

# Migration from MySQL to Databricks



## What We Offer

Organizations are migrating from Microsoft SQL Server (MS SQL) due to its scalability limitations, high licensing and infrastructure costs, and performance bottlenecks when handling large-scale analytics. Traditional MS SQL environments often require manual query tuning, indexing, and partitioning, making it challenging to scale efficiently for big data workloads. Additionally, ETL pipelines in SSIS and batch-based data processing can slow down analytics, limiting real-time insights.

Migrating to Databricks enables businesses to leverage a cloud-native, distributed computing platform built on Apache Spark for high-performance analytics and AI-driven workloads. Unlike MS SQL's traditional row-based processing, Databricks leverages parallelized execution and in-memory computing, significantly improving performance for large datasets and machine learning applications. With Delta Lake, organizations benefit from ACID-compliant transactions, schema evolution, and optimized storage for real-time analytics. Databricks also simplifies ETL and ELT workflows through Spark SQL, Python, and notebooks, providing a scalable, cost-efficient, and AI-ready environment for modern data processing and advanced analytics.

### What our customers are saying

Being able to work on the latest technology, instead of having to manage legacy systems, creates a much better employee experience that promotes improved career progression, engagement, and retention.

**Chloe Scott**

Vice President of Data Insights and Governance, Parexel

## Why Databricks?

- o **Unified:** One architecture for integration, storage, processing, governance, sharing, analytics and AI. One approach to how you work with structured and unstructured data. One end-to-end view of data lineage and provenance. One toolset for Python and SQL, notebooks and IDEs, batch and streaming, and all major cloud providers.
- o **Scalable:** Automatic optimization for performance and storage ensures the lowest TCO of any data platform together with world-record-setting performance for both data warehousing and AI use cases — including generative techniques like large language models (LLMs). Whether you're a startup or global enterprise, Databricks is built to meet the demands of your business at any scale.
- o **Optimize Costs and Improve ROI:** The Databricks Data Intelligence Platform helps customers realize an average ROI of 482% with a payback period of as few as 4 months.
- o **Open:** With Databricks, your data is always under your control, free from proprietary formats and closed ecosystems. Lakehouse is underpinned by widely adopted open-source projects Apache Spark™, Delta Lake and MLflow, and is globally supported by the Databricks Partner Network. And Delta Sharing provides an open solution to securely share live data from your lakehouse to any computing platform without replication and complicated ETL.

## Key Use Cases



### Real-time Analytics and Business Intelligence

Ability to process large datasets in near real-time, along with integration with BI tools like Tableau and Power BI, allows users to run complex queries and create up-to-date dashboards.



### Cloud-Native Data Warehousing

Allows an organization to scale storage and compute resources independently, significantly improving performance and cost efficiency.



### Simplified ETL Pipelines

Native support for SQL, integration with cloud data orchestration tools, and capability to handle semi-structured data reduces the complexity of ETL pipelines.



### Cost Optimization for Data Storage

Can scale resources based on demand, optimizing costs for processing large amounts of historical data.

## About Databricks

Databricks is a Data and AI company. More than 10,000 organizations worldwide — including Block, Comcast, Condé Nast, Rivian, Shell and 60% of the Fortune 500 — rely on the Databricks Data Intelligence Platform to take control of their data and put it to work with AI. Databricks is headquartered in San Francisco, with offices around the globe, and was founded by the original creators of Lakehouse, Apache Spark™, Delta Lake and MLflow. To learn more, follow Databricks on LinkedIn, X and Facebook.

## About Persistent

Persistent Systems (BSE & NSE: PERSISTENT) is a global services and solutions company delivering Digital Engineering and Enterprise Modernization to businesses across industries. With over 24,500 employees located in 19 countries, the Company is committed to innovation and client success. Persistent offers a comprehensive suite of services, including AI-enabled software engineering, product development, data and analytics, CX transformation, cloud computing, and intelligent automation. The Company is part of the MSCI India Index and is included in key indices of the National Stock Exchange of India, including the Nifty Midcap 50, Nifty IT, and Nifty MidCap Liquid 15 as well as several on the BSE such as the S&P BSE 100 and S&P BSE SENSEX Next 50. Persistent is also a constituent of the Dow Jones Sustainability World Index. The Company has achieved carbon neutrality, reinforcing its commitment to sustainability and responsible business practices. As a participant of the United Nations Global Compact, Persistent is committed to aligning strategies and operations with universal principles on human rights, labor, environment, and anti-corruption, as well as take actions that advance societal goals. With 327% growth in brand value since 2020, Persistent is the fastest-growing IT services brand in the 2024 Brand Finance India 100 Report.

### USA

Persistent Systems, Inc.  
2055 Laurelwood Road, Suite 210  
Santa Clara, CA 95054  
Tel: +1 (408) 216 7010  
Fax: +1 (408) 451 9177  
Email: [info@persistent.com](mailto:info@persistent.com)

### India

Persistent Systems Limited  
Bhageerath, 402  
Senapati Bapat Road  
Pune 411016  
Tel: +91 (20) 6703 0000  
Fax: +91 (20) 6703 0008



**Persistent**  
Re(AI)maging the World