

Automated replatforming of Informatica workloads to native Google Cloud Dataproc

VIDEO TRANSCRIPT

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

Enterprises are modernizing their Informatica workloads to Google Cloud Dataproc to improve scalability, reduce costs, enhance flexibility, and enable faster deployment.

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

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

It first assesses the existing Informatica ETL scripts and provides comprehensive insights. You can get an overall summary of the assessment and additional details for the complexity level of each file.

For instance, we've uploaded 4 small sample files, consisting of 15 transformations, 4 mappings, and 4 workflows. LeapLogic has successfully analyzed 85% of the ETL scripts, indicating that it can automate the transformation of nearly all Informatica scripts to their Dataproc 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 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. You can also view details related to missing tables from here.

LeapLogic also provides information regarding various schedulers. In the Workflows section, you can view a summary of workflows with details related to mappings, complexity, and associated

files. 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.

You can also download comprehensive reports with detailed insights for offline use. These reports provide information on files, transformation patterns, generated artifacts, and more.

You can go to the Report Summary section to view details regarding all the generated artifacts. The Volumetric Info tab offers a summary of the aggregated inventory after analyzing the source files. The Transformation Summary section lists all input source files along with the count of workflows, worklets, sessions, and more available in 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.

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

You can undergo different stages of the transformation lifecycle on LeapLogic's user-friendly interface. Subsequently, you can provide database connection details and configure the target data source, job, source, and execution types.

That's it! You can now execute the automation-driven conversion of Informatica scripts to Dataproc 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 Dataproc 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 Dataproc. You can now download the converted package containing the PySpark equivalent code and run it optimally on Google Cloud Dataproc.

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

LeapLogic streamlines your migration journey, ensuring a successful transformation of legacy Informatica workloads to Google Cloud Dataproc.

Explore LeapLogic's automation capabilities for the end-to-end transformation of data warehouse, ETL, Hadoop, analytics, and reporting systems to a 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.