DATA MIGRATION BETWEEN SERVERS USING BTEQ SCRIPTS IN INFORMATICA CLOUD

Abstract

This document presents an enterprise data migration approach using BTEQ (Basic Teradata Query) scripts orchestrated via Informatica Intelligent Cloud Services (IICS) Command Tasks. The solution illustrates how pre-processing logic can be applied to protect existing data in Teradata target tables before loading new records from source systems.

The implementation is based on real-world project experience and focuses on a controlled, repeatable, and auditable data loading process. It covers task configuration, execution flow, and validation techniques, demonstrating how Informatica Cloud and native Teradata utilities can be combined to ensure data integrity and operational reliability in hybrid cloud environments.

Hira Shahbaz

Data Engineer / ETL Developer
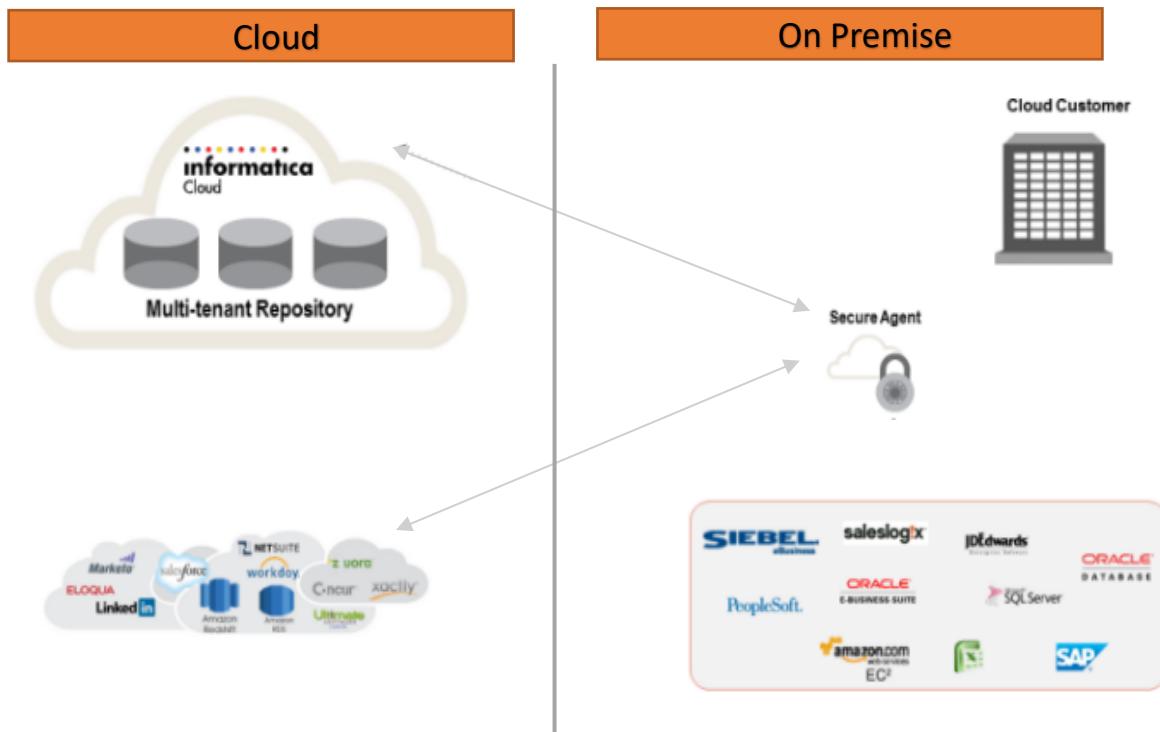
Contents

Informatica Cloud Services (ICS)

Informatica Cloud Services (ICS) is an on-demand cloud-based data integration and ETL platform designed to connect applications and data across cloud and on-premises environments. It is Informatica's cloud offering and provides many of the core capabilities available in its traditional on-premises products.

ICS enables users to design, execute, and monitor ETL processes through a web-based interface without requiring local development tools. It supports integration between various cloud applications and enterprise systems while offering centralized administration, scheduling, and monitoring. This makes it suitable for enterprise-scale data integration use cases where reliability and operational visibility are critical.

Overview of data loading into a Teradata target table using BTEQ scripts under Informatica Cloud Command Tasks

In this implementation, data is loaded into a Teradata target table using BTEQ scripts executed through Informatica Cloud Command Tasks. The approach is designed to ensure data safety, repeatability and controlled execution during the load process.

For this example, three tables are involved:

- Salesforce object – Source table
- Product2_tbl – Target table in Teradata
- Product2_bkp – Backup table created in Teradata using a BTEQ script

The Salesforce object serves as the source of incoming data. Before loading new data into the Teradata target table (Product2_tbl), a BTEQ script is executed to safeguard existing data.

The BTEQ script performs the following actions:

- Checks for the existence of the backup table (Product2_bkp) and drops it if present

- Creates a new backup table (Product2_bkp) based on the structure of the target table
- Copies existing data from Product2_tbl into Product2_bkp
- Deletes all records from the target table (Product2_tbl) after the backup is completed

Once the backup process is successfully executed, the BTEQ script completes its operation. The actual loading of the latest data from the Salesforce source into the Teradata target table (Product2_tbl) is then performed using an Informatica Cloud mapping task.
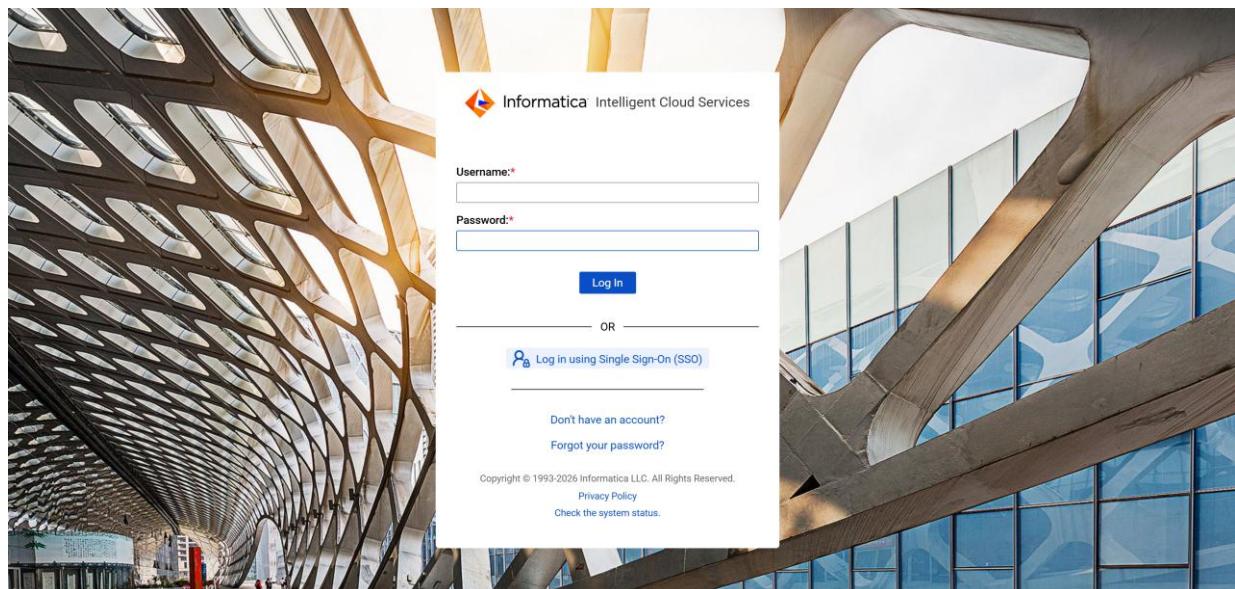
This two-step approach backup followed by load helps ensure data integrity and provides a rollback option in case of load failures or validation issues.

Steps to load data into a Teradata target table using BTEQ scripts under Informatica Cloud Command Tasks
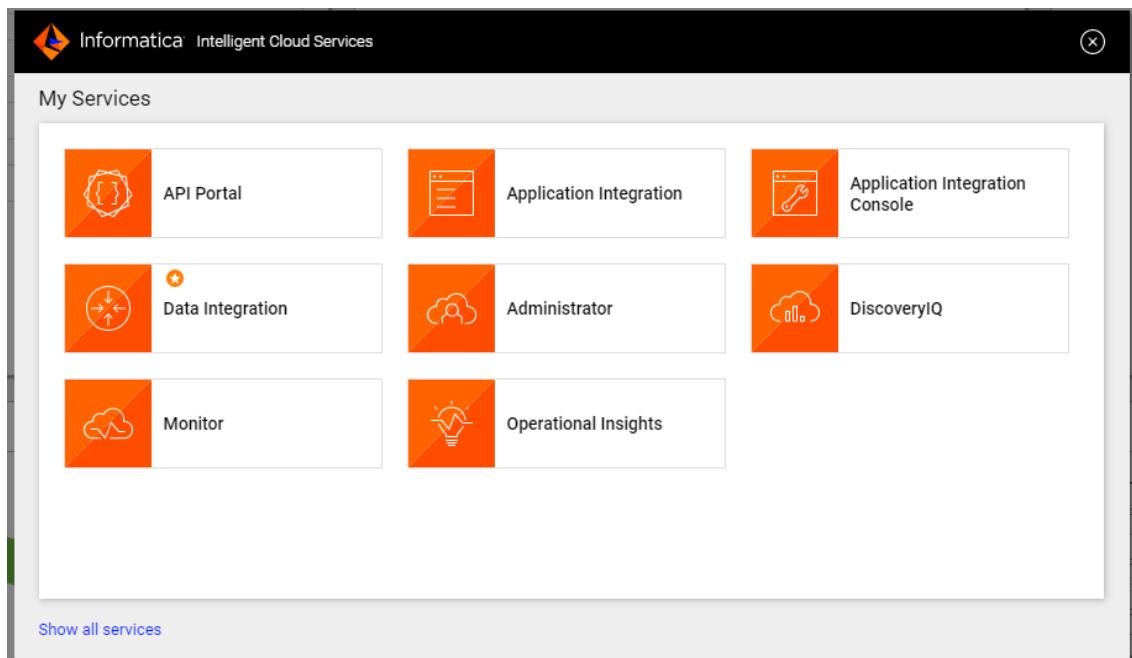
Step 1: Sign in to Informatica

Open the Informatica Cloud login page using the following URL and sign in with valid credentials: https://dm-us.informaticacloud.com/identity-service/home

After successful authentication, navigate to the Data Integration service. This will take you to the Informatica Cloud Data Integration dashboard, where integration tasks and mappings are created and managed.


Informatica Cloud login page
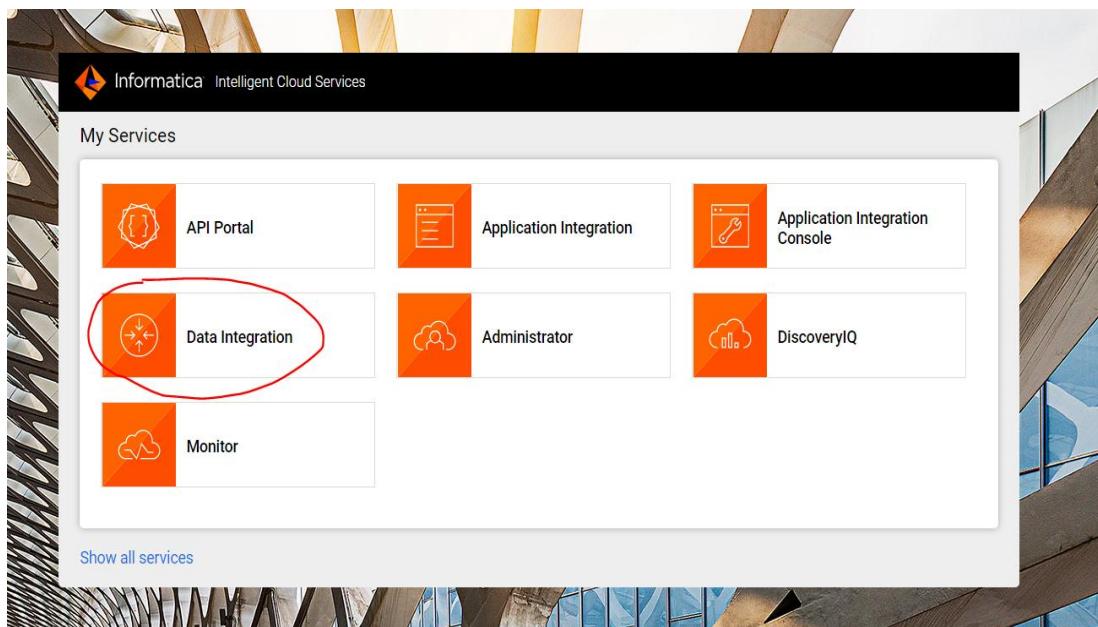
Landing dashboard

Informatica Cloud Data Integration Dashboard

The Informatica Cloud Data Integration dashboard provides access to all integration-related functionalities. It acts as the cloud-based equivalent of Informatica PowerCenter, delivering secure, scalable, and reliable data integration capabilities.

Informatica Cloud Data Integration enables organizations to design and execute global, distributed data warehouse and analytics workflows. It supports both cloud and on-premises data sources, making it suitable for hybrid integration scenarios.
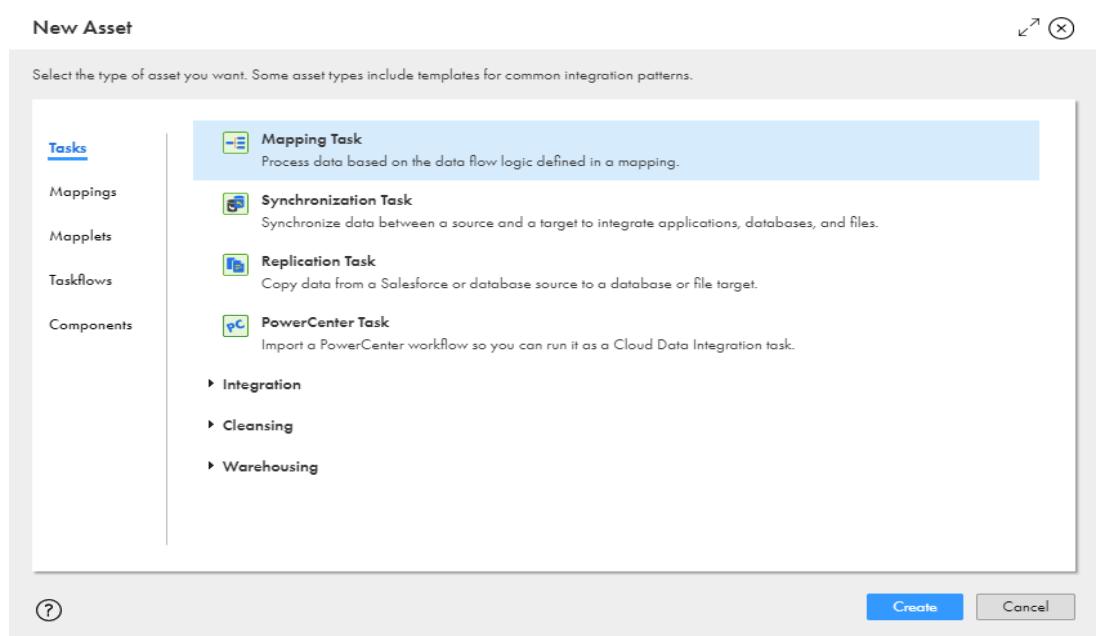
To begin creating a new task or mapping, click on Data Integration from the dashboard.

Data Integration dashboard

Creating a New Task or Mapping

To create a new mapping or task, click on the Task tab available on the screen.
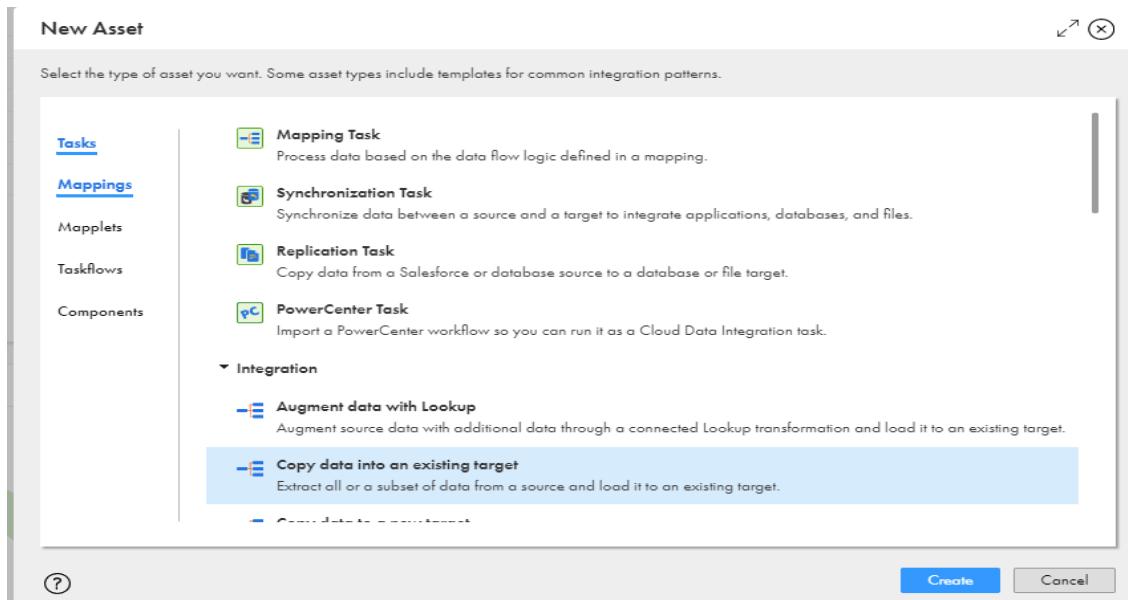

Click Tasks tab

Step 2: Task Creation

From the available integration task options, select Copy data into an existing target. This task type is used to extract all or a subset of data from a source system and load it into an already existing target table.

If the target table does not exist in the required database, it must be created prior to task execution.
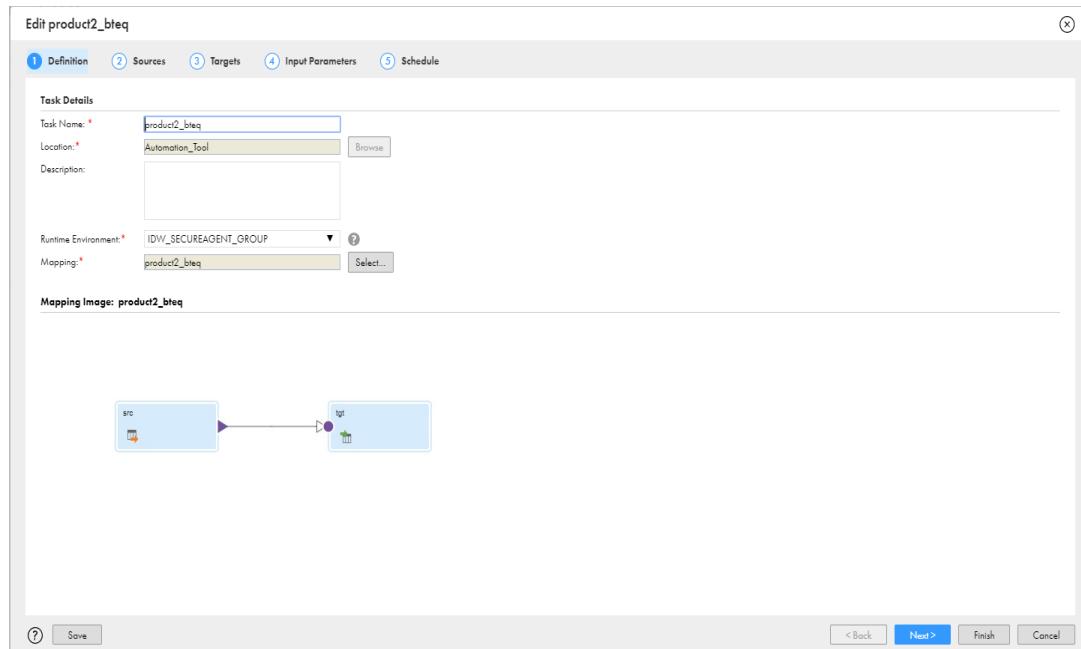
Task selection screen - Copy data into an existing target

Step 3: Task Configuration

Step 3.1: Definition

In the Definition section, provide an appropriate Task Name and Mapping Name based on project standards. For demonstration purposes, a dummy example is used where both the task name and mapping name are set to product2_bteq.

Select the required Runtime Environment which determines where the task will be executed. In this example, the runtime environment selected is: **IDW_SECUREAGENT_GROUP**

Task definition screen showing task name, mapping name and runtime environment

Step 3.2: Source

Configure the source properties as follows:
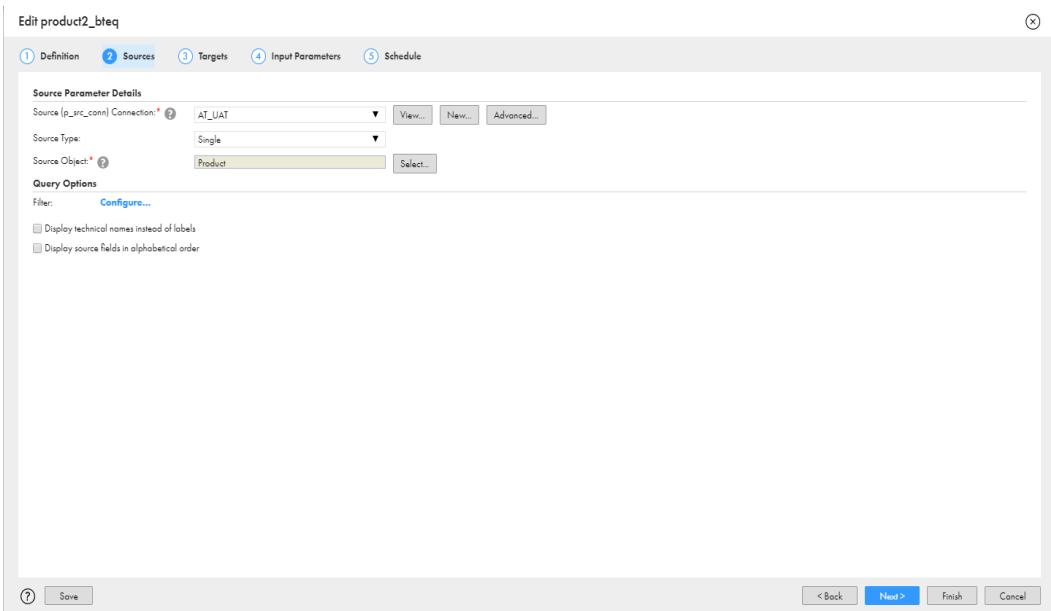
- **Source Connection:**
  Select the appropriate source connection. In this example, a Salesforce connection (AT_UAT) is used as the source.

- **Source Type:**
  Select Object.

- **Source Object:**
  Select the required source object. In this case, **Product2** is selected as the source table.
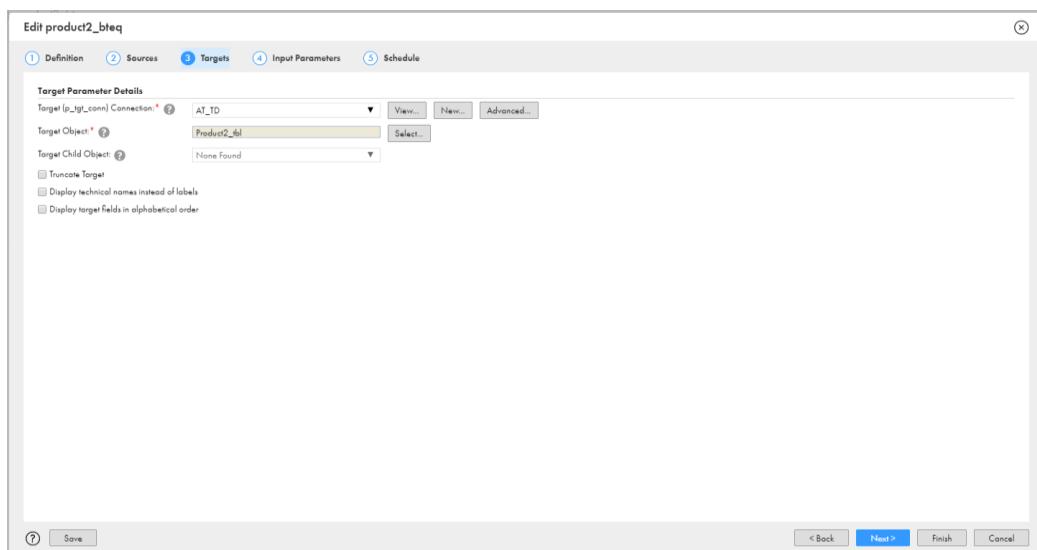
Source configuration screen

Step 3.3: Target

Configure the target properties as follows:

- **Target Connection:**
  Select the appropriate target connection pointing to the Teradata environment.

- **Target Type:**
  Select Object.

- **Object:**
  Select the target table into which data will be loaded. For demonstration purposes, Product2_tbl is selected as the target table.

After completing the target configuration, click Next to proceed.
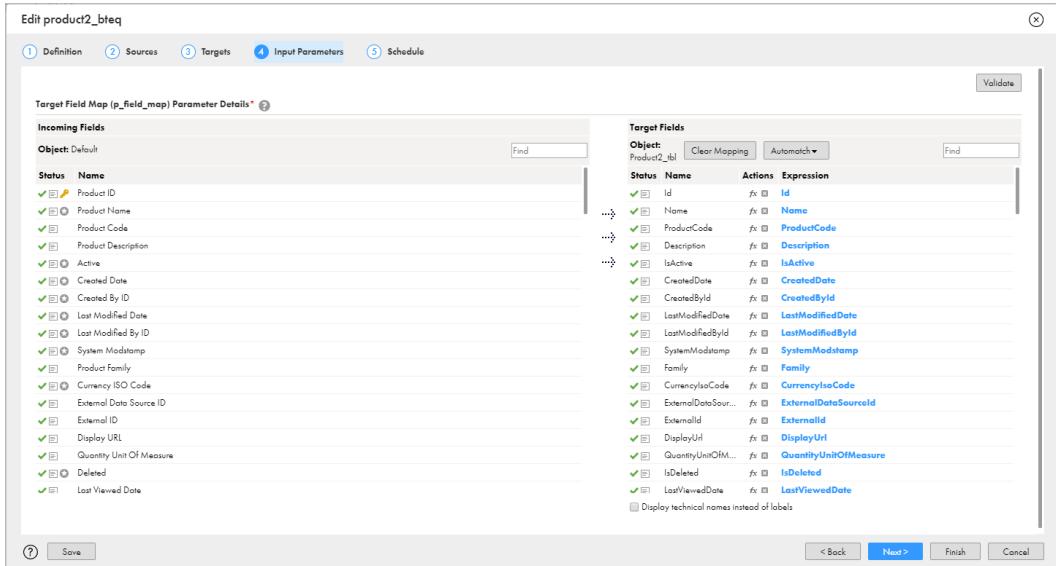
Target configuration screen

Step 3.4: Input Parameters

To configure field mappings, click on Field Mapping. Map the required fields from source to target. Remaining fields can be automated by the task if applicable.

After reviewing and confirming the mappings, click Next to continue.
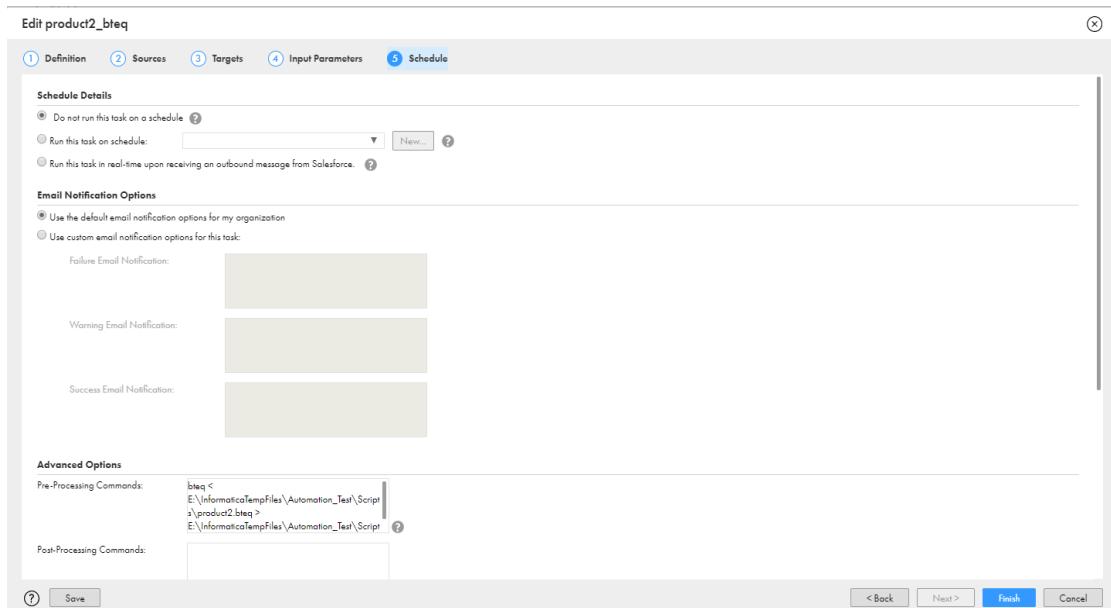
Field mapping screen

Step 3.5: Schedule

In the Schedule section, the task can be configured to run immediately or scheduled for execution later. Scheduling options include hourly, daily, or monthly execution based on business requirements.

If scheduling is not required, proceed by clicking Finish. The task configuration will be saved automatically.

Schedule configuration screen

Pre-processing and postprocessing commands

Informatica Cloud allows the use of pre-processing and post-processing commands to perform additional operations before or after the data load.

- Pre-processing commands run before the task reads data from the source.
- Post-processing commands run after data is written to the target.

Supported command types include:

- SQL commands for database-related operations
- Operating system commands such as shell or DOS commands

If any command in the pre-processing or post-processing step fails, the entire task execution fails. In this implementation, a BTEQ script is executed as a pre-processing command to back up and clean the target Teradata table before loading new data.

The BTEQ file is placed on the Secure Agent server at the following location: "E:\InformaticaTempFiles\Automation_Test\Scripts"

This file path is referenced in the pre-processing command configuration. During task execution, Informatica Cloud invokes the BTEQ script from this location and executes it on the Teradata database.
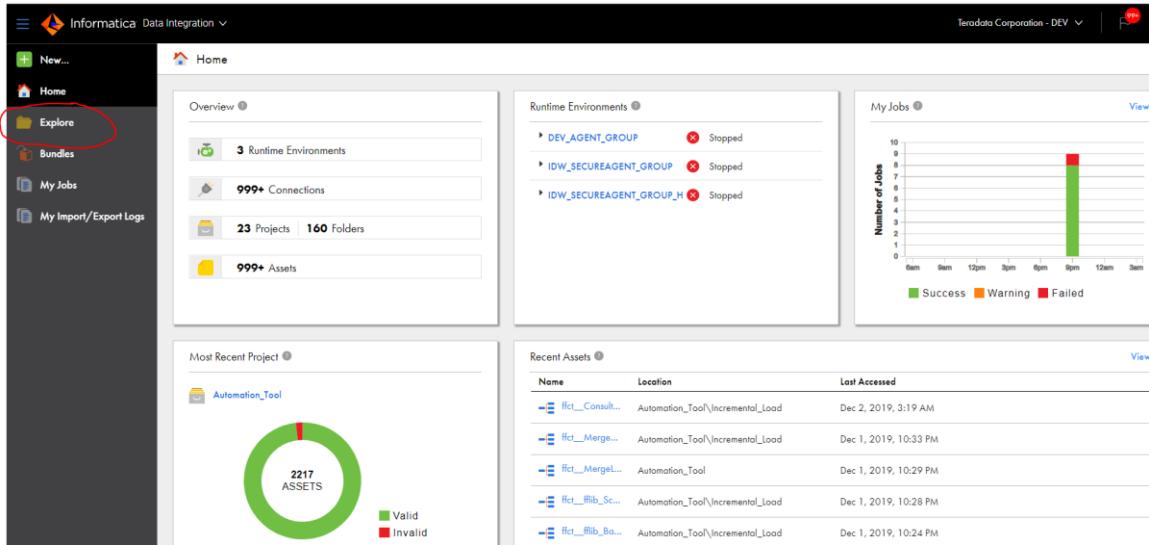
Pre-processing command configuration screen

Step 4: Exploring and Executing created task

After completing the task configuration, return to the Data Integration home screen. This is the default dashboard where all created tasks and mappings are listed.

To locate the task created in the previous steps, click on the Explore tab.

Data Integration dashboard with Explore tab

Locating the task

Use the search bar to find the required task by name. In the search results, two entries will be displayed:

- One entry for the task
- One entry for the associated mapping created for that task

This is expected behaviour, as each task in Informatica Cloud is linked to an underlying mapping.


Search results showing task and mapping entries

Executing the task

Click on the task name to open the task details page. From this screen, the task can be reviewed, edited, or executed.

If any configuration changes are required, the task can be edited before execution. Once confirmed, click Run to start the task execution.

Task details screen with Run option

Task execution

After triggering the task, Informatica Cloud begins execution using the configured runtime environment. The task runs the pre-processing command first (BTEQ script execution), followed by the mapping that loads data into the target Teradata table.
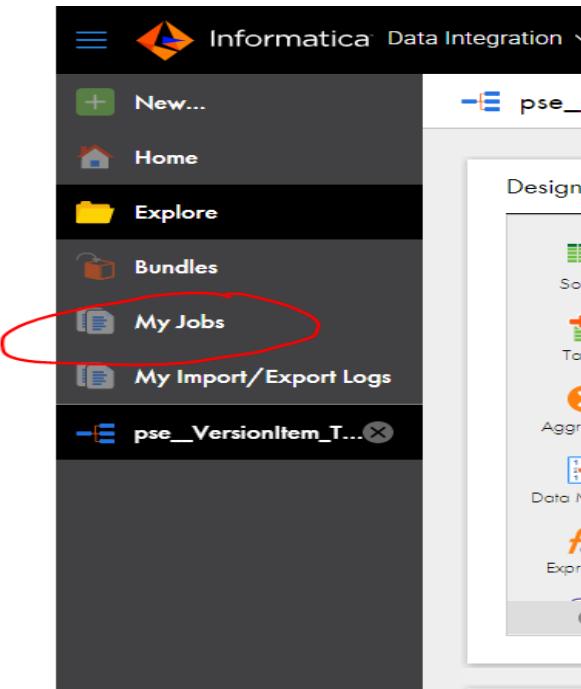
Task execution in progress

Step 5: Viewing job status and validating data load

To monitor the execution status of the task, navigate to the My Jobs tab. This section provides detailed information about all recent and running jobs, including execution time, status, and logs.

My Jobs tab showing job execution status

Data validation

Once the job completes successfully, the data load can be validated at the database level. In this example, data has been successfully loaded into the dummy target table.

As a basic validation step, a record count can be performed on the target table to confirm that rows have been inserted or updated as expected.

Instance Name	Location	Subtasks	Start Time ▼	End Time	Rows Processed	Status
product2_bteq-8	Automation_Tool		Feb 2, 2020, 10:32 PM	Feb 2, 2020, 10:32 PM	1549	✓ Success

Database row count validation

Log verification

For detailed verification, download and review the session logs from the job details page. The logs provide clear information about the number of rows processed during execution.

In this case, the logs confirm that 1,549 rows were successfully inserted into the dummy target table created for this demonstration.

Session logs showing inserted row count

This implementation demonstrates a controlled and reliable approach to loading data into Teradata using BTEQ scripts orchestrated through Informatica Cloud. By combining pre-processing safeguards with structured task execution and validation, the solution ensures data integrity and operational transparency.

All screenshots and examples are based on a real enterprise implementation. Sensitive information has been anonymized.