

Data Warehouse and Data Lake Modernization

Mphasis enables enterprises to modernize legacy Data Warehouses and Data Lakes into cloud-native, scalable, governed, real-time platforms using proven frameworks (Xenon, SAGE), strong cloud partnerships, extensive tooling automation, and deep data engineering expertise.

Features

- Seamless cloud migration (DW & DL)
- No-ETL platform for simplified operations
- Integration with ElasticSearch
- ETL modernization using DataOps.Live, Streamsets, Datametica, Delphix
- Data Fabric implementation for unified access & governance

Benefits

- **Enterprise-scale modernization** with 780+ PB of data migrated and 10,500+ data practitioners supporting large, complex transformations.
- **Significant cost savings**, including multi-million-dollar reductions in operational and modernization program costs.
- **Enhanced scalability, performance, and flexibility** through cloud-native architectures (AWS, Azure, GCP).
- **Improved data governance, quality, lineage, compliance, and security** using automated validation and governance frameworks.
- **Automation-led modernization** via tools like DB Migrator, Polyjuice, Pharos, MD-Cert ensuring accuracy, faster delivery, and reduced risk.
- **Legacy ETL modernization** to cloud-agnostic ELT pipelines, improving efficiency and lowering maintenance overhead.
- **Adoption of modern Lakehouse architectures** (Databricks, Delta Lake) enabling real-time analytics, streaming, and self-service BI.
- **Data Fabric & No-ETL patterns** improving agility, data accessibility, and reducing operational complexity.
- **Better decision-making** through real-time insights, dashboards, predictive analytics, and democratized data access.
- **Structured, low-risk modernization execution** via the Xenon™ 3-phase framework enabling parallel testing and progressive migration.

Mphasis Expertise in Data Warehouse & Data Lake Modernization

Mphasis has deep, large-scale experience modernizing **Enterprise Data Warehouses (EDW)** and **Data Lakes** across AWS, Azure, GCP, hybrid, and on-prem environments. The capabilities span strategy, assessment, migration, platform engineering, and modernization accelerators.

Below is a structured summary of the strongest, directly evidenced capabilities from your data.

1 Modernization Frameworks & Methodologies

Xenon™ – Mphasis' Flagship Modernization Framework

Xenon™ provides a **structured 3-phase execution model**—Assess → ADAPT → IMPLANT—for modernizing EDWs, Data Lakes, ETL workloads, and reporting systems.

It includes:

- Seven modernization principles & eight structured frameworks
- Automated discovery, lineage extraction (Datametica, Manta)
- ETL modernization accelerators (Datametica, NextPathway, Prophecy, mLogica)
- Patterns for appliance offloading (Exadata, Teradata, Netezza, Greenplum), Hadoop

2 Data Warehouse Modernization Experience

Mphasis has modernized **petabyte-scale** EDWs for enterprise customers. Key highlights:

Data Warehouse Appliance Migration

- Migration patterns for **Exadata, Teradata, Netezza, Hadoop → EMR, HDFS → S3, Sybase/Oracle/SQL Server → Cloud**
- Petabyte-scale ingestion, transformation, SQL tuning, and staging/EDW/data mart migrations

3 Data Lake Modernization Experience

Data Lake Migration & Modernization

Mphasis has modernized **450+ PB** of data from mainframes and legacy systems and migrated **27+ PB** from a single on-prem Data Lake.

Key capabilities:

- Migration to cloud data lakes: **AWS S3, GCP BigQuery, Azure Data Lake, Delta Lake**
- Building streaming and batch ingestion: **MSK/Kafka, Glue, Airflow, Spark, Lambda, Step Functions**

- Governance frameworks: **Lake Formation, Delta Lake**, metadata & quality frameworks

4 Cloud-Native Data Platform Engineering

AWS Modernization Experience

Deep experience across the AWS data stack:

- EMR, EKS, EC2, Lambda
- MSK, SQS/SNS
- Athena, SageMaker, Glue
- S3, LakeFormation
- Multiple reusable codified architectures

Google Cloud / BigQuery Modernization

- Enterprise-scale migration to **BigQuery** for Data Warehouse + Data Lake workloads
- Modern engineering platform using **Anthos, Apigee, Vertex AI**
- Hybrid & multi-cloud architectures for financial services, logistics, healthcare

Azure & Multi-Cloud

- Support for EDW modernization to Synapse/Snowflake/Databricks
- Transformation patterns: legacy ETL → Spark/Databricks

5 Databricks Modernization & Lakehouse Engineering

Mphasis has one of the strongest Databricks Lakehouse practices:

- **400+ Databricks & Spark data engineers**
- **50+ Databricks implementations**
- Data Lakehouse migration, Delta Lake engineering, ML Ops enablement
- Modernization of 450+ PB over the past 7 years and 210+ big data implementations

6 Modernization Patterns & Tools

Common Modernization Patterns

- On-prem Hadoop → EMR
- Netezza/DataStage → GCP + Databricks
- Oracle → Aurora/PG
- HBase → DynamoDB
- Kafka → MSK

- EDW → Cloud-native warehouses (Redshift, BigQuery, Synapse, Snowflake)

Accelerators & Toolkits

- **Cloud Migration Toolkit**, DB artifacts factory, SQL tuning frameworks
- **Polyjuice** for ETL refactoring
- **Automation for regression**, test automation frameworks

7 Industry Focus & Case Examples

Across industries—Banking, Insurance, Healthcare, Logistics, Retail—Mphasis has executed complex, large-scale modernization programs:

Examples include:

- Investment banking federated Data Lake with **1+ billion records**, real-time ingestion via MSK
- 100M+ messages/day modernization pipelines
- Mission-critical risk systems moving from Ab Initio to Spark/EMR

8 Mphasis Data Engineering Practice Strength

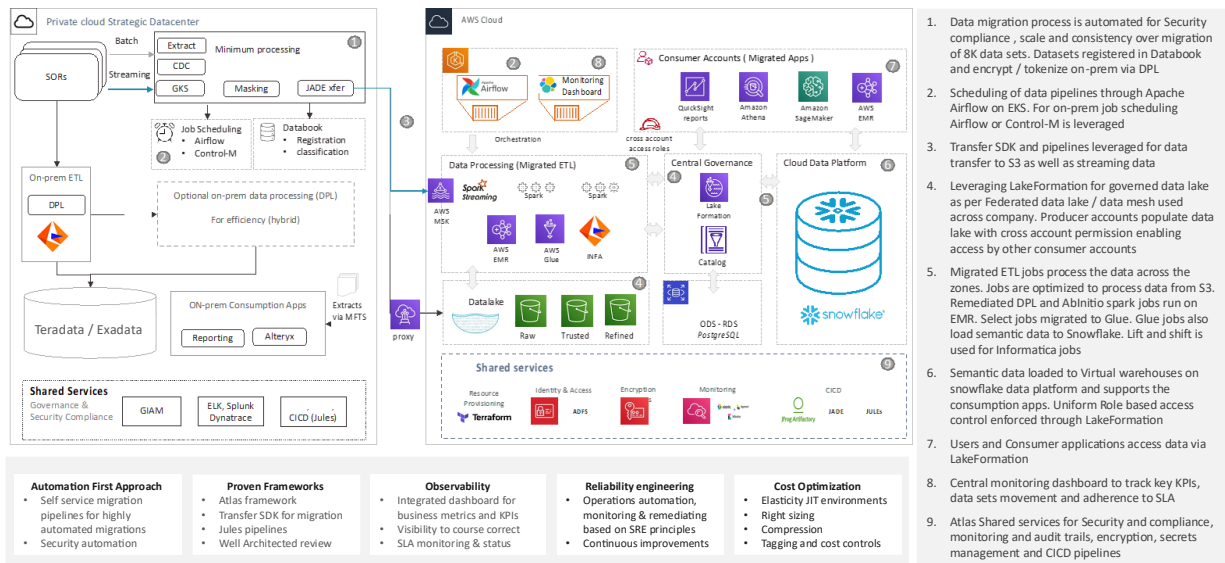
Your Data Engineering practice capabilities further reinforce modernization depth:

Core specializations:

- BI/DW design, development, and migration
- Data Lake design & modernization
- DataOps/MLOps
- Data governance, quality, metadata/lineage
- Cloud data platform implementation



Data Appliances (Teradata + Exadata) Migration For CCB LOB at a Global Bank



© Mphasis 2022 Proprietary and confidential information | 1/25/2026

| 12 |



Mphasis Experience of App & Data Modernization @ Global Bank

Mphasis has successfully delivered more than 150 Data Specific Programs at this Global Bank since 2012

CCB & AWM

- Digital Verification API Management
- Credit Journey app migration
- Contact Engine Migration
- Cards Risk Balance Transfer Migration to AWS
- BMW Fusion Exit
- 2023 Home Lending Servicing DCM
- PANDA AWS Migration
- Annual Account Review - AWS Migration
- ADE Touchstone Data
- Gateway Services Migration (Credit card history)
- DNC Privacy preference

CIB

- Securities Services -Revenue management
- Payment Optimizer - SMB
- Post Trade - Regulatory Reporting
- Post Trade – Thunderdome
- Post Trade - iRecs Discover
- Engg. & Arch – AWS Enablement
- Digital Platform Services (DPS)
- Trade Plus data migration
- Banking Tech (Payment optimizer, COLT)

CT - FRDC

- WFT DC Migration
- CCB Risk - URDS Modernization
- CCB Risk - Business Risk application migration
- Legal Tech data migration
- CCB Risk initiatives - Data Access Layer Migration
- CCB Risk Initiatives - CMDS migration
- CCB Risk Initiatives - DIAS pond migration
- Chase 360 - Risk data migration

Solve for AWS Adoption Blockers - Sparkle Projects

- CT - Legal Tech – AI based Legal Document Tagging
- CCB Digital Innovation - Auth Confidence Scoring
- CCB - OPT – ECCS – AWS rails for ML on OmniAI
- CCB – Digital Marketing – Identity Graph

© Mphasis 2022 Proprietary and confidential information | 1/25/2026

| 15 |



Mphasis Experience of AWS Data Modernization within Financial Industry

As the complex Atlas eco-system continuous to evolve it requires well versed talent pool with knowledge of Atlas landscape

Atlas Landscape			Migration Pattern Experience			
AWS Services	Managed Blueprints	Managed Patterns	Compute	Messaging/Integration	Database	Analytics
~57 and counting	~12 and counting	~60 Migration Patterns	Services	- MSK (multi-region) - Step Functions - SQS - SNS - Airflow on EKS - ODM on EMR - Pentaho on EKS - MFTS	- RDS Oracle (multi-region) - Aurora (multi-region) - DynamoDB (multi-region) - RDS postgresSQL(multi-region) - Redis on EC2 (multi-region) - MongoDB on EKS - MongoDB Atlas	- SageMaker - Comprehend - Glue - Athena - DataLake - S3, FDL (Federated DataLake) - LakeFormation - Redshift
GA & Beta	Infra and App	Across multiple categories		- Springboot/.net Core on EKS - VSI/GKP to EKS - Hadoop to EMR	- GOS to RDS Oracle - Oracle on VSI to RDS Oracle - Oracle to Aurora postgresSQL - S3 Integration, Data pump JADE SDK for transfer - Hbase to DynamoDB - Sybase to RDS SQL Server	- HDFS to S3 DataLake - ML Model migration - Decision Mgmt. migration - DataMesh - real-time inference - Tagging legal docs - Stream processing - Ad-hoc queries
Pipelines	Tools	Kickstarts	Migration Patterns			
Jules DPL/JADE JET Streams	JET OmniAI DevX Fabric Spinnaker/AIM Terraform	~45 Codified solution architectures				
Config. During Development	More enabled	More in development				

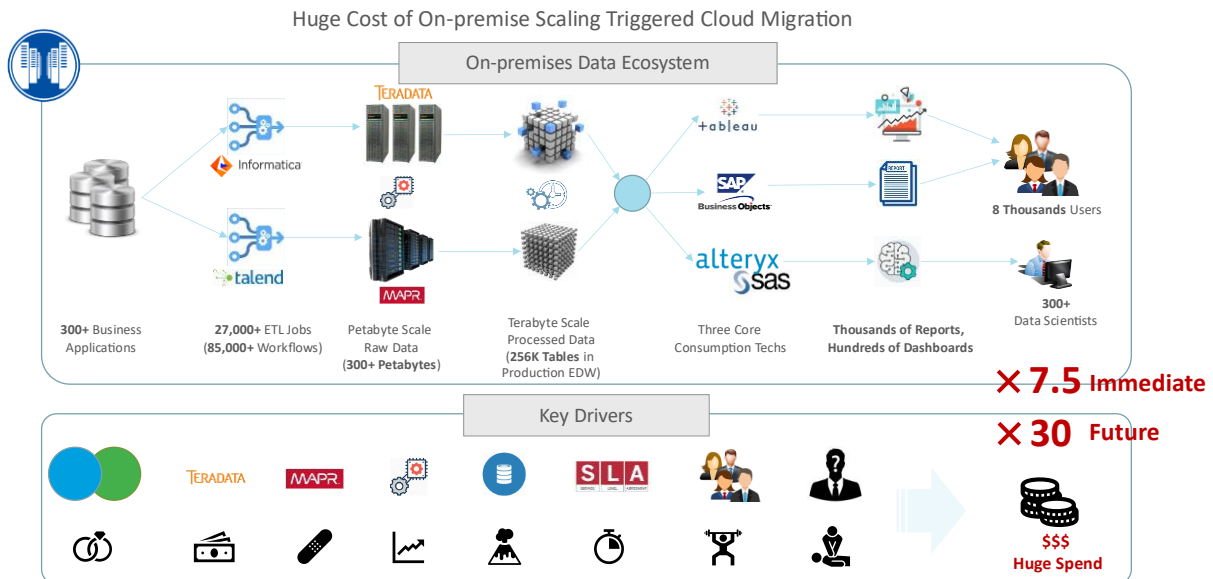
Patterns Categories				
Foundational	Integration	Storage and Database	Compute & Containers	SRE
~6	~18	~12	~8	~16

© Mphasis 2022 Proprietary and confidential information | 1/25/2026

| 16 |



Lay of Land and Key Drivers for the Transformation



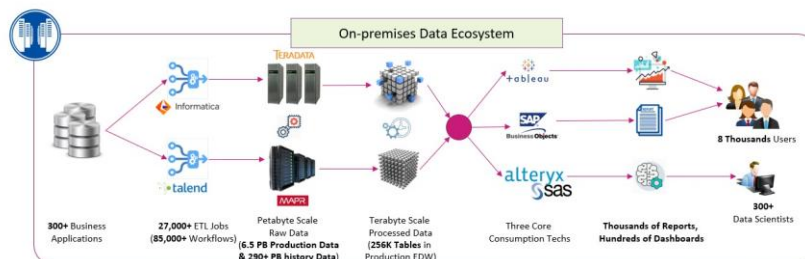
© Mphasis 2022 Proprietary and confidential information | 1/25/2026

| 3 |

The Largest Data Migration and Modernization Program in the Financial Services industry also presented many complexities / challenges

“No Vendor in the world has done such large-scale Cloud Data Migration, period!”

by Google Team



The Volumetrics

-  **6.5 PB** of production data
- 290+ PB** of Historical data
-  **256,000** Data Tables in production
-  **85,000** ETL mappings & workflows

Some Burning Complexities

1. Unclear data lineage leading to complex mining
2. Product limitations & evolving technology
3. Platform Integration friction
4. Custom consumption frameworks
5. Control M complexities
6. Absence of test data and NPI restrictions
7. Parallel runs causing extra burden on Production

Why is this Case Relevant?

- ✓ Many clients want to Modernize their Data
- ✓ All challenges & solutions applicable
- ✓ Pro-active risk management techniques directly applicable to many clients
- ✓ The Automation Platform is relevant “as-is”
- ✓ Our Learnings can help save **at least 6 months** of elapsed time for similar programs

Contact Us



Further Information: For more information about our G-Cloud services, please contact our Public Sector Team at mphasisukps@mpphasis.com