

Informatica PowerCenter 9

What is Informatica PowerCenter Tool?

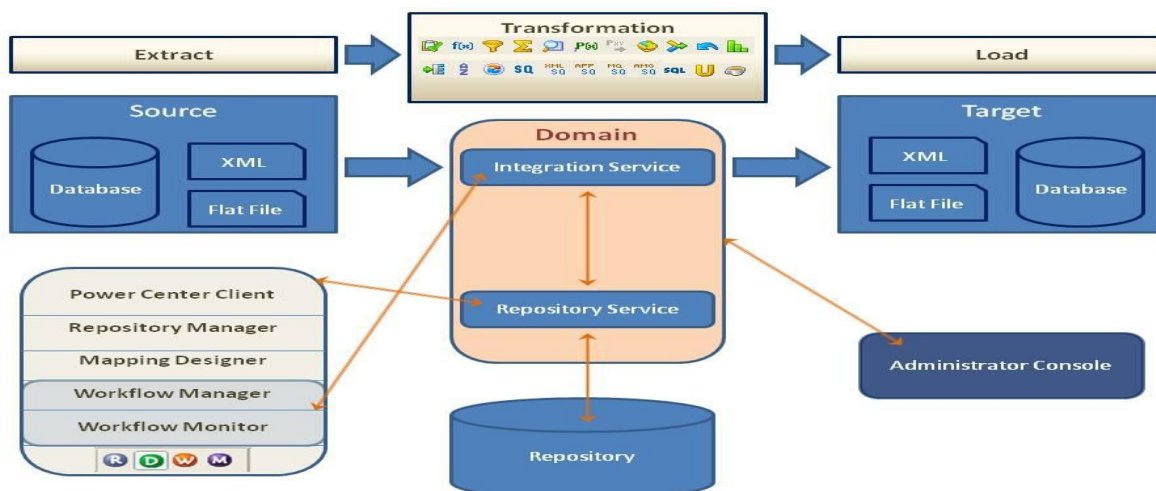
Informatica is a data integration/ETL tool. It gathers data from various sources and loads it into distinctive targets. It just moves data from one place to another, but does not store anything. Informatica utilizes a service oriented architecture that carries the ability to promote services and resources over various machines.

Informatica Powercenter Components

Informatica PowerCenter provides an environment that allows to load data into a centralized location, such as a data mart, data warehouse, or operational data store (ODS). Informatica provides the following integrated components:

1. **PowerCenter domain.** The Power Center domain is the primary unit for management and administration within PowerCenter. The Service Manager runs on a PowerCenter domain.
2. **PowerCenter repository.** The PowerCenter repository resides in a relational database. The repository database tables contain the instructions required to extract, transform, and load data.
3. **Administration Console.** The Administration Console is a web application that you use to administer the PowerCenter domain and PowerCenter security.
4. **PowerCenter Client.** The PowerCenter Client is an application used to define sources and targets, build mappings and mapplets with the transformation logic, and create workflows to run the mapping logic.
5. **Repository Service.** The Repository Service accepts requests from the PowerCenter Client to create and modify repository metadata and accepts requests from the Integration Service for metadata when a workflow runs.
6. **Integration Service.** The Integration Service extracts data from sources and loads data to targets.

Informatica Architecture



Informatica Power Center 9 Architecture (Framework)

Informatica PowerCenter Client Tools

The Power Center Client consists of the following applications that we use to manage the repository, design mappings, mapplets, and create sessions to load the data:

1. PowerCenter Designer

Use the Designer to create mappings that contain transformation instructions for the Integration Service. The Designer has the following tools that you use to analyze sources, design target Schemas, and build source-to-target mappings:

- **Source Analyzer:** Import or create source definitions.
- **Target Designer:** Import or create target definitions.
- **Transformation Developer:** Develop transformations to use in mappings.
- **Mapplet Designer:** Create sets of transformations to use in mappings.
- **Mapping Designer:** Create mappings that the Integration Service uses to Extract, Transform, and load data.

2. Repository Manager

Use the Repository Manager to administer repositories. Common task we perform are Manage users and groups, Perform folder functions, View metadata.

3. Workflow Manager

Use the Workflow Manager to create, schedule, and run workflows. A workflow is a set of instructions that describes how and when to run tasks related to extracting, transforming, and loading data. The Workflow Manager has the following tools

- **Task Developer:** Create tasks we want to accomplish in the workflow.
- **Work let Designer:** Create a worklet in the Worklet Designer. A worklet is an object that groups a set of tasks. A worklet is similar to a workflow, but without scheduling information. We can nest worklets inside a workflow.
- **Workflow Designer:** Create a workflow by connecting tasks with links in the Workflow Designer. You can also create tasks in the Workflow Designer as you develop the workflow

4. Workflow Monitor

Use the Workflow Monitor to monitor scheduled and running workflows for each Integration Service. We can view details about a workflow or task in Gantt chart view or Task view. We can run, stop, abort, and resume workflows from the Workflow Monitor. We can view Sessions and workflow log events in the Workflow Monitor Log Viewer.

Transformations in Informatica

What is a Transformation?

A transformation is a repository object which reads the data, modifies the data and passes the data. Transformations in a mapping represent the operations that the integration service performs on the data

Transformations can be classified as active or passive, connected or unconnected.

List of Transformation

Transformation	Type	Description
Aggregator	Active/Connected	Performs aggregate calculations.
Expression	Passive/Connected	Calculates a value.
External Procedure	Passive/Connected or Unconnected	Calls a procedure in a shared library or in the COM layer of Windows.
Filter	Active/Connected	Filters data.
Mapplet Input	Passive/Connected	Defines mapplet input rows. Available in the Mapplet Designer
Joiner	Active/Connected	Joins data from different databases or flat file systems.
Lookup	Active or Passive/ Connected or Unconnected	Lookup and return data from a flat file, relational table, view, or synonym.
Normalizer	Active/Connected	Source qualifier for COBOL sources. Can also use in the pipeline to normalize data from relational or flat file sources.
Mapplet Output	Passive/Connected	Defines mapplet output rows. Available in the Mapplet Designer.
Rank	Active/Connected	Limits records to a top or bottom range.
Router	Active/Connected	Routes data into multiple transformations based on group conditions.
Sequence Generator	Passive/Connected	Generates primary keys.
Sorter	Active/Connected	Sorts data based on a sort key.
Source Qualifier	Active/Connected	Represents the rows that the Integration Service reads from a relational or flat file source when it runs a session.
SQL	Active or Passive/Connected	Executes SQL queries against a database.
Stored Procedure	Passive/Connected or Unconnected	Calls a stored procedure.
Transaction Control	Active/Connected	Defines commit and rollback transactions.
Union	Active/Connected	Merges data from different databases or flat file systems.
Update Strategy	Active/Connected	Determines whether to insert, delete, update, or reject rows.
XML Generator	Active/Connected	Reads data from one or more input ports and outputs XML through a single output port.
XML Parser	Active/Connected	Reads XML from one input port and outputs data to one or more output ports.
XML SourceQualifier	Active/Connected	Represents the rows that the Integration Service reads from an XML source when it runs a session

Slowly Changing Dimensions (SCD) – Types

Slowly Changing Dimensions: Slowly changing dimensions are the dimensions in which the data changes slowly, rather than changing regularly on a time basis.

For example, you may have a customer dimension in a retail domain. Let say the customer is in India and every month he does some shopping. Now creating the sales report for the customers is easy. Now assume that the customer is transferred to United States and he does shopping there. How to record such a change in your customer dimension?

The different types of slowly changing dimensions are explained below.

SCD Type 1: SCD type 1 methodology is used when there is no need to store historical data in the dimension table

SCD Type 2: SCD type 2 stores the entire history the data in the dimension table. With type 2 we can store unlimited history in the dimension table. In type 2, you can store the data in three different ways. They are Versioning, Flagging, and Effective Date.

SCD Type 3: In type 3 method, only the current status and previous status of the row is maintained in the table. To track these changes two separate columns are created in the table

Mapping Example

