

# Intelligent and automated transformation of SAS analytic workloads to a cloud-native stack or Spark

Moving legacy SAS analytic workloads to modern platforms on the cloud, Spark, or R has become a strategic imperative for enterprises struggling with petabytes of unstructured data from multiple sources, and high cost of ownership and operations.

However, migrating SAS workloads, including procedures, reports, and datasets to the cloud or Spark is complex, as the scripts contain thousands of conditional statements (logic), macros, procedural statements, and outputs. It is also difficult to ascertain the accuracy and precision of migrated SAS analytical models.

Because of these challenges, enterprises are still skeptical about moving to the cloud. While migrating workloads to the cloud seems easy, businesses have several concerns:

- Will there be any business downtime?
- How do I transform years of complex business logic and code?
- Will my workloads be optimized for the new environment?
- Do I need to identify and prioritize workloads manually?
- How do I seamlessly operationalize SAS ETL scripts, analytic scripts, and reports on the target environment?

The Impetus Workload Transformation Solution helps address all these concerns. Its intelligent grammar engine identifies optimization opportunities at schema, code, and execution level and automatically converts all types of workloads, logic, and workflows to a cloud-native stack of your choice or Spark.

## KEY BENEFITS

---

- **4x** faster
- **2x** cheaper
- **4x** developer productivity
- **100%** cloud-ready
- **100%** automation across the migration lifecycle

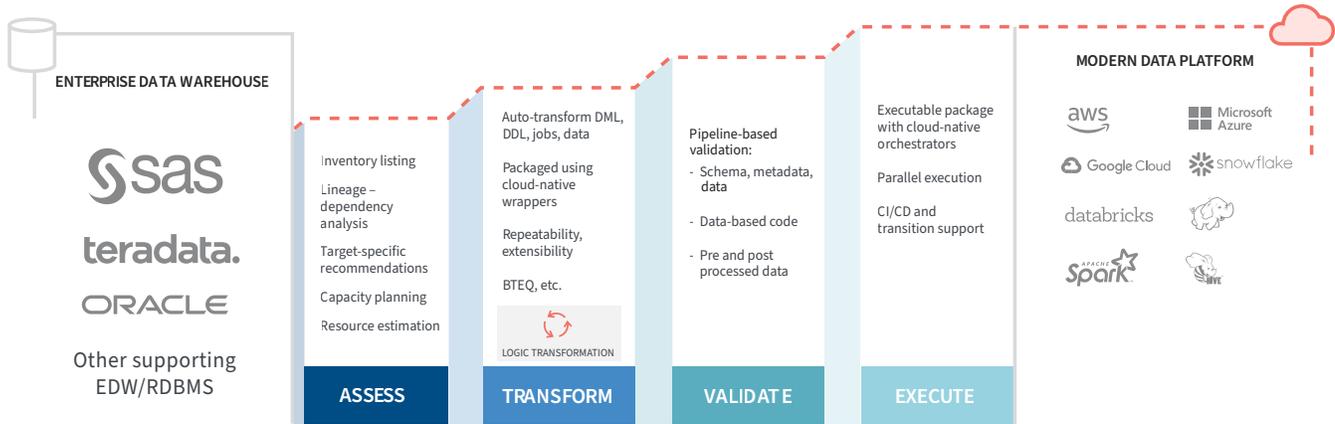
## WHAT'S POSSIBLE

---

1. ML-based assessment and recommendation for the target architecture and tech stack
2. Business logic conversion
3. End-to-end transformation to cloud-native services of your choice
4. End-to-end packaging, orchestration, and execution for the target
5. Code optimization and query validation to avoid business disruption
6. Automated legacy code translation to a cloud-compatible equivalent with multiple query engine support
7. Cost-performance ratio optimization
8. Data governance and security compliance

# How it works

The Impetus solution enables end-to-end transformation and operationalization in four steps. Automatically.



## STEP 1: Assessment and prescription

- Lists entire inventory for diverse workloads
  - Assesses ETL scripts, DML and DDL scripts, procedures, scheduler/orchestrator scripts (jobs), etc.
  - Provides actionable insights and prescriptive recommendations
- Identifies complex interdependencies to group workloads for offload
  - Plots an end-to-end lineage showing interdependencies between different kinds of workloads
  - Provides advanced filters according to workload type and an interactive graphical interface to deep dive into certain flows
- Advanced blueprinting of target architecture
  - Strategizes partitioning, bucketing, clustering, sorting, and distribution keys for improving CPU usage, memory usage, cache hit ratio, and disk I/O
  - Provides comprehensive, configurable recommendations for workload parallelism to ensure optimal performance on target
  - Provides actionable recommendations for future-state functional component architecture and tech stack components
- Assesses usage patterns to categorize workloads into three major buckets

**ETL usage**

- SQL operations + SAS procedural
- Operations like PROC SQL, REPORT, PROC CONTENTS, MACRO, etc.

**SAS statistical**

- Base operations + SAS procedural
- Operations like PROC UNIVARIATE, PROC SUMMARY, etc.

**Advanced SAS algorithms**

- SAS packages like SAS/OR, SAS/ETS, SAS/STAT, etc. + SAS procedural
- Operations like PROC LIFEREG, PROC FORECAST, etc.

## STEP 2: Transformation

- Transforms diverse workloads and migrates schema and data to the target store of choice
  - DML scripts, DDL scripts, ETL scripts, scheduler scripts, stored procedures, etc.
- End-to-end packaging
  - Transforms core business logic to cloud-native wrappers or orchestrators
  - Re-packages to an open programming language of choice
  - Ensures end-to-end transformation of scheduler/orchestrator scripts to production-ready jobs on target
  - Ensures end-to-end execution on staging and production environments after system integration testing

- Notebook-based inline query editor to optimize or address errors
- Extensible tool and methodology
  - Converts code to a variety of target stores and formats, enabling enterprise-wide use

### Accurate and precise SAS analytical models on the target

SAS workloads are transformed to a target environment of your choice. For example:

- SAS programs → Python programs
- Macros → Python functions
- PROC SQLs → SnowSQL, BigQuery, or Spark SQL
- SYMPUT and SYMGET → Spark constructs
- SAS inbuilt functions → Python inbuilt and custom functions in a usable library

The Impetus SAS Workload Transformation Solution generates production-ready code, validates, and optimizes workloads for the target platform.

### STEP 3: Validation

- Pipeline-based automated validation of the transformed code
  - Validates code at the row and cell-level and reports errors with a pluggable tool
- Data-based validation of transformed code
  - Auto-generates sample dataset based on complex query conditions – ideal for unit testing of the transformed queries
  - Feeds the customer-provided dataset for testing on real datasets – suitable for integration testing of the transformed queries

### STEP 4: Execution

- Delivers a target-specific executable package for cloud-native orchestration
- Ensures optimal performance through parallel execution
  - Provides parallel execution recommendations through exhaustive data-driven assessment
  - Generates required artifacts in the transformation output and executes them in parallel on production
- Supports productionalization
  - Supports end-to-end transitioning into production and operationalization
  - Optimizes capacity
  - Stabilizes environment through a parallel-run period
  - Monitors operations to ensure implicit data governance and compliance on the cloud
  - Ensures continuous integration and delivery
  - Provides runbook documentation, training, and handholding

Start your SAS workload transformation journey today!

To learn more, write to [inquiry@impetus.com](mailto:inquiry@impetus.com)

## IMPETUS

Impetus Technologies is focused on enabling a unified, clear, and present view for the intelligent enterprise by enabling data warehouse modernization, unification of data sources, self-service ETL, advanced analytics, and BI consumption. For more than a decade, Impetus has been the 'Partner of Choice' for several Fortune 500 enterprises in transforming their data and analytics lifecycle. The company brings together a unique mix of software products, consulting services, and technology expertise. Our solutions include industry's only platform for the automated transformation of legacy systems to the cloud/big data environment and StreamAnalytix – a self-service ETL and machine learning platform.

To learn more, visit [www.impetus.com](http://www.impetus.com) or write to us at [inquiry@impetus.com](mailto:inquiry@impetus.com).