leaplogic

An Impetus Product

Solution Brief

# Intelligent and automated data platform and workloads modernization to the Databricks Lakehouse

## Auto-migrate ETL, EDW, analytics, and Hadoop workloads

Modernizing legacy workloads to a unified ecosystem like Databricks can help enterprises drive the speed, scale, and agility needed to gain a competitive edge. The Databricks Lakehouse Platform simplifies data architecture, breaks down data silos, and provides a single source of truth for data and analytics.

However, migrating years of legacy business code, logic, and workloads to Databricks has its challenges like:

- Risk of business disruption
- Inadequate, out-of-sync, or missing documentation of legacy systems
- No insight into usage, complexity, and dependencies across workloads and environments
- Uninformed decision-making without proper data-driven assessment and recommendations
- Risk of losing business logic and legacy code during conversion
- Meeting the SLAs using Databricks services while keeping the price-performance ratio in control
- Huge manual effort that turns costly with overruns and inefficiency

LeapLogic, an Impetus product for automated workload transformation, addresses all these concerns. Its intelligent grammar engine identifies optimization opportunities at schema, code, and execution levels and automatically converts all types of workloads, logic, and workflows to Databricks-native stack.

### Key Benefits

**Up to 95%** automation

**50-75%** reduced time-to-market

**90%** reduction in risk

**2x** cheaper

**1.5x** faster validation

**100%** preservation of investment in business logic

# What's possible?

## For data warehouse

- Plan a phased migration instead of a 'big bang'
- Mitigate risks beforehand with no downtime
- Handle proprietary elements like BTEQs efficiently
- Ensure optimization to meet production SLAs
- Consider all workloads in totality
- Stabilize fast with a minimal parallel run period
- Cut over and retire your legacy data warehouse
- Strategize for people and processes

## For analytics and reporting workloads

- Assess usage patterns with automation accelerators.
  E.g., SAS
  - ETL – mostly SQL + some SAS procedural
  - SAS procedural – mostly statistical
  - SAS advanced algorithms
- Convert code to Databricks-native stacks
- Map the conversion target for each usage pattern
- Enable datasets and migrate them as cloud stores or access them via JDBC
- Use a staggered approach to convert and validate scripts for Databricks-equivalence across all use cases
- Execute thorough integration testing on staging
- Execute on production

## For ETL

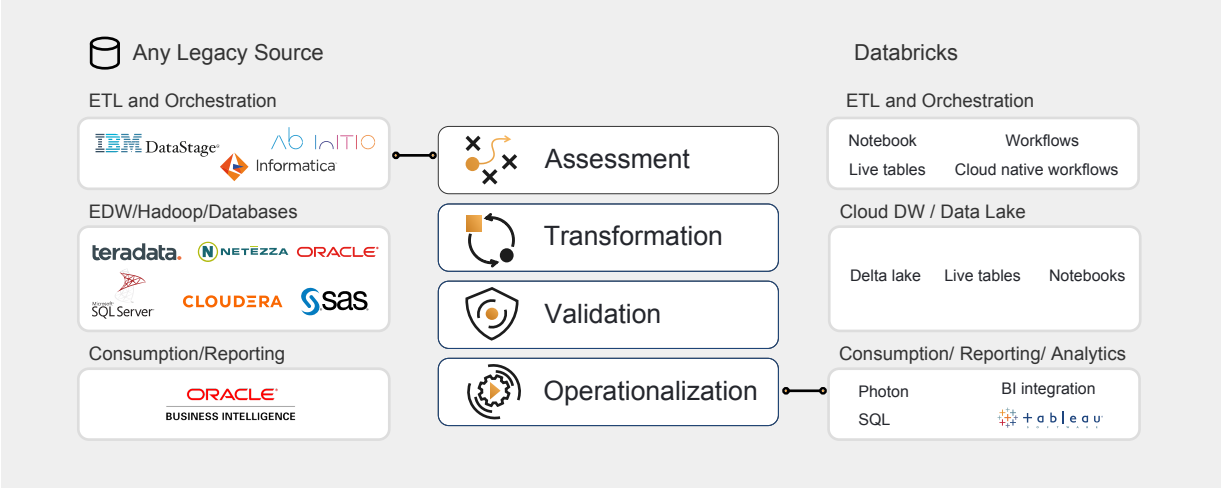
- Assess code complexity, usage patterns, etc.
- Identify and analyze complex interdependencies. For example, for Informatica, it can be from XML files to workflows and then to mappings and transformations. Similarly, for DataStage, identify and analyze all jobs and components for each script and job activity, sequencer, lookup, aggregator, Transformer stage, join, etc.
- Transform core business logic to Databricks-equivalent format
- Package for production-ready jobs
- Validate scripts for Databricks-equivalence of all use cases

## For Hadoop

- Assess infrastructure and workload inventory
- Map infrastructure and workloads on Databricks
- Assess TCO and forecast for Databricks
- Detect and optimize patterns
- Navigate complexities and risks
- Migrate efficiently at scale
- Assure quality through technology mapping
- Save cost and time with automation
- Convert workloads to Databricks-native equivalent
- Ensure validation equivalence and acceptance
- Provision through infrastructure as code

# How it works

LeapLogic enables end-to-end transformation, operationalization, and transitioning of workloads in four steps:




## STEP 1: Assessment

- Integrated assessment for several workload types
- Comprehensive inventory listing
- Workload prioritization as per the business use case
- Dependency analysis
- Advanced blueprint of Databricks stack
- Optimization opportunities for schema, code, and orchestration
- Prescriptive recommendations
- Phased migration plan with timelines and cost estimates

## STEP 2: Transformation

- Intelligent grammar engine supporting a variety of workloads and formats
- End-to-end transformation, packaging, and orchestration to Databricks-native format
- Notebook-based inline editor for query optimizations
- Delivers a verified, executable package with performance SLAs met
- Extensible, repeatable, and verifiable methodology

## STEP 3: Validation

- Pipeline-based automated validation
  - Sample or customer data-based code validation
- Auto-generation of reconciliation scripts
- Automated SQL/query and business level validation
- Cell-to-cell validation reports
- Data type and entity-level matching

## STEP 4: Operationalization

- Productionization and go-live
- Infrastructure as code
- Execution using cloud-native orchestrators
- Automated DevOps, including CI/CD, etc.
- Target environment stabilization
- Smooth cut-over

# Enterprise success stories

LeapLogic has helped several large enterprises transform their workloads to the cloud while preserving years of business logic, workflows, and execution rules.

**30% performance improvement by converting Netezza and Informatica to Azure-Databricks stack**

140 Informatica ETL scripts transformed using automation and operationalized in 16 weeks

READ MORE

**Automated assessment and transformation of Informatica workflows and Oracle EDW to AWS**

~80% auto-conversion to save 50% cost and time compared to the manual transformation

READ MORE

**Sample LeapLogic demos**

- Teradata to Databricks
- Informatica to Databricks
- SAS to Databricks
- Hadoop to Databricks
- Oracle to Databricks

---

Start your end-to-end workload transformation journey today!
To learn more, write to us at **info@leaplogic.io**

leaplogic

LeapLogic automates the transformation of legacy data warehouse, ETL, analytics, and Hadoop to native cloud platforms. Owned by Impetus Technologies Inc., LeapLogic partners with AWS, Azure, Databricks, GCP, and Snowflake to de-risk migrations. For over a decade, Impetus Technologies has been the 'Partner of Choice' for several Fortune 500 enterprises in transforming their data and analytics lifecycle, including modernization to the cloud, data lake creation, advanced analytics, and BI consumption. The company brings together a unique mix of engineering services, technology expertise, and software products.

To learn more, visit **www.leaplogic.io** or **info@leaplogic.io.**