leaplogic | an IMPETUS product

Automated ETL transformation form Informatica to Google Cloud Composer

VIDEO TRANSCRIPT

Are you planning to modernize your ETL workloads from Informatica to Google BigQuery and Cloud Composer, but worried about business disruption?

Enterprises are modernizing their ETL workloads to Google BigQuery and Cloud Composer to leverage cloud scalability, reduce costs, and access advanced analytics.

LeapLogic, Impetus' cloud migration accelerator, can help you achieve this by ensuring an end-to-end automated transformation from Informatica to BigQuery and Google Cloud Composer.

Here's a demo of how LeapLogic automates and simplifies the transformation of Informatica workloads to Google Cloud stack – starting with a comprehensive assessment, followed by automated transformation, validation, and all the way up to operationalization.

LeapLogic first conducts a thorough assessment of your existing Informatica workloads and provides comprehensive complexity analysis and end-to-end data and process lineage. The assessment helps in planning and streamlining the migration to Google Cloud in a phased manner.

You can get an overall summary of the assessment and additional details for the complexity level of each file. For instance, we've uploaded 25 sample files, consisting of 1238 transformations, 299 mappings, and 25 workflows.

As you can see, LeapLogic has successfully analyzed 92% of the ETL scripts, indicating that it can automate the transformation of nearly all Informatica scripts to their BigQuery and Google Cloud Composer-equivalent.

LeapLogic calculates the complexity of each file and workflow based on various parameters, helping you identify the nature of workloads and the level of transformation effort required.

You can go to the Source Analysis section in the Analysis tab to view summarized statistics for all files, including source databases, workflows, mappings, transformations, and more. Similarly,

in the Entities section, you can view details regarding different types of tables, such as lookup, source, and target.

LeapLogic also provides information regarding various schedulers. In the Workflows section, you can view a summary of workflows with details related to mappings and complexity. The Artifacts section lists missing, additional, and unparsed artifacts.

The Lineage section provides a graphical lineage showcasing complex interdependencies between different workloads, along with end-to-end data and process lineage.

From here, you can also download comprehensive reports with detailed insights for offline use. These reports provide information on transformations, workflows, mappings, and more.

The Report Summary section provides details regarding all the generated artifacts. The Volumetric Info tab offers a summary of the aggregated inventory after analyzing the source files.

You can also visit the Workflow Summary section to view all workflows existing in source Informatica files. It also provides vital statistics related to various worklets, sessions, mappings, etc.

The Mapping Summary section describes the flow of data between source and target. It lists all kinds of mappings and the count of associated transformations. You can also visit the Transformation Summary section to view all the input source files along with the count of workflows, worklets, sessions, and more available in the source files.

You can also download additional reports that provide actionable prescriptive insights.

Now, let's have a look at how LeapLogic's intelligent transformation engine automates the end-to-end conversion of Informatica ETL scripts to Composer.

You can view different stages of the transformation lifecycle on LeapLogic's user-friendly interface. Firstly, select the ETL source as Informatica and upload any complex Informatica ETL file that contains unique ETL constructs and transformations you want to transform.

Select the target as Google BigQuery.

From here, you can select the orchestration technique where the LeapLogic's intelligent transformation engine generates an output to BigQuery and Google Cloud Composer-equivalent through which you can create, monitor, and manage the defined workflows.

Select Google Cloud Composer in the orchestration section. Next, you need to select the Data Interaction Technique.

The Google BigQuery-native technique accesses input data, processes it, and stores the output data in BigQuery. On the other hand, Google BigQuery: External takes input from an external source, processes it in BigQuery, and moves the output or processed data into an external target.

Select the Google BigQuery-native option. Next, select the validation type as Cluster and upload the required data source file. Specify the GCS base path where external files are stored. Now, save the ETL Conversion stage from here. You can now save the pipeline.

That's it! You can now execute the automation-driven conversion of Informatica scripts to Google Composer equivalent. Simply go to the Output section to check the converted code. As you can see, LeapLogic automates component-to-component transformation for Informatica ETL scripts to BigQuery and Google Cloud Composer equivalent, thereby preserving and optimizing the core business logic. However, where a direct equivalent is unavailable, it does component-to-query conversion to ensure seamless end-to-end transformation.

LeapLogic also provides a target-compatible code, which can be orchestrated and executed as-is on

Google Cloud Composer. Here you can see the auto-generated mapping file. You can also view the Google Cloud Composer equivalent workflow output.

You can now download the converted package containing the BigQuery and Google Cloud Composer equivalent code.

LeapLogic's unit and integration testing ensure that your transformed code is rigorously validated before productionization.

Next, log into Google Cloud BigQuery and import the converted code into the Composer DAGs folder. Here, you can see the workflow DAG that includes all the tasks involved in the process. Click on Trigger Tag to view whether all the tasks have been executed successfully.

That's it! LeapLogic streamlines your migration journey, right from assessment to transformation to execution on the target environment, ensuring a successful, end-to-end transformation of legacy Informatica workloads to Google Cloud stack.

Explore LeapLogic's automation capabilities for the end-to-end transformation of data warehouse, ETL, Hadoop, analytics, and reporting systems to Google Cloud-native stack– faster, at a lower cost, and with minimal risk.

It's more than the next step. It's a leap into the future of your business.