files/folders, to do ftp of files** etc.

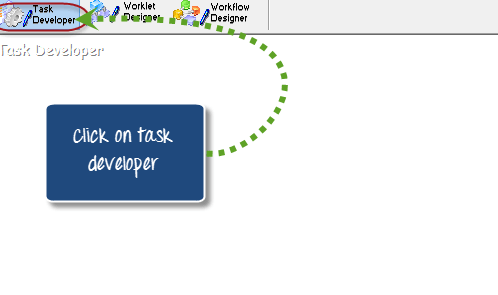
**Session Task –**A session task in Informatica is required to run a mapping.

* Without a session task, you cannot execute or run a mapping
* A session task can execute only a single mapping. So, there is a one to one relationship between a mapping and a session
* A session task is an object with the help of which informatica gets to know how and where to execute a mapping and at which time
* Sessions cannot be executed independently, a session must be added to a workflow
* In session object cache properties can be configured and also advanced performance optimization configuration.

**Email task –**With the help of email task you can send email to defined recipients when the Integration Service runs a workflow. For example, if you want to monitor how long a session takes to complete, you can configure the session to send an email containing the details of session start and end time. Or, if you want the Integration Service to notify you when a workflow completes/fails, you can configure the email task for the same.

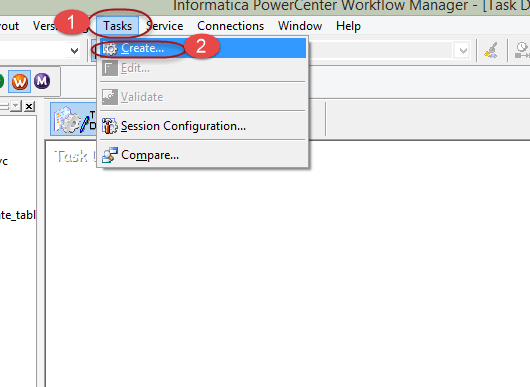
**How to create command task**

**Step 1**– To create a command task we are going to use Task Developer. In Workflow Manager, open the task developer by clicking on tab “task developer” from the menu.



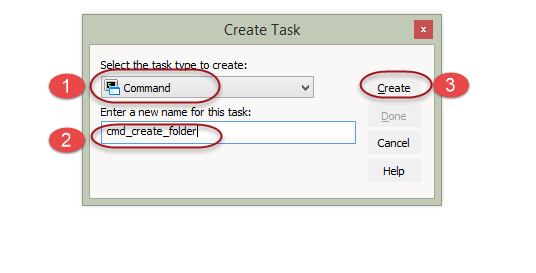
**Step 2** – Once task developer is opened up, follow these steps

1. Select Tasks menu
2. Select Create option

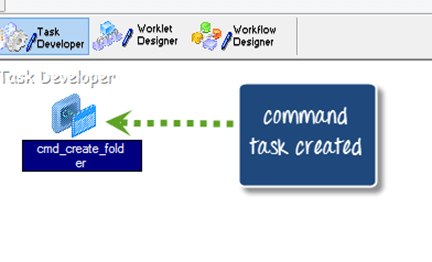


**Step 3** – In the create task window

1. Select command as type of task to create
2. Enter task name
3. Select create button



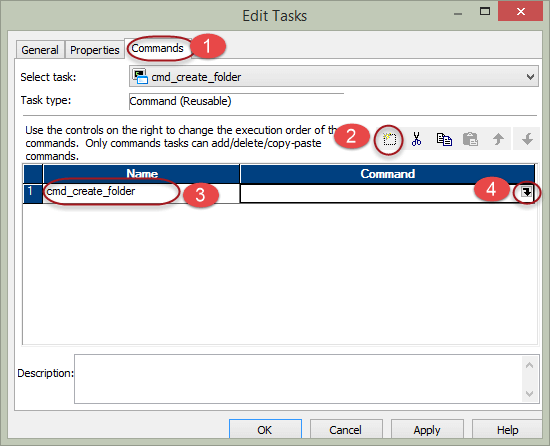
This will create command task folder. Now you have to configure the task to add command in it, that we will see in next step.



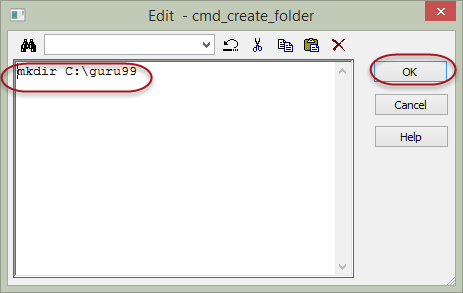
**Step 4** – To configure the task, double click on the command task icon and it will open an “edit task window”. On the new edit task window

1. Select the commands menu
2. Click on the add new command icon
3. Enter command name
4. Click on the command icon to add command text

This will open a command editor box.

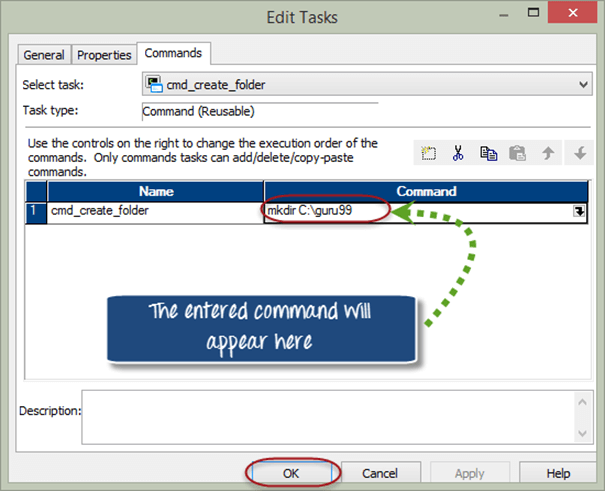


**Step 5** – On the command editor box, enter the command “mkdir C:\guru99” (this is the windows command to create a folder named “guru99”) and select OK.

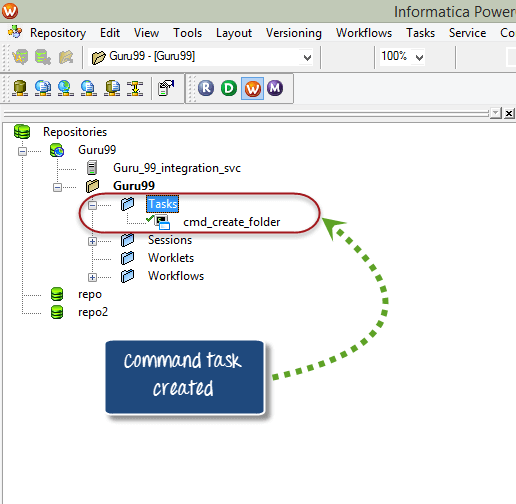


Afther this step you will return to the edit tasks window and you will be able to see the command you added in to the command text box.

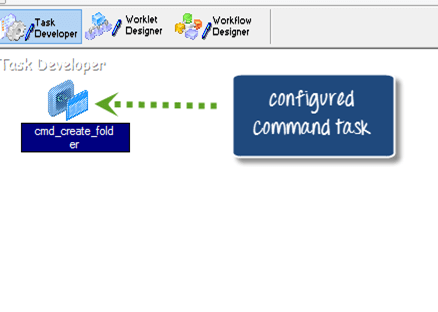
**Step 6** – Click OK on the edit task window,



The command task will be created in the task developer under “Guru99” repository.



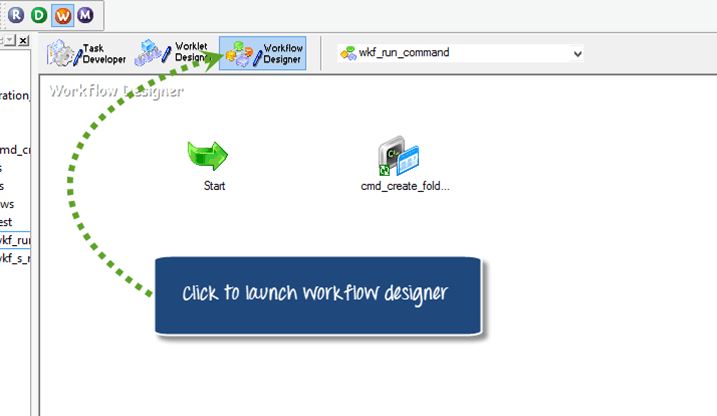
**Note** – use ctrl+s shortcut to save the changes in repository



**How to create workflow to execute command task**

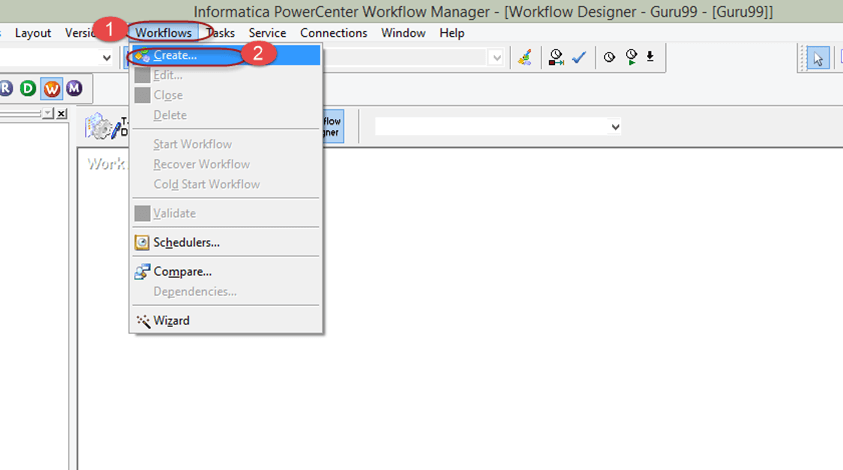
To execute command taks you have to switch on to workflow designer. A workflow designer is a parent or container object in which you can add multiple tasks and when workflow is executed, all the added tasks will execute. To create a workflow

**Step 1** – Open the workflow designer by clicking on workflow designer menu



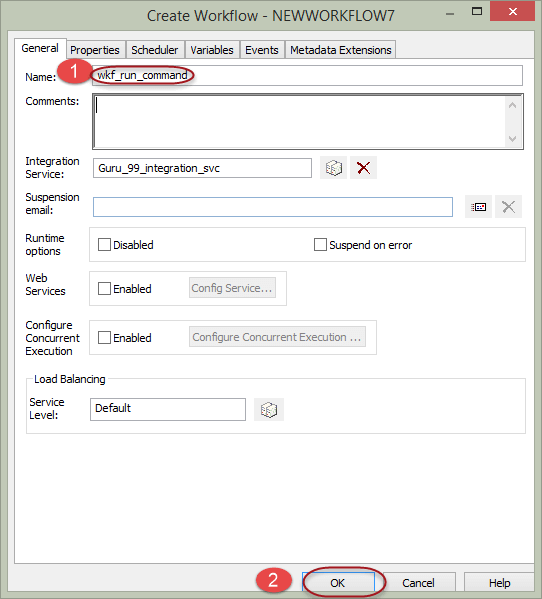
**Step 2** – In workflow designer

1. Select workflows menu
2. Select create option



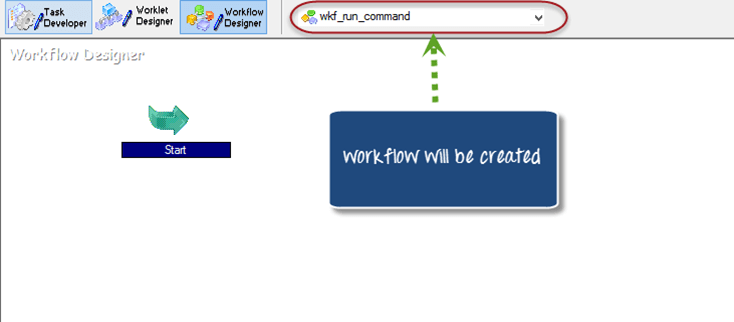
**Step 3** – In create workflow window

1. Enter workflow name
2. Select OK Button ( leave other options as default)



This will create the workflow.

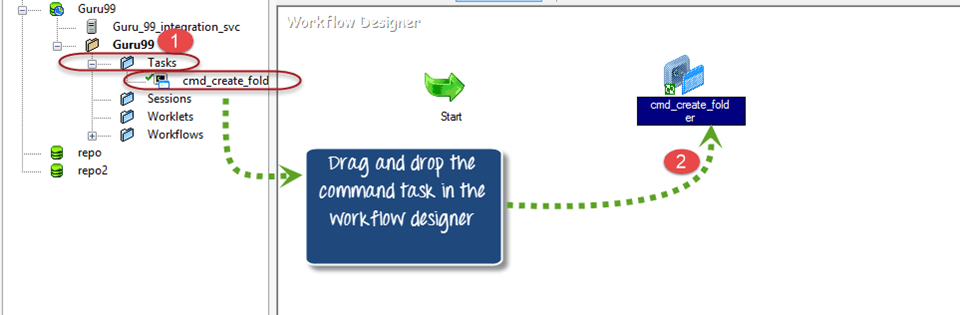
**Naming Convention –**Workflow names are prefixed with using ‘**wkf\_’**, if you have a session named ‘**s\_m\_employee\_detail**‘ then workflow for the same can be named as ‘**wkf\_s\_m\_employee\_detail’**.



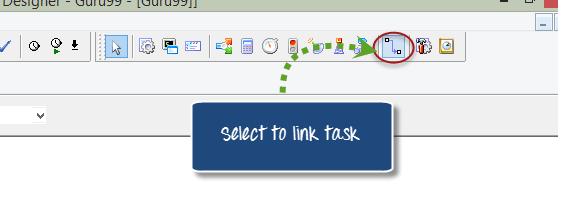
When you create a workflow, it does not consist of any tasks. So, to execute any task in a workflow you have to add task in it.

**Step 4** – To add command task that we have created in Task developer to the workflow desinger

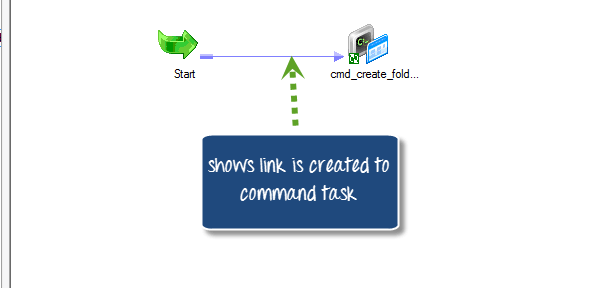
1. In the navigator tree, expand the tasks folder
2. Drag and drop the command task to workflow designer



**Step 5** – Select the “link task option” from the toolbox from the top menu. (The link task option links various tasks in a workflow to the start task, so that the order of execution of tasks can be defined).



**Step 6** – Once you select the link task icon, it will allow you to drag the link between start task and command task. Now select the start task and drag a link to the command task.

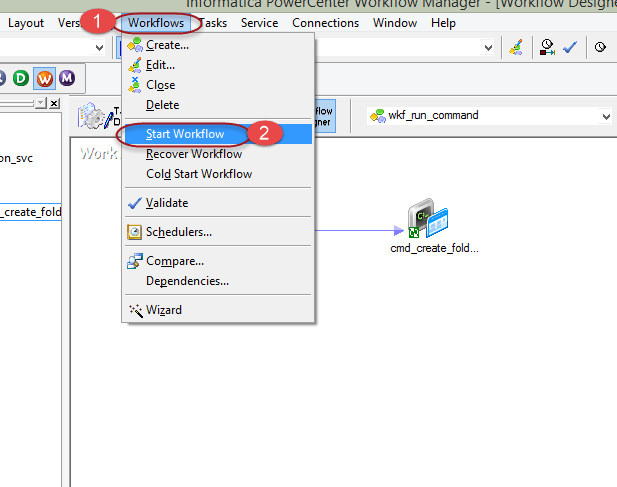


Now you are ready with the workflow having a command task to be executed.

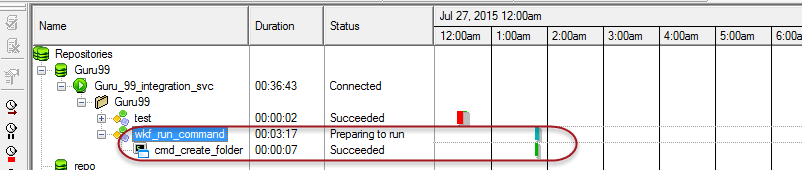
**How to execute workflow**

**Step 1** – To execute the workflow

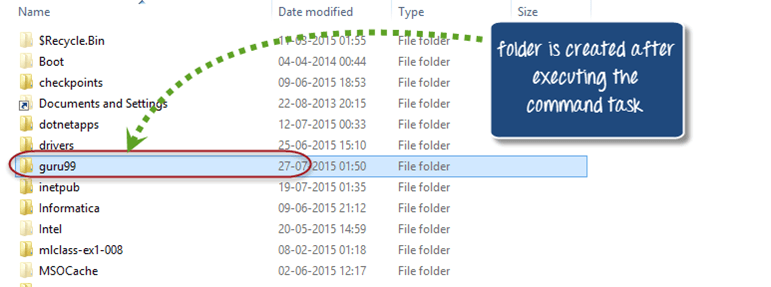
1. Select workflows option from the menu
2. Select start workflow option



This will open workflow monitor window and executes the workflow



Once the workflow is executed, it will execute the command task to create a folder (guru99 folder) in the defined directory.



**Session Task**

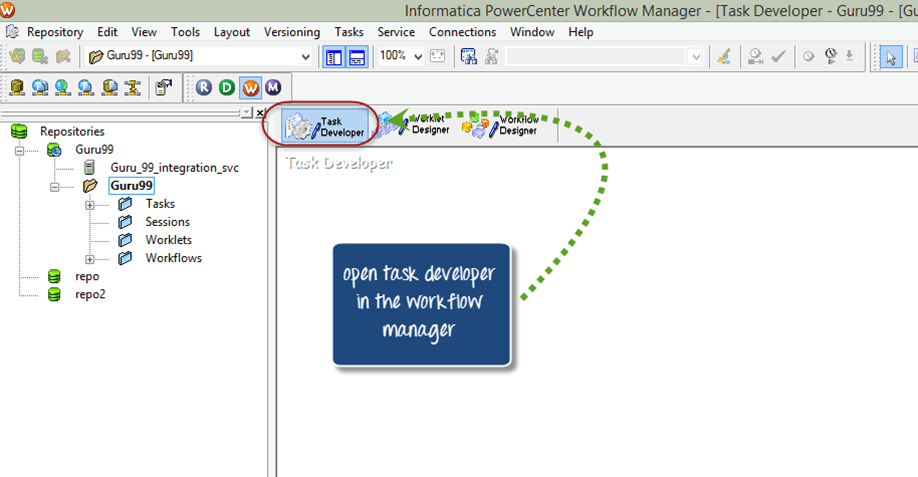
A session task in Informatica is required to run a mapping.

Without a session task, you cannot execute or run a mapping and a session task can execute only a single mapping. So, there is a one to one relationship between a mapping and a session. A session task is an object with the help of which Informatica gets to know how and where to execute a mapping and at which time. Sessions cannot be executed independently, a session must be added to a workflow. In session object cache properties can be configured and also advanced performance optimization configuration.

**How to create a session task.**

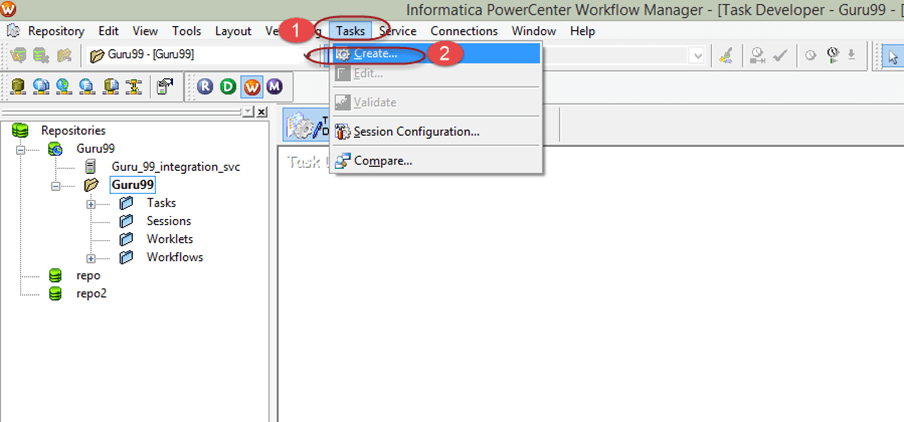
In this exercise you will create a session task for the mapping “m\_emp\_emp\_target” which you created in the previous article.

**Step1** – Open Workflow manager and open task developer



**Step 2** – Now once the task developer opens, in the workflow manager go to main menu

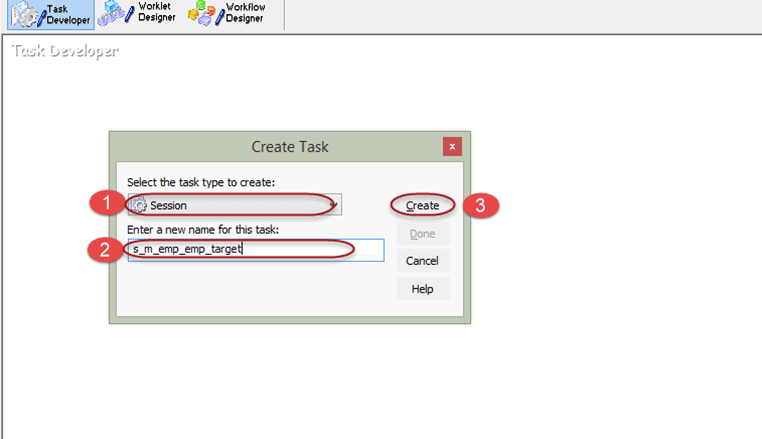
1. Click on task menu
2. Select create option



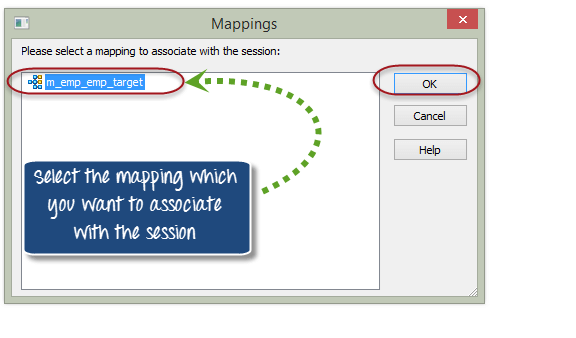
This will open a new window “Create Task”

**Step 3** – In the create task window

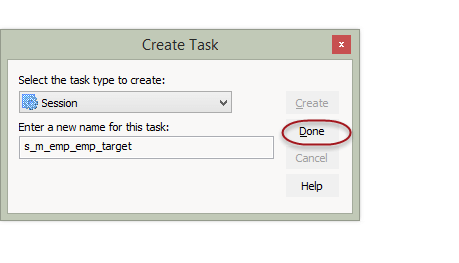
1. Select session task as type of task.
2. Enter name of task.
3. Click create button



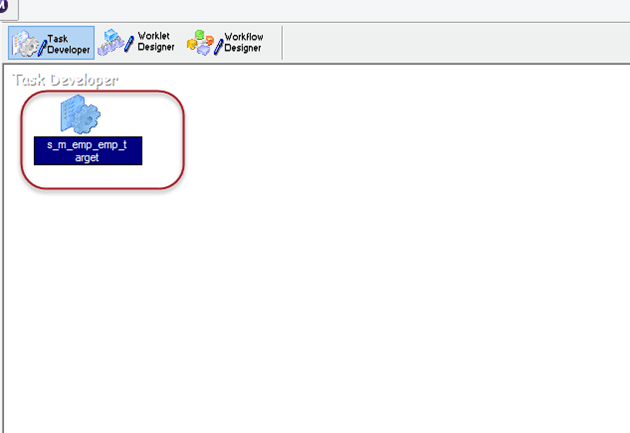
**Step 4** – A window for selecting the mapping will appear. Select the mapping which you want to associate with this session, for this example select “m\_emp\_emp\_target” mapping and click OK Button.



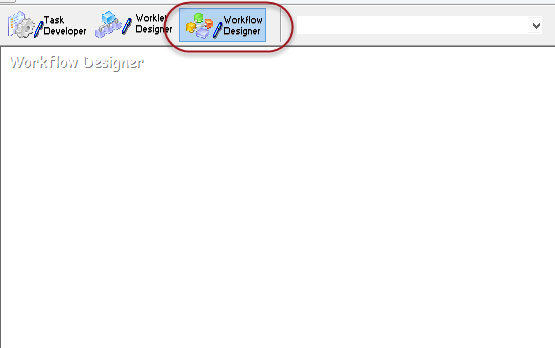
**Step 5** – After that, click on “Done” button



Session object will appear in the task developer

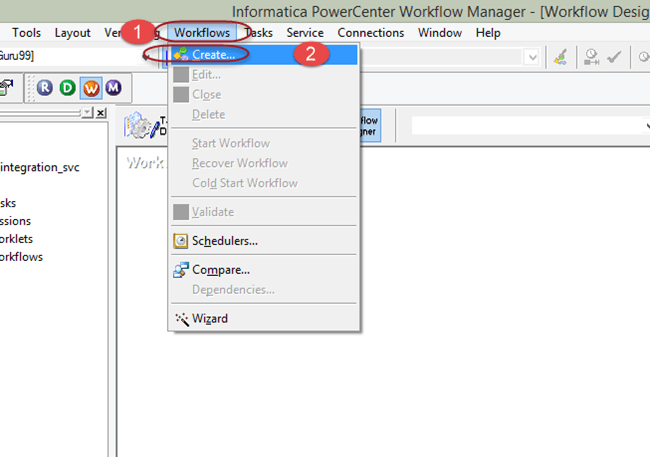


**Step 6** – In this step you will create a workflow for the session task. Click on the workflow designer icon.



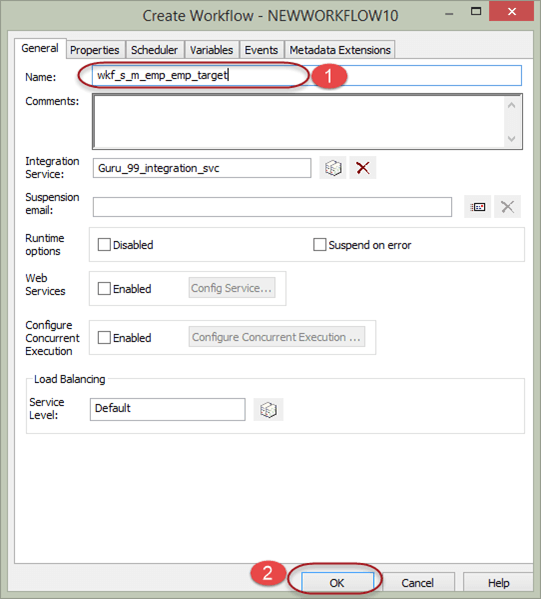
**Step 7** – In the workflow designer tool

1. Click on workflow menu
2. Select create option

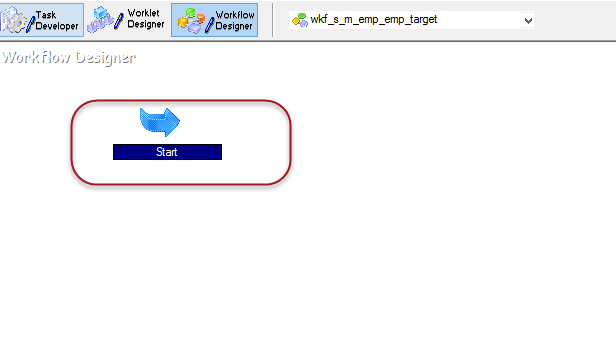


**Step 8** – In the create workflow window

1. Enter workflow name
2. Select OK. ( leave other properties as default, no need to change any properties)

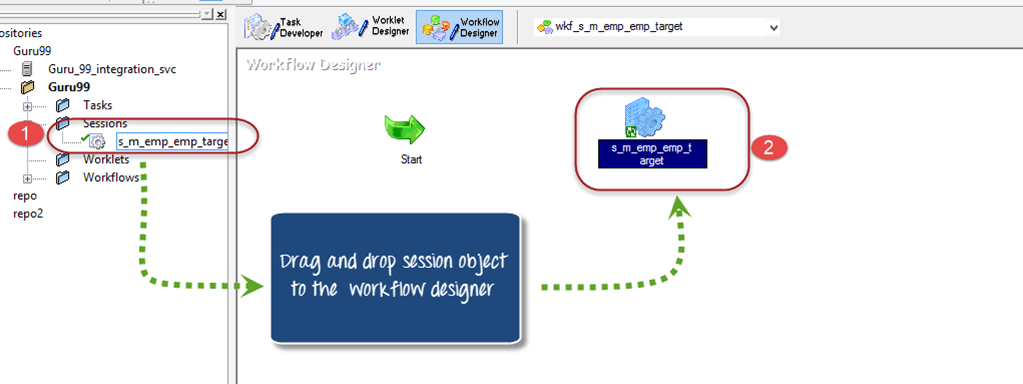


In workflow manager a start task will appear, it’s a starting point of execution of workflow.

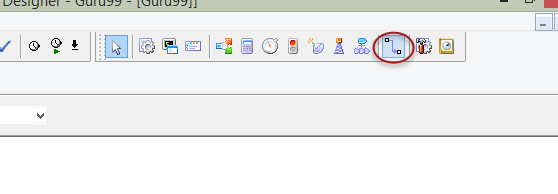


**Step 9** – In workflow manager

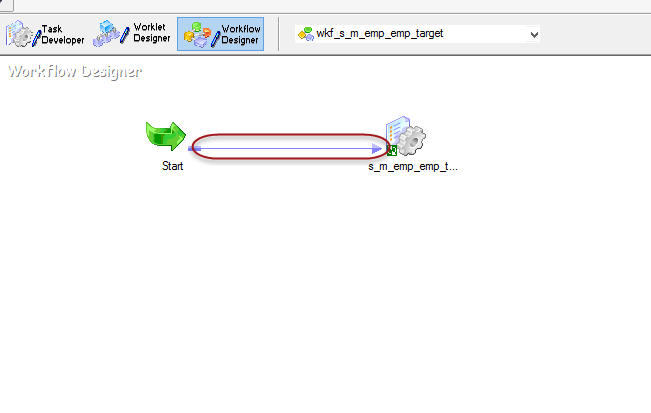
1. Expand the sessions folder under navigation tree.
2. Drag and drop the session you created in the workflow manager workspace.



**Step 10** – Click on the link task option in the tool box.



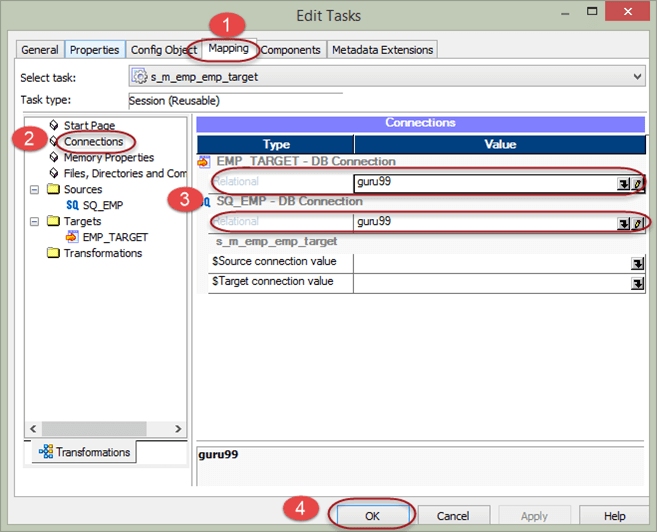
**Step 11**– Link the start task and session task using the link.



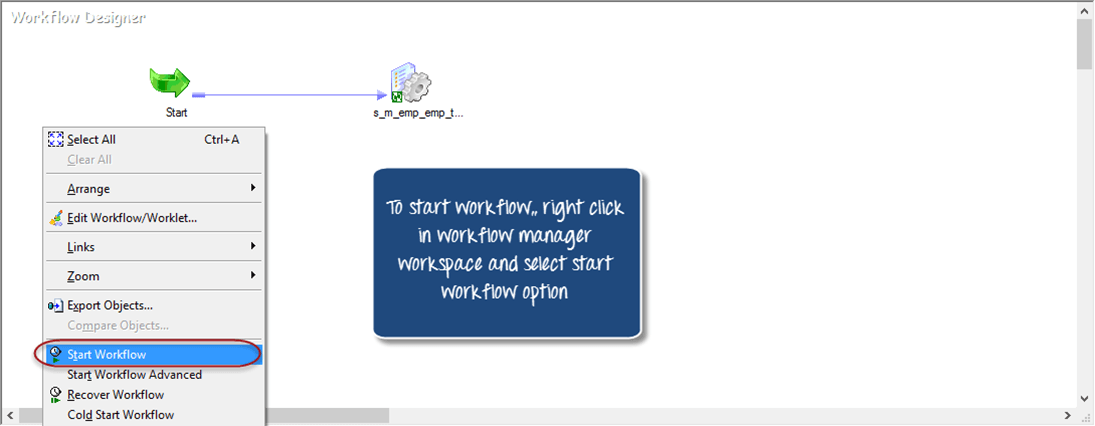
**Step 12** – Double click on the session object in wokflow manager. It will open a task window to modify the task properties.

**Step 13** – In the edit task window

1. Select mapping tab
2. Select connection property
3. Assign the connection to source and target, the connection which we created in early steps.
4. Select OK Button



Now your configuration of workflow is complete, and you can execute the workflow.



**How to add multiple tasks to a start task**

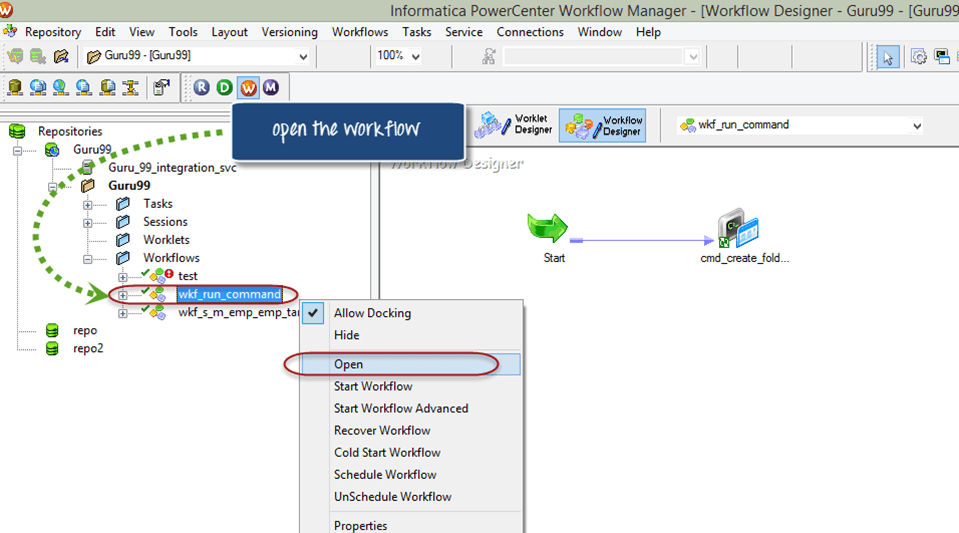
The start task is a starting point for the execution of workflow. There are two ways of linking multiple tasks to a start task.

1. Parallel
2. Serial

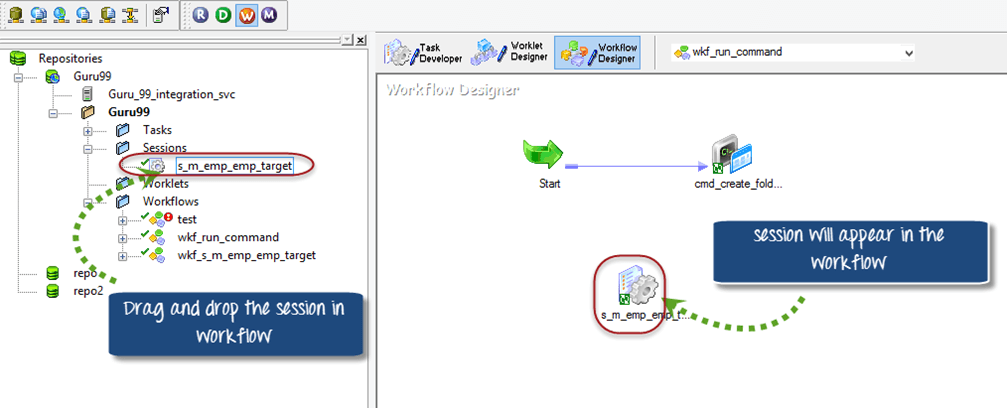
In parallel linking the tasks are linked directly to the start task and all tasks start executing in parallel at same time.

**How to add tasks in parallel**

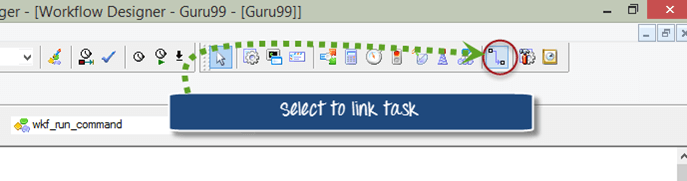
**Step 1 –**In the workflow manager, open the workflow “wkf\_run\_command”



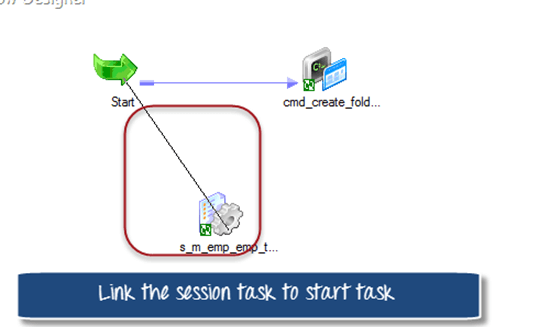
**Step 2 –**In the workflow, add session task “s\_m\_emp\_emp\_target”. ( by selecting session and then drag and drop)



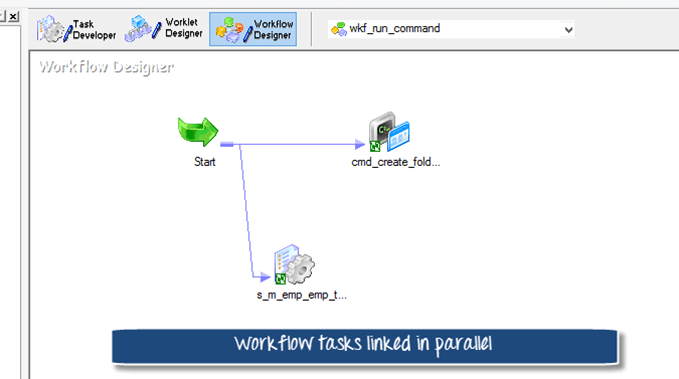
**Step 3 –**Select the link task option from the toolbox



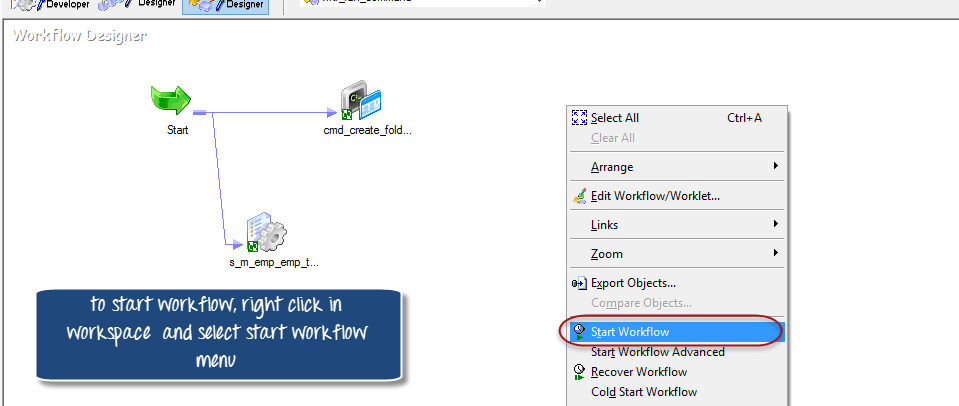
**Step 4 –**link the session task to the start task (by clicking on start taks, holding the click and connecting to session task)

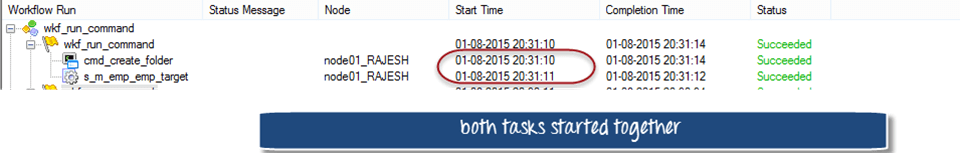


After linking the session task, the workflow will look like this.



**Step 5 –**Start the workflow and monitor in the workflow monitor.



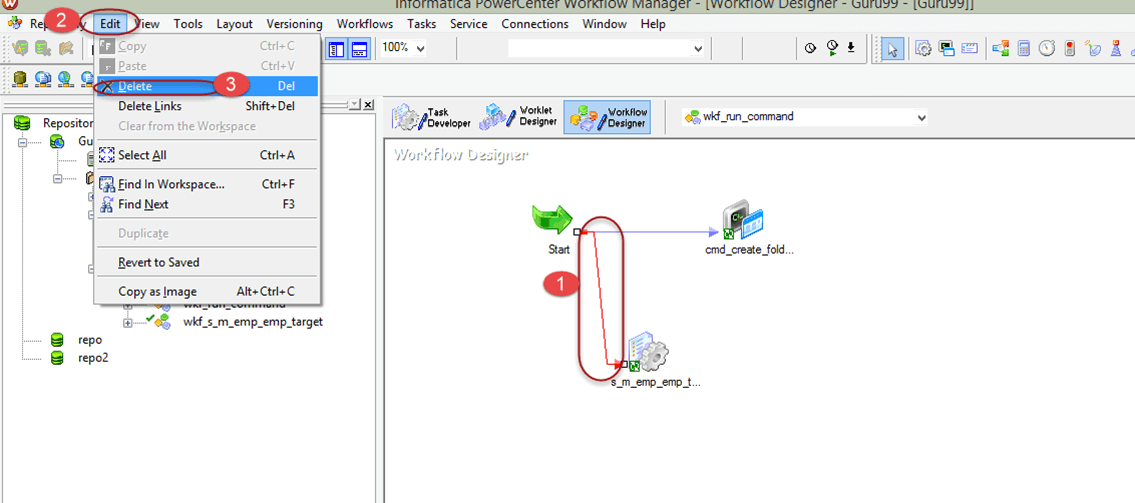


**How to add tasks in serial mode**

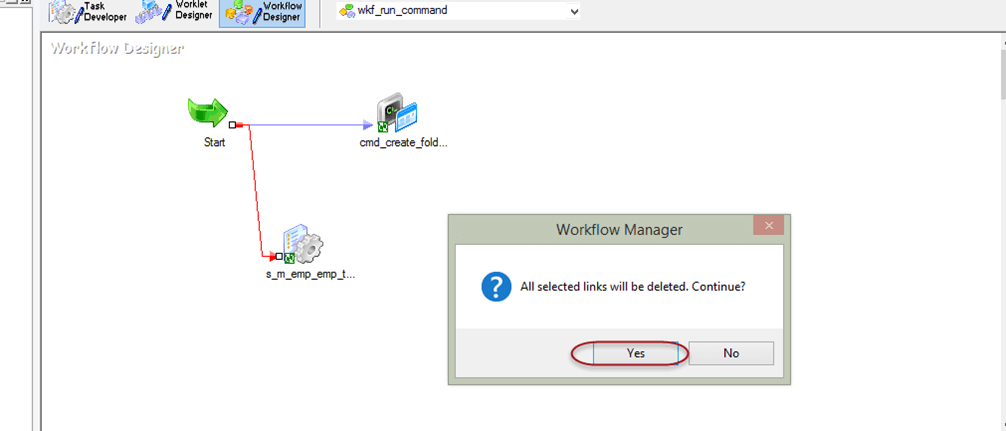
But before we add tasks in serial mode, we have to delete the task that we added to demonstrate parallel execution of task. For that

**Step 1 –**Open the workflow “w.kf\_run\_command”

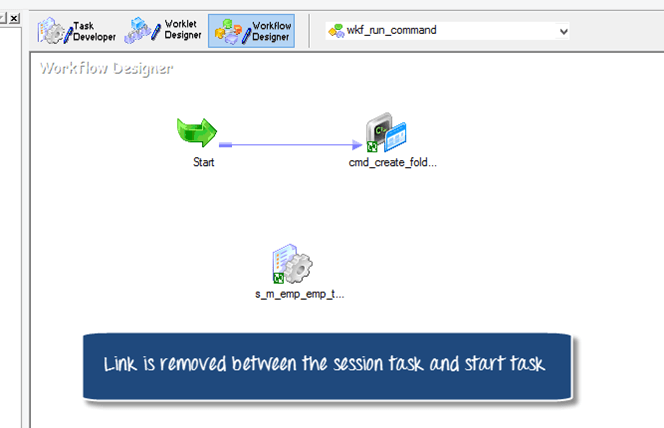
1. Select the link to the session task.
2. Select edit option in the menu
3. Select delete option



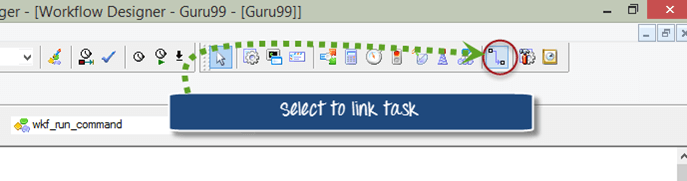
**Step 2 –**Confirmation dialogue box will appear in a window, select yes option



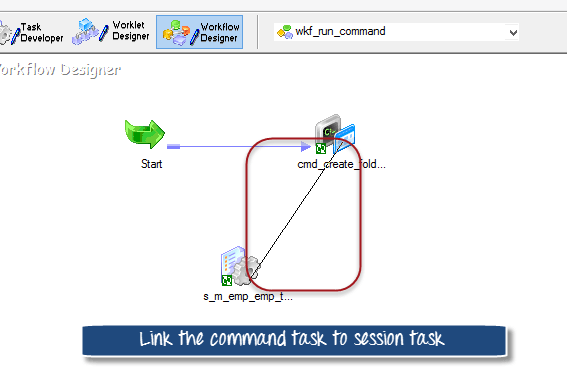
The link between the start task and session task will be removed.



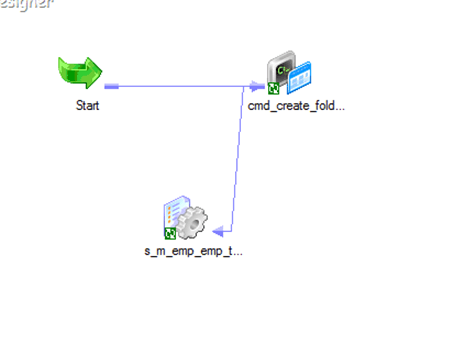
**Step 3 –**Now again go to top menu and select the link task option from the toolbox



**Step 4 –**link the session task to the command task

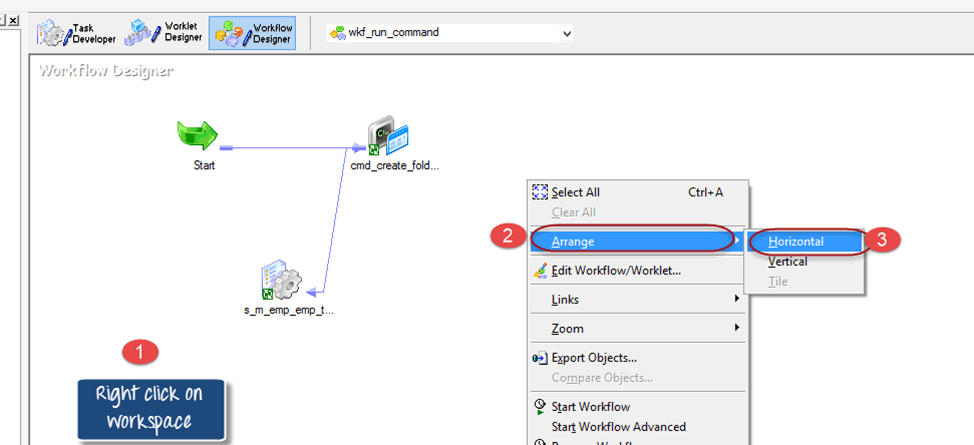


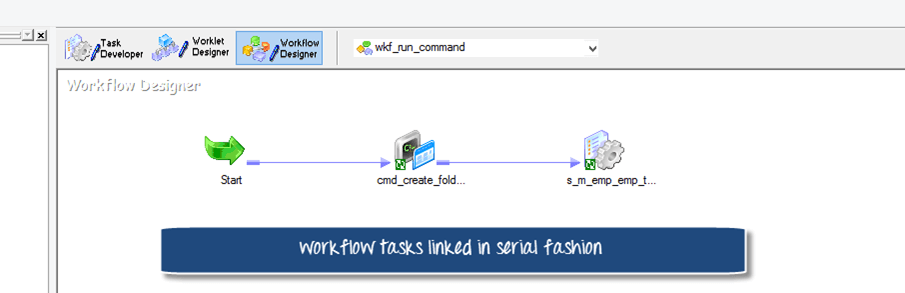
After linking the workflow will look like this



**Step 5 –**To make the visual appearance of workflow more clear

1. Right click on wokspace of workflow
2. Select arrange menu
3. Select Horizontal option





If you start the workflow the command task will execute first and after its execution, session task will start.

**Workflow Variable**

Workflow variables allows different tasks in a workflow to exchange information with each other and also allows tasks to access certain properties of other tasks in a workflow. For example, to get the current date you can use the inbuilt variable “sysdate”.

Most common scenario is when you have multiple tasks in a workflow and in one task you access the variable of another task. For example, if you have two tasks in a workflow and the requirement is to execute the second task only when first task is executed successfully. You can implement such scenario using predefined variable in the workflow.

**Implementing the scenario**

We had a workflow “wkf\_run\_command” having tasks added in serial mode. Now we will add a condition to the link between session task and command task, so that, only after the success of command task the session task will be executed.

**Step 1 –**Open the workflow “wkf\_run\_command”

**Step 2 –**Double click on the link between session and command task

An Expression window will appear

**Step 3 –**Double click the status variable under “cmd\_create\_folder” menu. A variable “$cmd\_create\_folder.status” will appear in the editor window on right side.

**Step 4 –**Now we will set the variable “$cmd\_create\_folder.status” condition to succeeded status . which means when the previous tasks is executed and the execution was success, then only execute the next session task.

1. Change the variable to “$cmd\_create\_folder.status=SUCCEEDED” value.
2. Click OK Button

The workflow will look like this

When you execute this workflow, the command task executes first and only when it succeeds then only the session task will get executed.

**Workflow Parameter**

Workflow parameters are those values which remain constant throughout the run. once their value is assigned it remains same. Parameters can be used in workflow properties and their values can be defined in parameter files. For example, instead of using hard coded connection value you can use a parameter/variable in the connection name and value can be defined in the parameter file.

Parameter files are the files in which we define the values of mapping/workflow variables or parameters. There files have the extension of “.par”. As a general standard a parameter file is created for a workflow.

Advantages of Parameter file

* Helps in migration of code from one environment to other
* Alows easy debugging and testing
* Values can be modified with ease without change in code

**Structure of parameter file**

The structure of parameter file

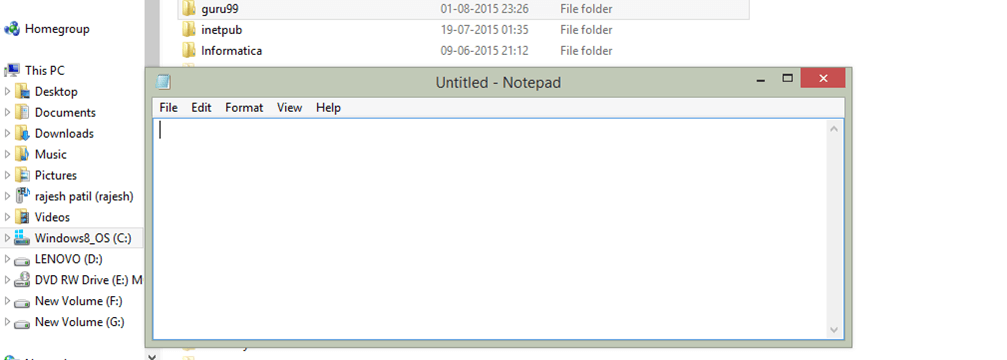
* [folder\_name.WF:Workflow\_name]
* $Parameter\_name=Parameter\_value

Folder\_name is the name of repository folder, workflow name is the name of workflow for which you are creating the parameter file.

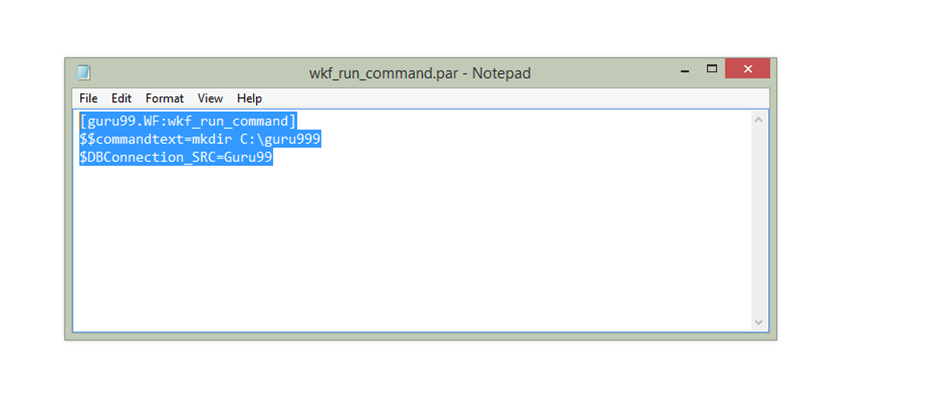
We will be creating a parameter file for the database connection “guru99” which we assigned in our early sessions for sources and targets.

**How to create parameter file**

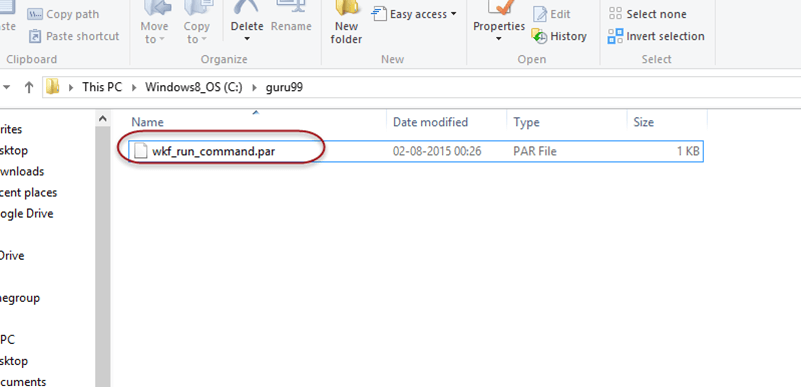
**Step 1 –**Create a new empty file (notepad file)



**Step 2 –**In the file enter text as shown in figure



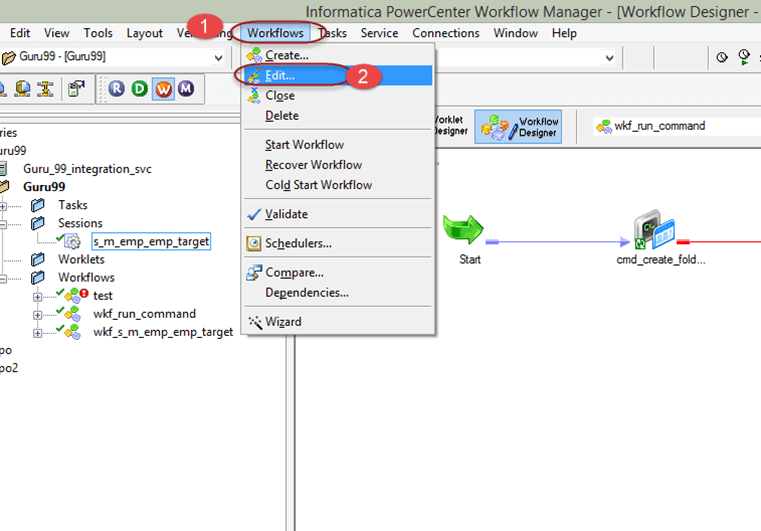
**Step 3 –** Save the file under a folder guru99 at the location “C:\guru99” as “wkf\_run\_command.par”



In the file we have created a parameter “$DBConnection\_SRC”, we will assign the same to a connection in our workflow.

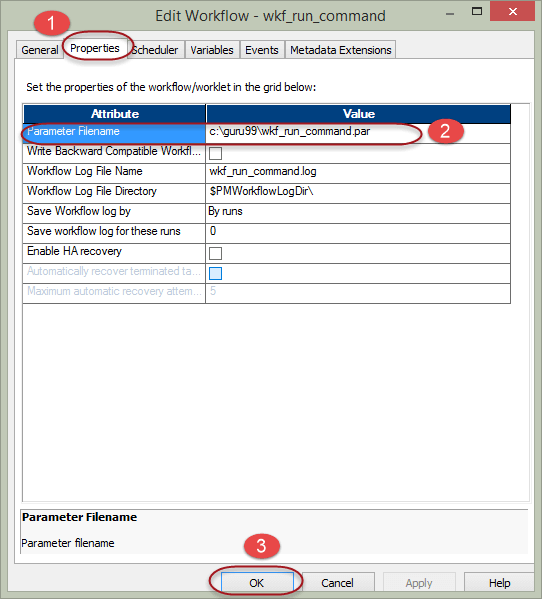
**Step 4-**Open the workflow “wkf\_run\_command”

1. Select workflows menu
2. Select edit option



**Step 5 –**This will open up edit workflow window, in this window

1. Go to properties tab menu
2. Enter the parameter file name as “c:\guru99\wkf\_run\_command.par”
3. Select OK Button

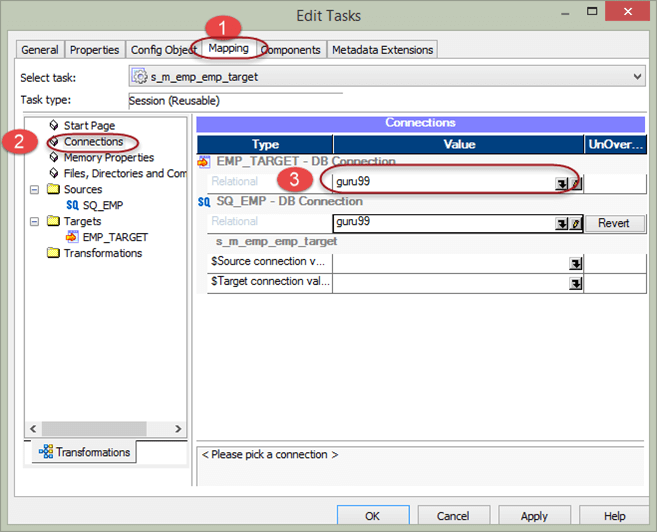


Now we are done with defining the parameter file content and point it to a workflow.

Next step is to use the parameter in session.

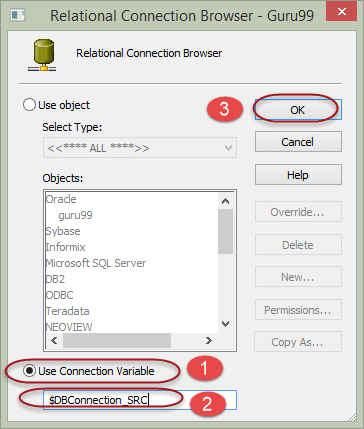
**Step 6 –**In workflow double click on the session “s\_m\_emp\_emp\_target”, then

1. Select mappings tab menu
2. Select connection property in the left panel
3. Click on the target connection, which is hardcoded now as “guru99”

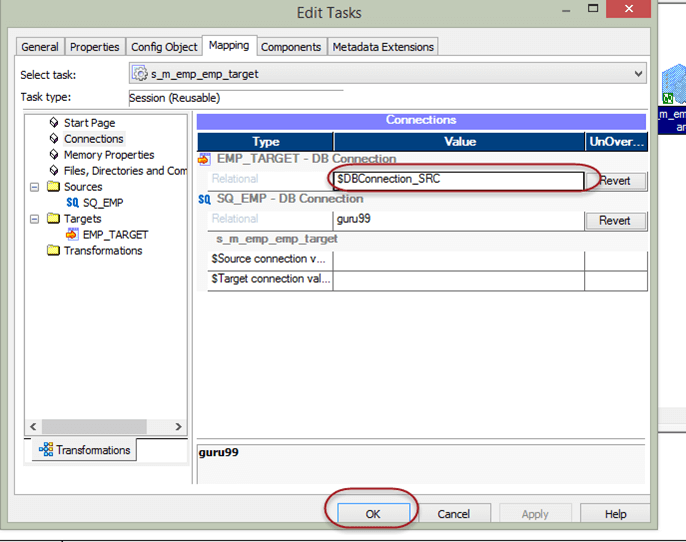


**Step 7 –**A connection browser window will appear, in that window

1. Select the option to use connection variable
2. Enter connection variable name as “$DBConnection\_SRC”
3. Select Ok Button



**Step 8** – In the edit task window connection variable will appear for the target, Select OK button in the edit task window.



Now we are done with creating parameter for a connection and assigning its value to parameter file.

When we execute the workflow, the workflow picks the parameter file looks for the value of its paramters/variables in the parameter file and takes those values.

**What is Workflow Monitor?**

Workflow monitor is a tool with the help of which you can monitor the execution of workflows and task assigned to the workflow.

In workflow monitor you can,

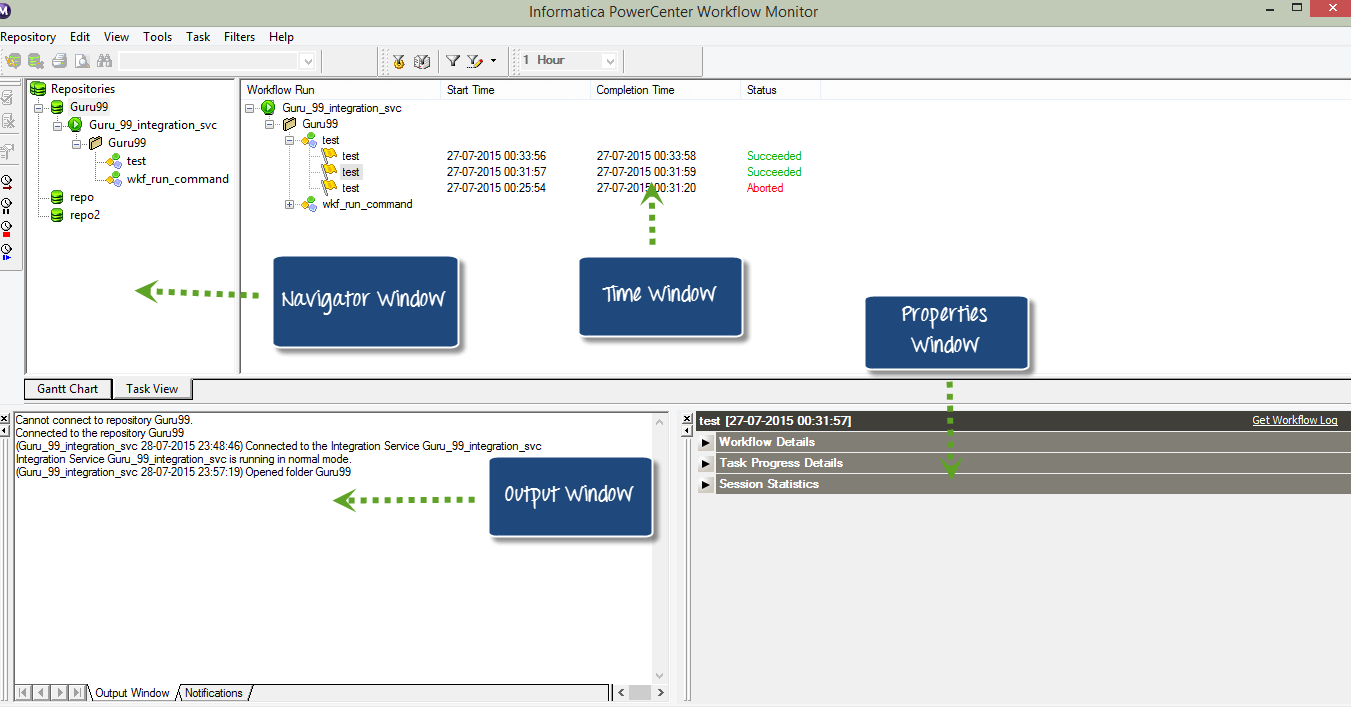
* See the details of execution
* See the history of the workflow execution
* Stop, abort or restart workflows, and tasks
* Display the workflows those who are executed at least one time

In this Tutorial – you will learn

* [How to open Workflow Monitor](https://www.guru99.com/workflow-monitor-informatica.html#1)
* [Task View](https://www.guru99.com/workflow-monitor-informatica.html#2)
* [Gantt Chart View](https://www.guru99.com/workflow-monitor-informatica.html#3)
* [Example- How to monitor and view details](https://www.guru99.com/workflow-monitor-informatica.html#4)

Workflow monitor consists of following windows –

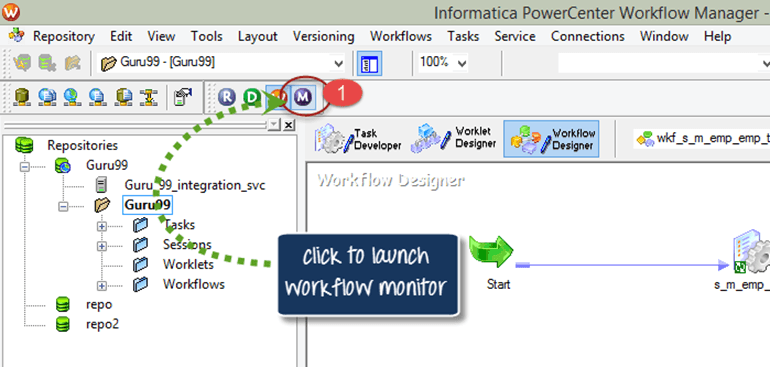
* Navigator window- shows the monitored repositories, folders & integration service
* Output window – displays the messages from integration services and repository
* Properties window – displays the details/properties about tasks and workflows
* Time window – displays the progress of the running tasks & workflows with timing details.



Now, let see what we can do in Workflow Monitor

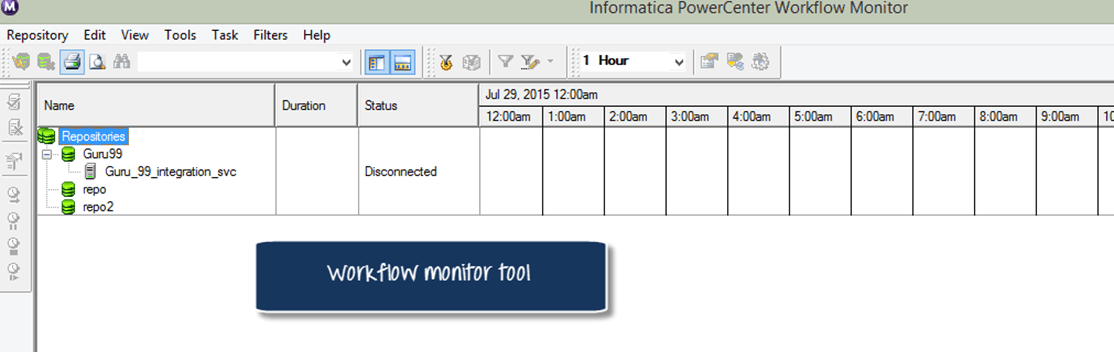
**How to open Workflow Monitor**

**Step 1** – In Informatica Designer or Workflow manager toolbox, click on the workflow monitor icon



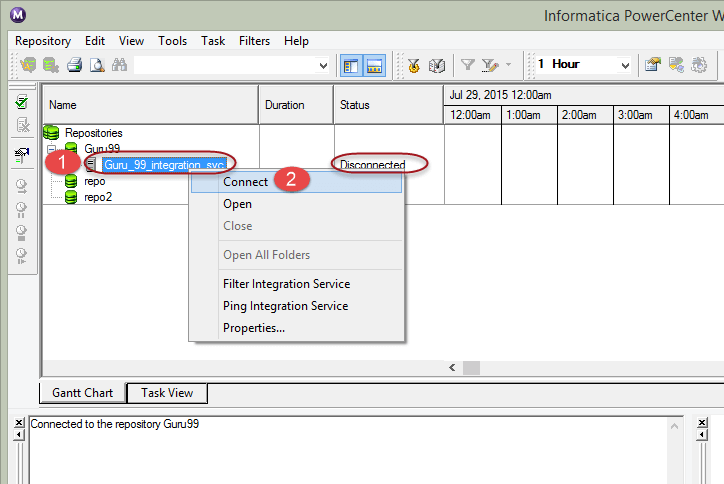
**Step 2** – This will open workflow monitor window

In the workflow monitor tool, you will see the repositories and associated integration services on the left side. Under the status column, you will see whether you are connected or disconnected to integration service. If you are in the disconnected mode, you won’t see any running workflows. There is a time bar which helps us to determine how long it took a task to execute.

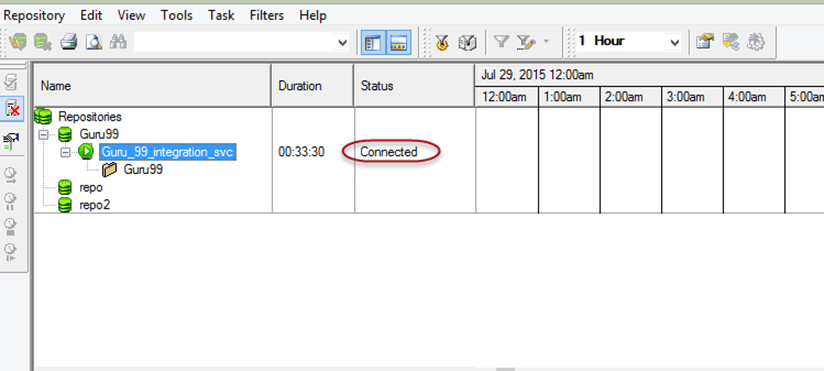


**Step 3** – The workflow monitor is in a disconnected mode by default. To connect to integration service.

1. Right click on the integration service
2. Select connect option



After connecting, the monitor will show the status as connected.



**Views in Workflow Monitor**

There are two types of views available in Informatica workflow monitor

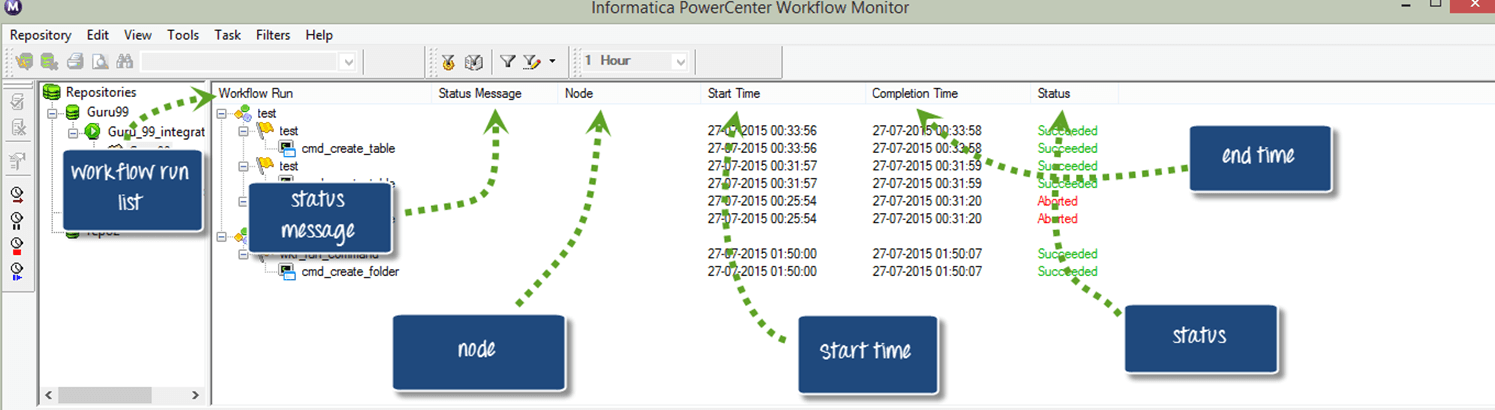
* Task view
* Gantt View

**Task View**

Task view displays the workflow runs in report format, and it is organized by workflow runs. It provides a convenient approach to compare workflow runs and filter details of workflow runs.

Task view shows the following details

* Workflow run list – Shows the list of workflow runs. It contains folder, workflow, worklet, and task names. It displays workflow runs in chronological order with the most recent run at the top. It displays folders and Integration Services alphabetically.
* Status message – Message from the Integration Service regarding the status of the task or workflow.
* Node – Node of the Integration Service executed the task.
* Start time – The time at which task or workflow started.
* Completion time – The time at which task or workflow completed the execution.
* Status – Shows status of the task or workflow, whether the workflow started, succeeded, failed or aborted.



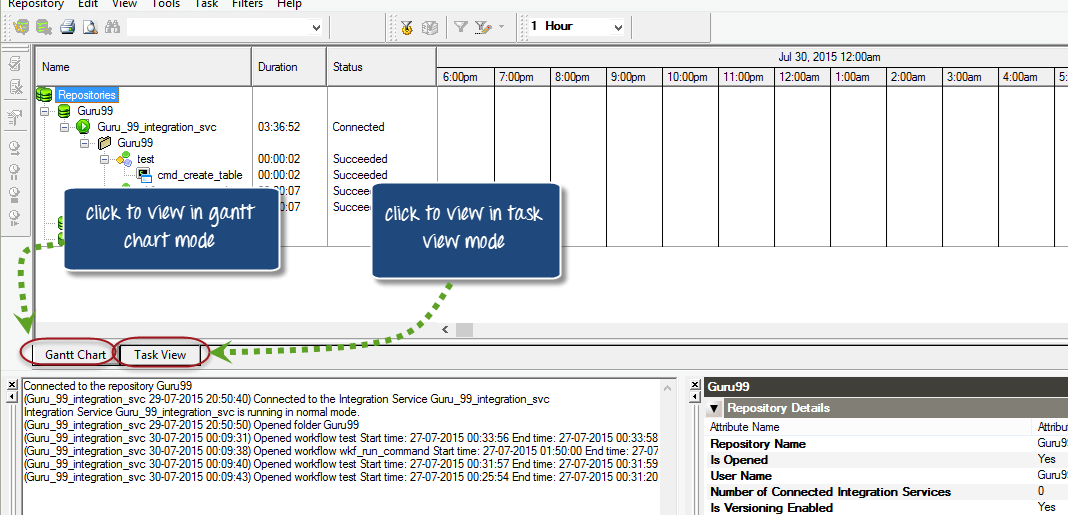
**Gantt Chart View**

In Gantt chart view, you can view chronological view of the workflow runs. Gantt chart displays the following information.

* Task name – Name of the task in the workflow
* Duration – The time taken to execute the task
* Status – The most recent status of the task or workflow

To switch between Gantt chart and task views

To switch from Gantt chart to Task view or vice versa, click on the respective button as shown in the screenshot to change the mode.



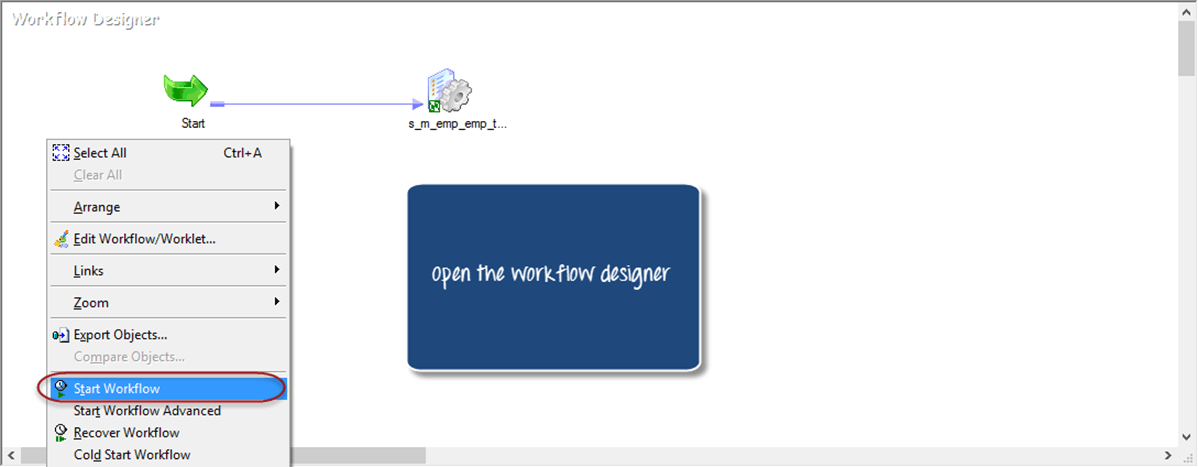
**Example- How to monitor and view details**

In previous examples, we have created a

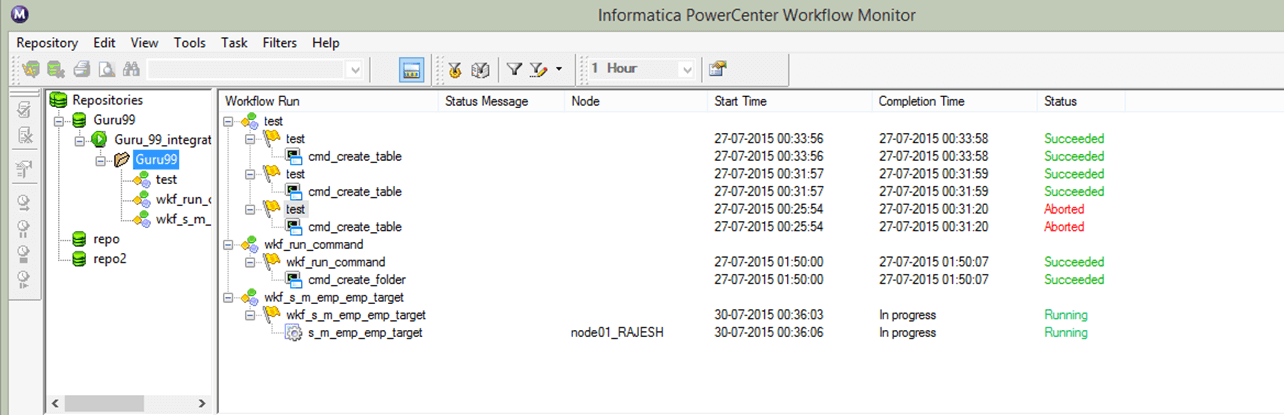
* **Mapping “m\_emp\_emp\_target”:** A mapping is a set of instructions on how to modify the data and processing of transformations that affects the record set.
* **Session “s\_ m\_emp\_emp\_target”** : A session is a higher level object to a mapping which specifies the properties of execution. For example performance tuning options, connection details of sources/targets, etc.
* **Workflow “wkf\_s\_m\_emp\_emp\_target”**: A workflow is a container for the session and other objects, and it defines the timing of the execution of tasks and the dependency or flow of execution.

Now, we will analyze the details of execution in this topic.

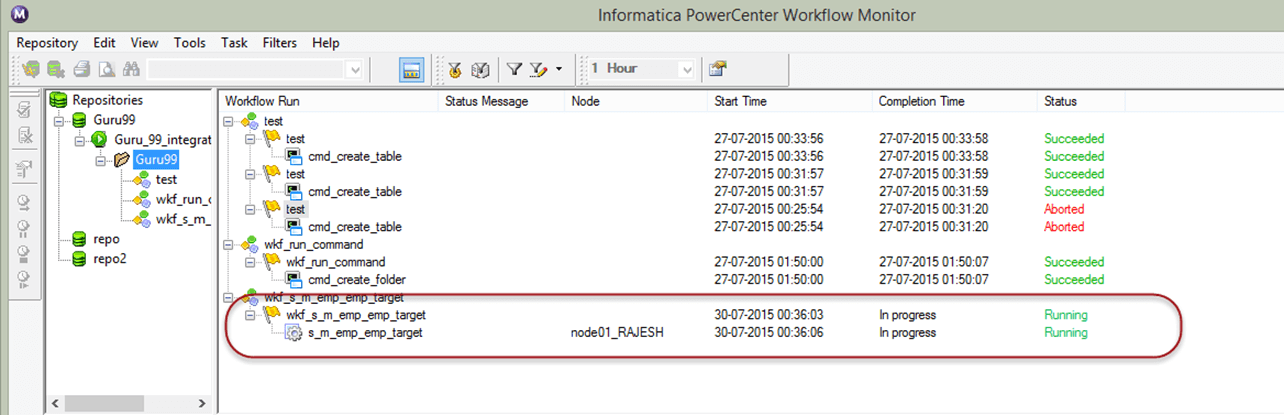
**Step 1** – Restart the workflow designer, as described in previous topic



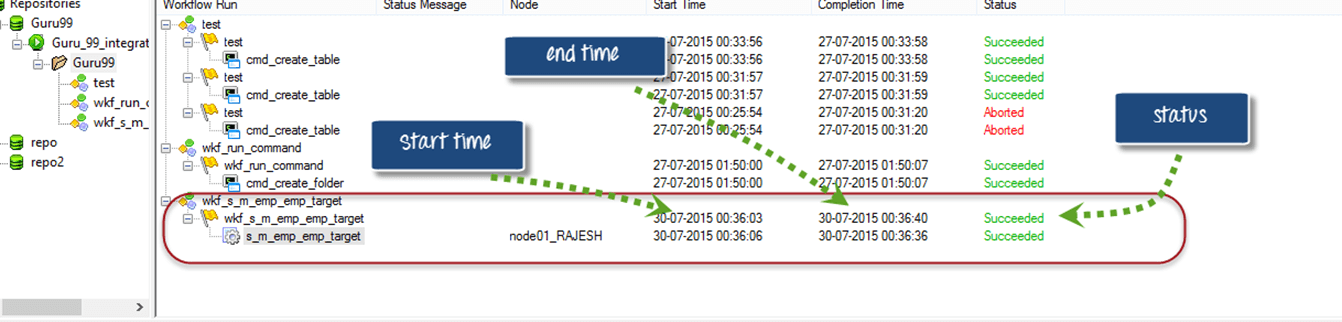
**Step 2** – Go to workflow monitor and in the monitor window you will see details as shown in screen shot like repository, workflow run details, node details, workflow run start time, workflow run completion time and status.



**Step 3** – Here you can view the currently running workflow, which is having status as “running”.

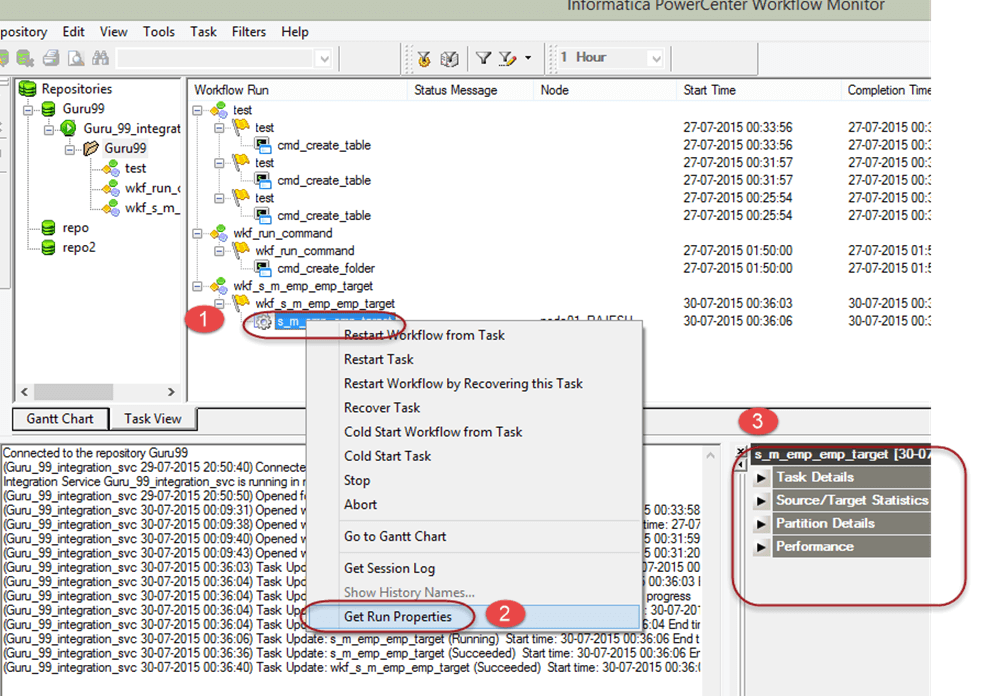


**Step 4** – Once the workflow execution completes, it status would change to succeeded/failed along with start and end time details.



**Step 5** – To view the task details

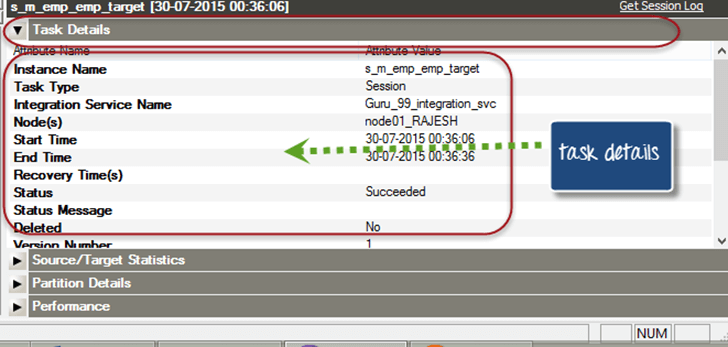
1. Right click on task name
2. In the pop-up window select “get run properties”
3. A properties window would appear with the task details



**Step 6** – Click on each of the menu of the properties window to view specific details.

Here we chose “Task Details” to view. It will display all the details like Instance Name, Task Type, Start Time, Integration Service Name, etc.

Task details –

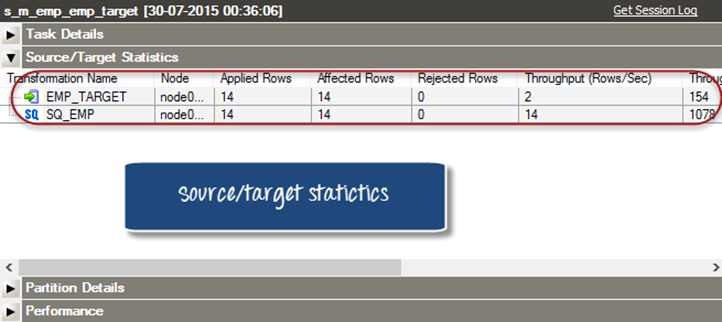


**Source and Target Statistics**

Source and target statistics gives the details of source and target. For example, how many rows are fetched from the source and how many rows are populated in the target the current throughput, etc

In the following screen, 14 records are fetched from the source, and all 14 are populated in the target table.

* **Applied rows** signify how many records Informatica had tried to update or insert the target
* **Affected rows** signify how many numbers of applied rows were actually succeeded.Here all 14 rows are successfully loaded in the target, so the count is same for both.
* **Rejected rows** signify how many rows are dropped due to target constraint or other issues.



In this tutorial, you have learned how to open and monitor the workflows and tasks using workflow monitor.

# Session Properties in Informatica: Complete

Session property is a set of instructions that instructs Informatica how and when to move the data from source to targets.

A session property is a task, just like other tasks that we create in workflow manager. Any session you create must have a mapping associated with it.

A session can have a single mapping at a time and once assigned, it cannot be changed. To execute a session task, it must be added to a workflow.

A session can be a reusable object or non-reusable. When you create a session in task developer, then it can be reused, but when you create a session in workflow designer, then it is non-reusable.

A reusable session can be added to multiple workflows.

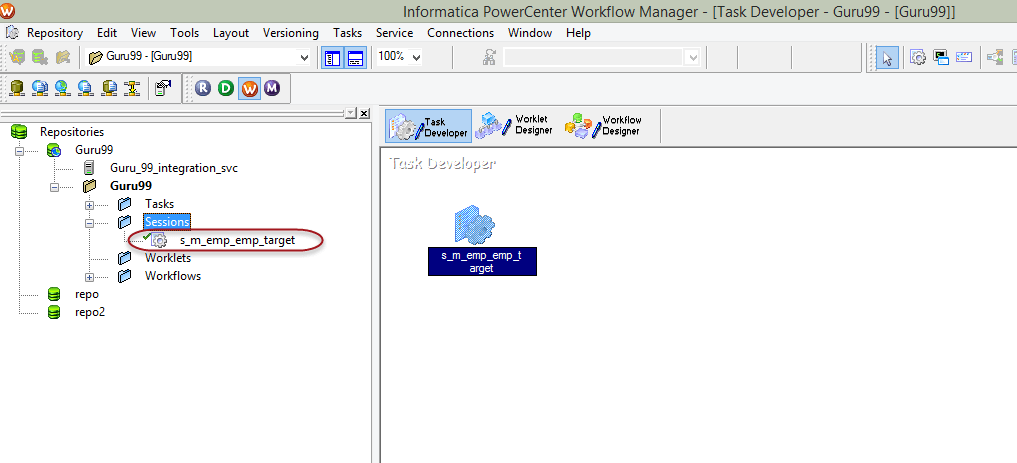
## Properties Of Session

Using the properties of the session you can configure various characteristics of the session like pre and post[SQL](https://www.guru99.com/sql.html)scripts, log file name and path, memory properties, etc.

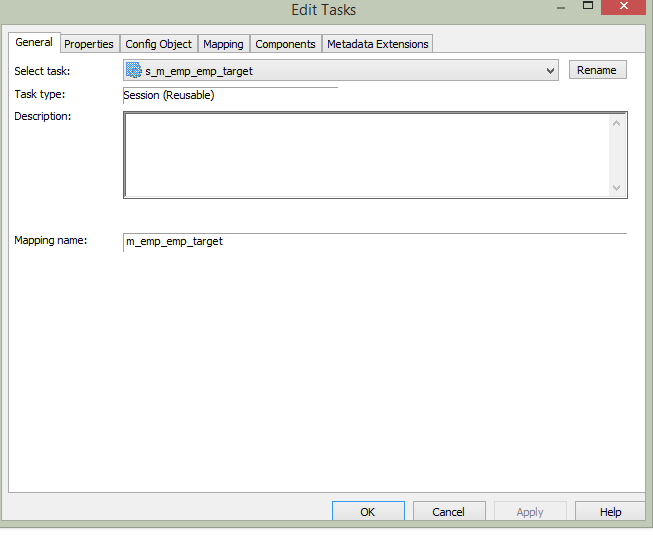
You can also override mapping properties in the session properties. In this section, we will discuss the following important properties of the session.

* Treat source rows as
* Commit Interval
* Session log file Property
* Test Load Property
* Log options
* Error Handling
* Source/Target Properties
* Connections

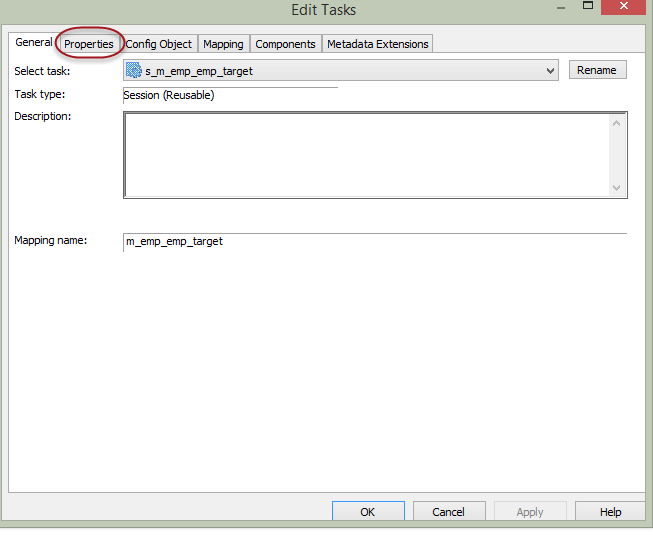
**Step 1)**Open the session “s\_m\_emp\_emp\_target” in task developer, which we created in the earlier tutorial.



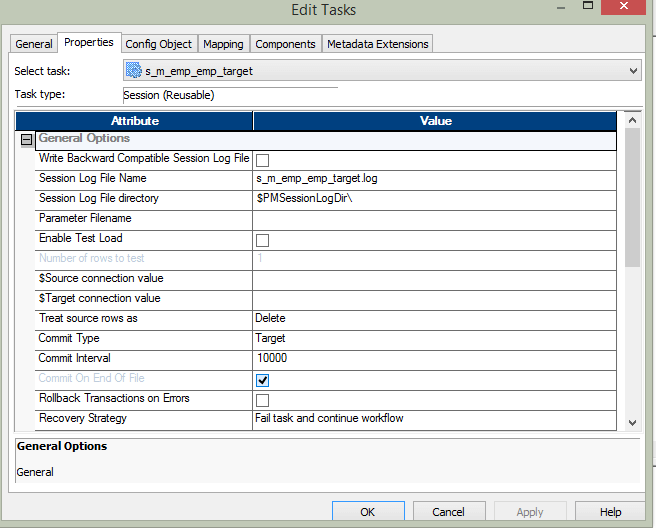
**Step 2)**Double click on the session icon inside Task Developer to open edit task window.



**Step 3)**Inside the “Edit Task” window clicks on the properties tab.



**Step 4)**In properties tab, it will show the properties of the session

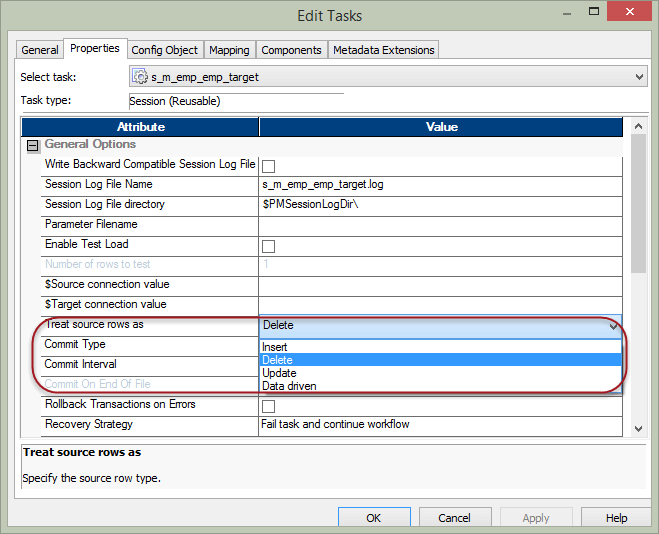


**Treat Source Rows As Property**

This property allows you to define how the source data affects the target table. For example, you can define that the source record should be inserted or deleted from the target.

This property has four options –

* Insert
* Update
* Delete
* Data-driven



* When this property is set to **insert**, the source data will be marked to be inserted. It means the data will only be inserted.
* When the property is set to **update**, the target data will be updated by the source data. For updating of data primary key needs to be defined in the target table.
* When property is set to **delete** the source data which is already present in the target will be deleted from the target table. For this property to execute and apply the changes, the primary key should be defined in the target table.
* With the property set to **data driven**, the Informatica checks what source records are marked. If in a mapping the source records are marked as insert then records will be inserted into the target. If records are marked as an update in the mapping, then the records will be updated in the target. So what operation will be performed at the target depends on how records are handled inside the mapping.

**How To Make Treat Source Rows – Delete**

**Step 1 –**

1. In the property tab of the session task, select “Delete” option in “Treat Source Rows as”
2. Select OK Button

**Step 2** – To define primary key in target table, open Informatica designer

1. Open target designer
2. Double click on the “emp\_target” table

This will open an “Edit Table” for our target table.

**Step 3** – In the edit window of target table

1. For the EmpNo column, select key type as “primary key” from the scroll down menu and
2. Select OK button.

**Step 4 –** Save the changes in Informatica and execute the workflow for this mapping.

When you execute this mapping, **the source records which are already present in the target will get deleted.**

**Commit Interval – Property**

This property defines the interval after which the Informatica makes commit operation to the target table.

For example, if you are inserting 20,000 records in a target table, and you define commit interval as 5,000, then after every 5,000 insertions of records in the target, a commit operation will be performed.

**Session Log File Name & Session Log File directory**

Configure this property to modify

* Default session log file name and
* Path of the log file

The $PMSessionLogDir\ is an Informatica variable and in windows it points to the following default location “C:\Informatica\9.6.1\server\infa\_shared\SessLogs”.

**Enable Test Load**

Using this property, you can test your session and mappings. When you use this feature and execute the sessions, records are fetched from the sources but they don’t get loaded into the target. So this feature helps in[Testing](https://www.guru99.com/software-testing.html)the correctness of mappings, parameter files, functioning of various transformations inside the mapping.

If you enable this feature, then there is another property – No of Rows to Test, this property should be configured for the no of records which you want to be fetched from the source for the test load.

**Memory Properties**

Memory properties give us the flexibility to fine tune the memory allocated to the Informatica for performance optimizations. When there are high bottleneck and performance is poor then you can try to improve the performance using the memory properties.

To configure memory properties click on the “config object” tab of the edit task window. It will open another window where you can configure the changes.

In this section, you can configure the memory properties. For example, default buffer block size, sequential buffer length, etc. Changes to this properties will determine how much memory should be allocated to Informatica services for their operation.

**Log options**

In this property section, you can configure the log properties of the session. You can set the no for how many no of logs you want to save for a session, session log file max size.

**Error Handling**

In this section, you can configure the error properties for the session.

Using **Stop on errors** you can configure after how many errors the session has to be stopped.

Using **override tracing** you can override the mapping tracing levels.

You can also configure the behaviour of the session for various errors encountered for example stored procedure error, pre-post SQL error, etc.

**Mapping and Source/Target Properties**

In the mapping tab of the edit task window of the session, you can configure the properties related to the mapping and its sources/targets and transformations. With this section of properties, you can override the properties of the source and targets. For the sources and targets we can override the table names, we can apply table name prefixes. Inside the mappings we can configure the properties of various transformations, sources and targets, in addition to that in this section we can also review and override those properties. It’s like a single place to review all those properties.

**Connection Properties in Mapping**

Using this property, you can define database connections for the source and targets.

**Source Properties**

In this section, you can configure the properties related to the source of the mapping. You can configure pre and post SQL scripts for the source.

Using SQL query property, you can override the SQL for the source. You can also override the source table name in this section.

**Target Properties**

In this section, you can configure the details of the target. You can define whether target load has to be a bulk load or a normal mode.

In bulk load, the performance gain is achieved as during the load there are no redo log buffers managed by the database.

On the other hand, normal load is slower as compared to bulk load, but in case of failure database recovery is possible.

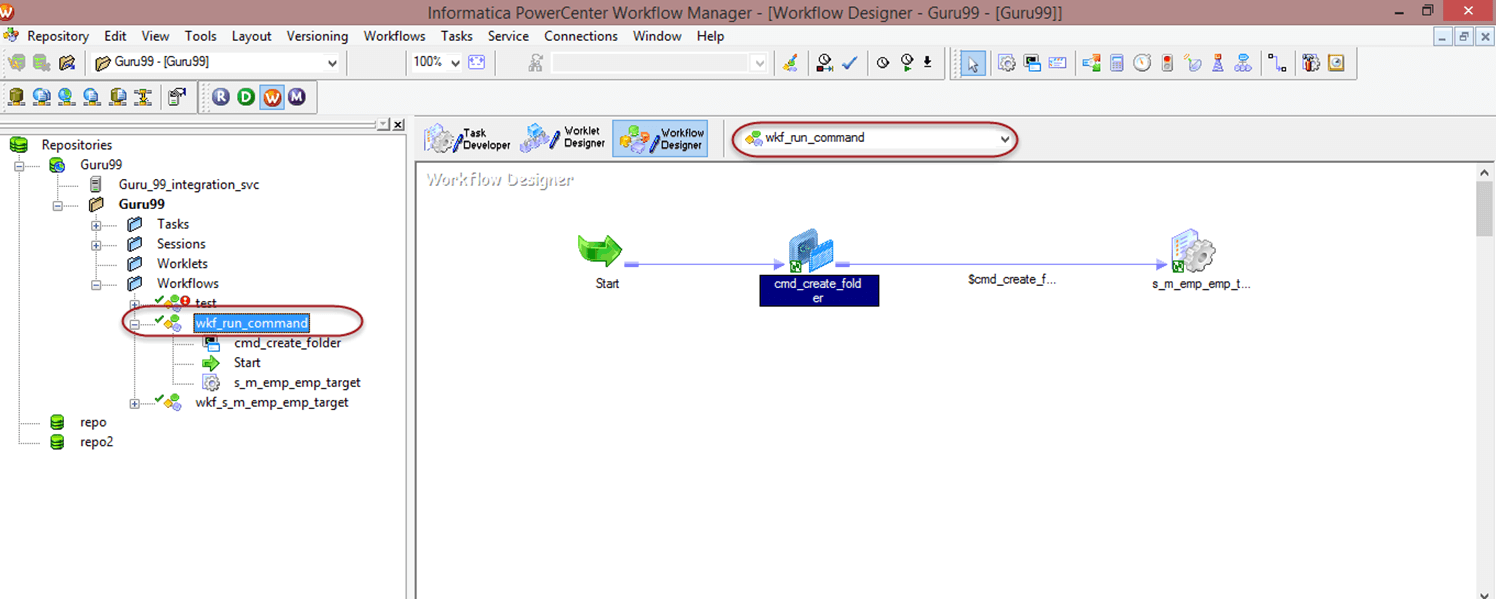
You can also define the property to truncate the target table before populating it. It means before loading any records in the target, the target table will be truncated, and then the load will be performed. This property is useful when we create mappings for stage load.

We can also define target table pre SQL and post SQL. Pre SQL is the piece of SQL code which will be executed before performing insert in the target table, and post SQL code will be executed after the load of target table is completed.

**Success or Failure Of Session Task**

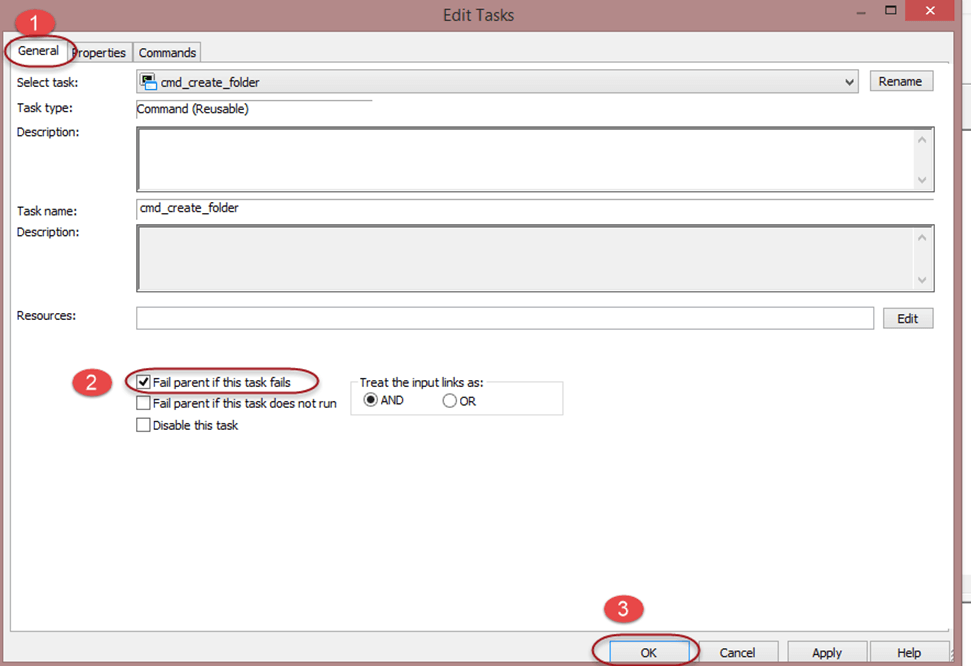
When you have multiple sessions inside a workflow, then there can be a scenario where one or more session fails. In such condition, there comes a question of what would be the status of the workflow because you are having a workflow in which few tasks have failed, and few task got succeeded. To handle such conditions, Informatica provides the option to set this failure specific property inside the workflow. To configure such behaviour –

**Step 1** – Open the workflow “wkf\_run\_command”, which we created earlier



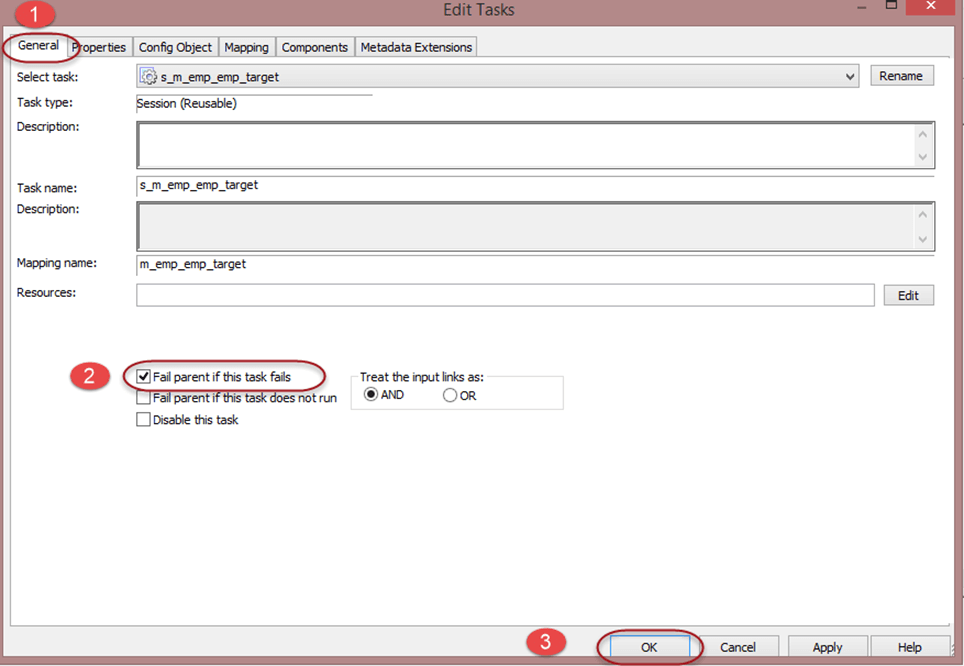
**Step 2** – Double click on the command task, this will open edit task window

1. Select the general tab
2. Select the check box against property “fail parent if this task fails”
3. Select Ok Button



**Step 3** – Double click on the session task, this will open edit task window

1. Select the general tab
2. Select the check box against property “fail parent if this task fails”
3. Select Ok Button



**Step 4** – Save the changes using a ctrl+s shortcut.

When you execute this workflow after making the above changes if any of the tasks fails the workflow status will be made as failed. so you can identify that during the execution of your workflow some of its task has failed.

**What is Transformation?**

Transformations is in Informatica are the objects which creates, modifies or passes data to the defined target structures (tables, files or any other target).

The purpose of the transformation in Informatica is to modify the source data as per the requirement of target system. It also ensures the quality of the data being loaded into the target.

Informatica provides various transformations to perform specific functionalities.

For example, performing tax calculation based upon source data, data cleansing operation, etc. In transformations, we connect the ports to pass data to it, and transformation returns the output through output ports.

In this tutorial- you will learn

* [Classification of Transformation](https://www.guru99.com/introduction-transformations-informatica-and-filter-transformation.html#1)
* [Filter Transformation](https://www.guru99.com/introduction-transformations-informatica-and-filter-transformation.html#2)

**Classification of Transformation**

Transformation is classified into two categories, one based on connectivity, and other based on the change in no of rows. First we will look the transformation based on connectivity.

Types of transformation based on connectivity

* **Connected** Transformations
* **Unconnected** Transformations

In Informatica, during mappings the transformations which are connected to other transformations are called connected transformations.

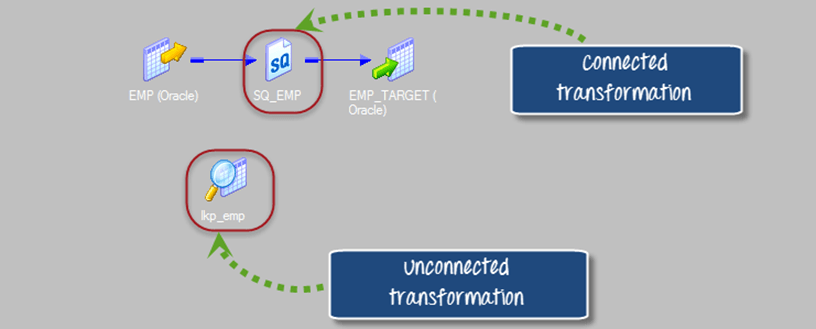
For example, Source qualifier transformation of Source table EMP is connected to filter transformation to filter employees of a dept.

Those transformations that are not connected to any other transformations are called unconnected transformations.

Their functionality is used by calling them inside other transformations like Expression transformation. These transformations are not part of the pipeline.

The connected transformations are preferred when for every input row, transformation is called or is expected to return a value. For example, for the zip codes in every row, the transformation returning city name.

The unconnected transformations are useful when their functionality is only required periodically or based upon certain conditions. For example, calculation the tax details if tax value is not available.



Types of transformations based on the change in no of rows

* Active Transformations
* Passive Transformations

Active Transformations are those who modifies the data rows and the number of input rows passed to them. For example, if a transformation receives ten number of rows as input, and it returns fifteen number of rows as an output then it is an active transformation. The data in the row is also modified in the active transformation.

Passive transformations are those who does not change the number of input rows. In passive transformations the number of input and output rows remain the same, only data is modified at row level.

In the passive transformation, no new rows are created, or existing rows are dropped.

**Following is the List of Transformations in Informatica**

* Source Qualifier Transformation
* Aggregator Transformation
* Router Transformation
* Joiner transformation
* Rank Transformation
* Sequence Generator Transformation
* Transaction Control Transformation
* Lookup and Re-usable transformation
* Normalizer Transformation
* Performance Tuning for Transformation
* External Transformation
* Expression Transformation

**What is Filter Transformation?**

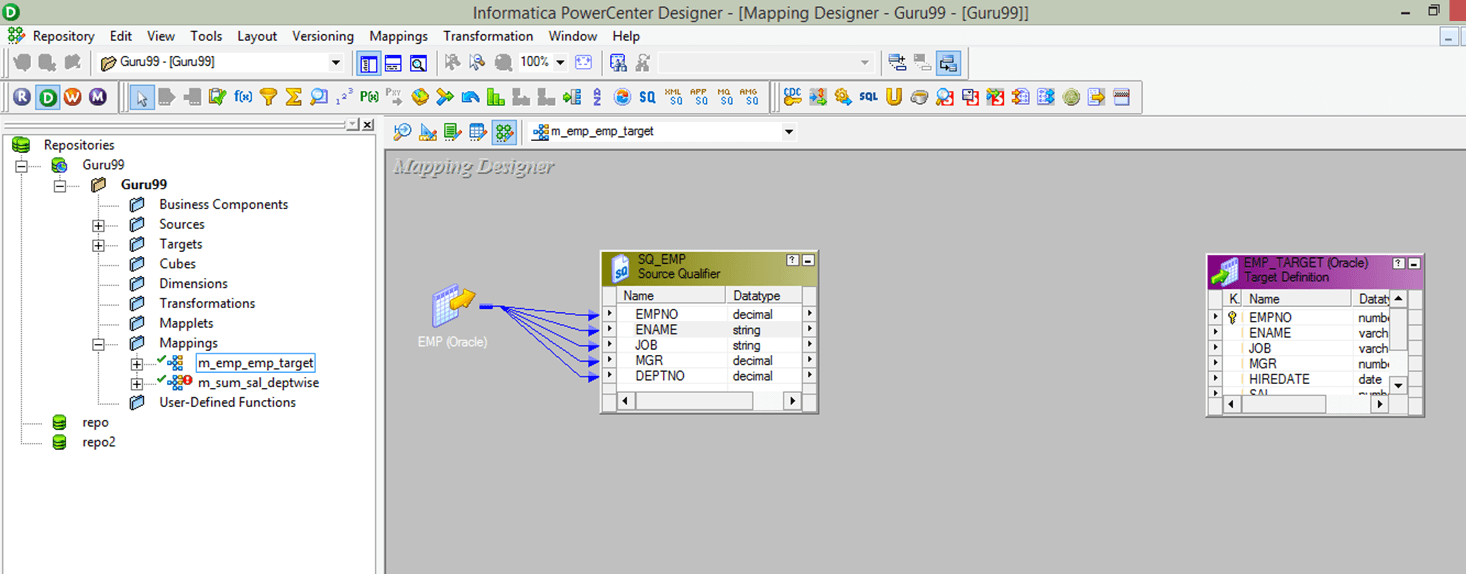
Filter Transformation is an active transformation as it changes the no of records.

Using the filter transformation, we can filter the records based on the filter condition. Filter transformation is an active transformation as it changes the no of records.

For example, for loading the employee records having deptno equal to 10 only, we can put filter transformation in the mapping with the filter condition deptno=10. So only those records which have deptno =10 will be passed by filter transformation, rest other records will be dropped.

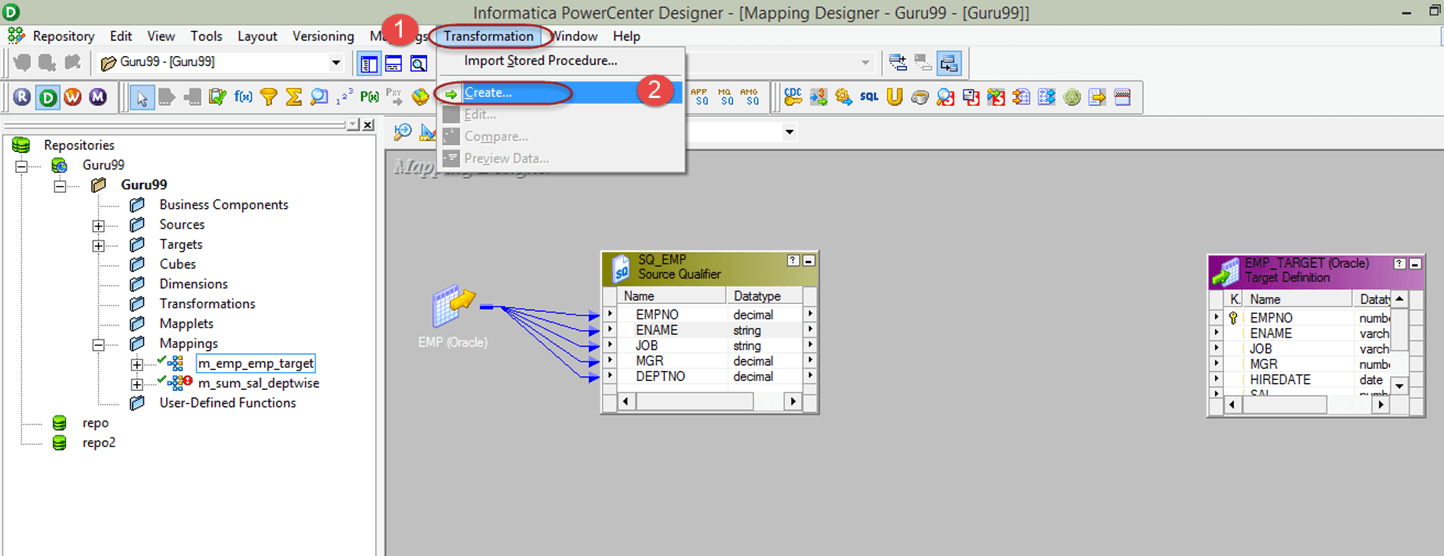
How to use filter transformation-

**Step 1** – Create a mapping having source “EMP” and target “EMP\_TARGET”



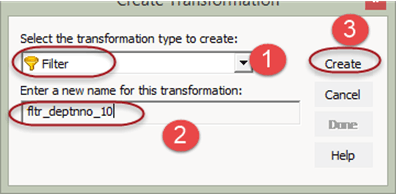
**Step 2** – Then in the mapping

1. Select Transformation menu
2. Select create option

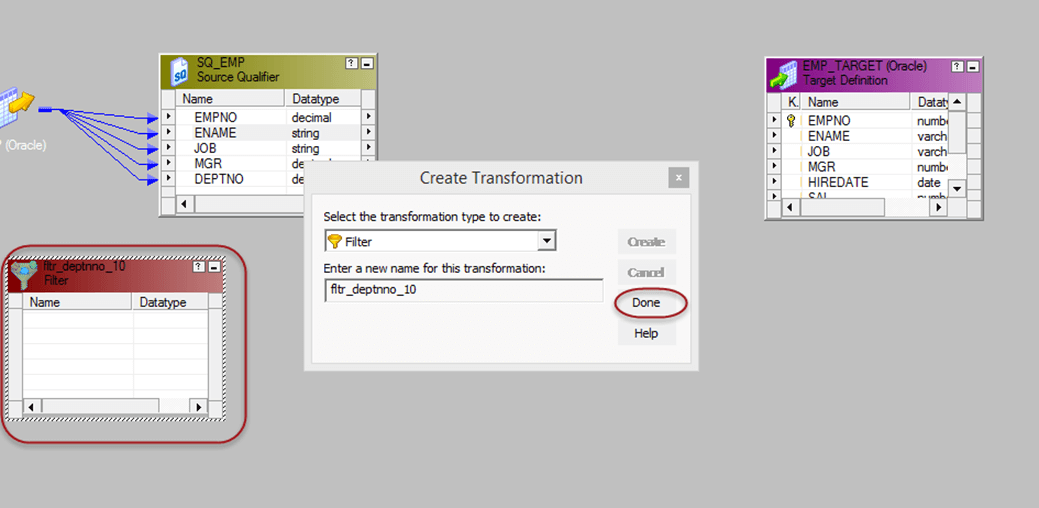


**Step 3** – Then in the create transformation window

1. Select Filter Transformation from the list
2. Enter Transformation name “fltr\_deptno\_10”
3. Select create option

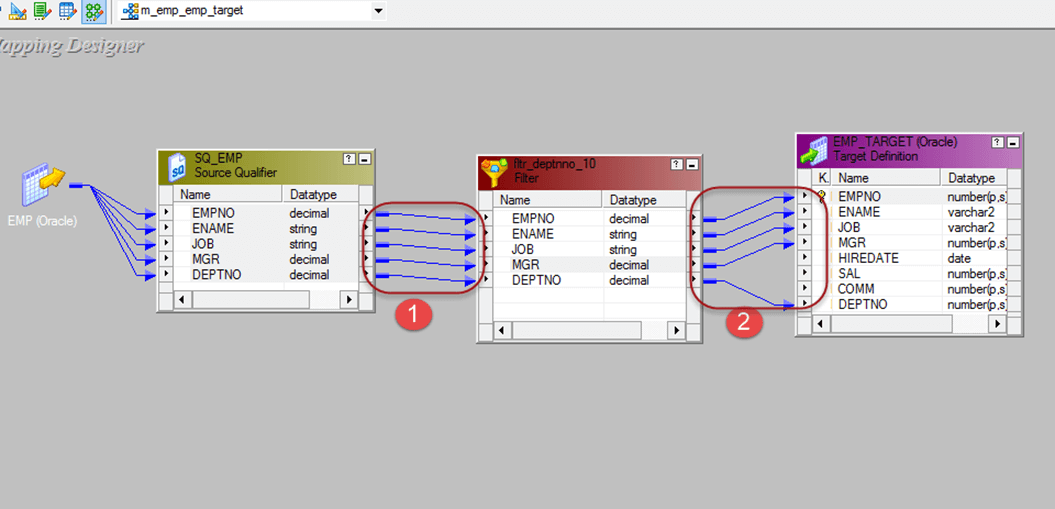


**Step 4** – The filter transformation will be created, Select “Done” button in the create transformation window



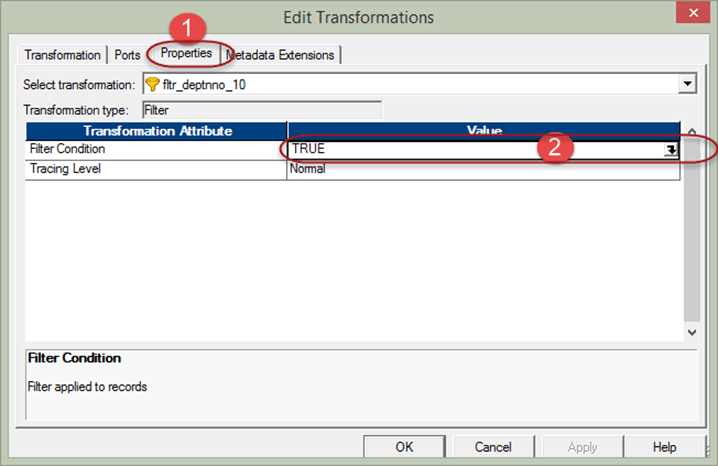
**Step 5** – In the mapping

1. Drag and drop all the Source qualifier columns to the filter transformation
2. Link the columns from filter transformation to the target table



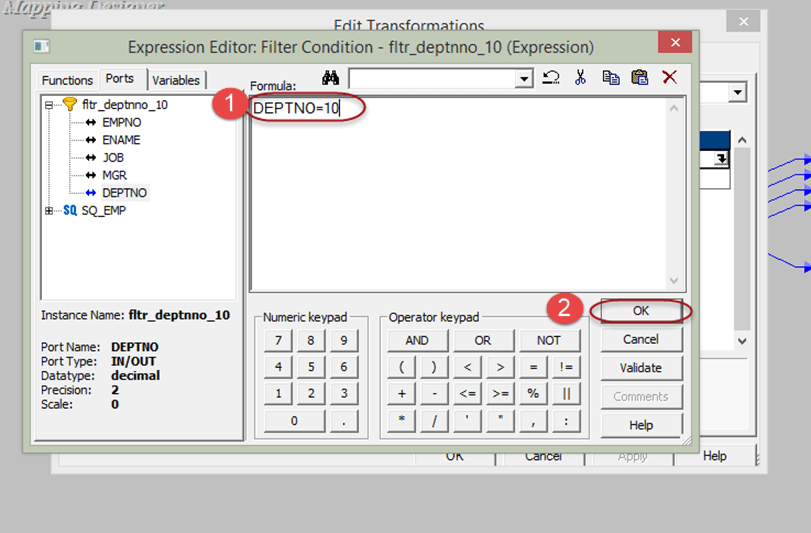
**Step 6** – Double click on the filter transformation to open its properties, and then

1. Select the properties menu
2. Click on the Filter condition editor

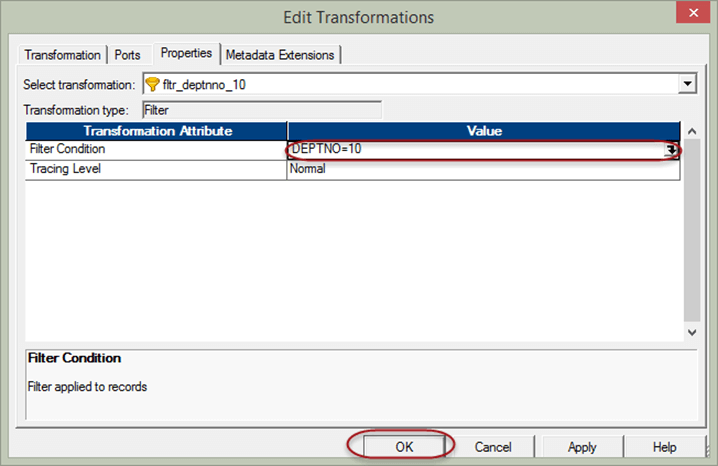


**Step 7** – Then in the filter condition expression editor

1. Enter filter condition – deptno=10
2. Select OK button



**Step 8** – Now again in the edit transformation window in Properties tab you will see the filter condition, select OK button



Now save the mapping and execute it after creating session and workflow. In the target table, the records having deptno=10 only will be loaded.

In this way, you can filter the source records using filter transformation.