

Modernize your data estate by migrating to Azure Databricks



Contents

SECTION 1:	The data-driven transformation imperative.....	4
	Data's role in strategic growth.....	4
	Key challenges with legacy systems	4
	Modernizing with Azure Databricks.....	5
	Strategic partnership for future growth.....	5
SECTION 2:	The new data reality — Why change is necessary.....	6
	The growing need for modern data infrastructure	6
	Rising data demands require new approaches.....	6
	Limitations of traditional data warehouses.....	7
	The benefits of a modern cloud-native platform.....	8
SECTION 3:	Azure Databricks as a catalyst for transformation	9
	Streamlining data operations.....	9
	Accelerating time to insight with real-time processing	10
	Driving innovation with AI and machine learning	10
	Advantages of the Azure Databricks platform.....	11
SECTION 4:	The strategic advantage of migrating to Azure Databricks — AI and real-time analytics.....	12
	Maximizing AI and ML capabilities	12
	Real-time analytics for rapid insights	12
	Scaling data operations to drive innovation.....	13
	Enhanced value through Azure integrations.....	14
SECTION 5:	Key benefits of migrating to Azure Databricks.....	15
	Reducing operational overhead	15
	Enhancing scalability and flexibility.....	15
	Leveraging advanced analytics and AI.....	16
	Improved data governance with Unity Catalog.....	16
	Cost savings and predictable pricing.....	16

Contents

SECTION 6:	Enhancing data governance and security with Unity Catalog	17
	Centralized governance and data access control.....	17
	Data lineage for transparency and compliance.....	17
	Simplified compliance with automated governance.....	18
	Enhanced security with Azure integration.....	18
	Supporting long-term governance goals	18
SECTION 7:	Preparing for a successful migration to Azure Databricks	19
	Step 1: Assess your current data infrastructure and needs	19
	Step 2: Develop a phased migration strategy	19
	Step 3: Leverage Azure Databricks migration tools and support.....	19
	Step 4: Prepare data for migration	20
	Step 5: Test and validate migrated data.....	20
	Step 6: Facilitate change management and team training	20
SECTION 8:	Measuring success and maximizing value after migration	21
	Establishing KPIs to measure migration success	21
	Optimizing workflows and operations on Azure Databricks	22
	Driving innovation with AI and advanced analytics.....	22
	Maximizing ROI with Azure integrations.....	23
	Continuous improvement for long-term value	23
SECTION 9:	Future-proofing your data strategy with Azure Databricks	24
	Building a scalable foundation for long-term growth.....	24
	Adapting to emerging technologies and analytics trends	24
	Driving innovation with AI and machine learning	25
	Reinforcing data governance for a secure future	25
	Building security with Azure integrations	25
SECTION 10:	Driving growth with Azure Databricks	26
	Building a resilient, future-ready data environment.....	27
	Taking the next step.....	27

SECTION 1:

The data-driven transformation imperative

Data's role in strategic growth

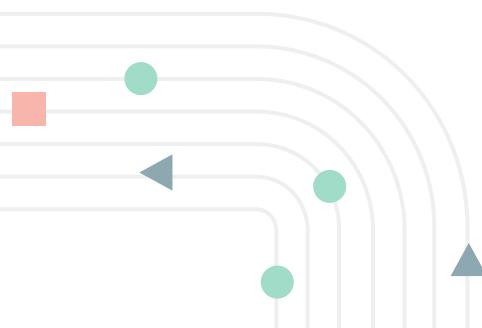
Data is integral to decision-making and operational efficiency across industries. Every day, companies generate extensive amounts of data from diverse sources, including sales transactions, social media, Internet of Things (IoT) devices, and AI systems. The convergence of traditional data sources with AI — including generative AI and large language models — has created new opportunities and challenges for organizations. The ability to process, analyze, and leverage this data effectively allows organizations to respond swiftly to market trends, optimize operations, and create new business opportunities through AI-powered insights.

However, many companies still operate with legacy, on-premises data warehouses that struggle to meet the growing demands of modern analytics and AI workloads. These systems, designed for traditional structured data processing, often fall short in delivering real-time insights, supporting advanced analytics capabilities, supporting organizational growth, and handling the computational requirements of modern AI applications, limiting an organization's ability to fully harness its data for strategic initiatives and AI-driven innovation.

Key challenges with legacy systems

Traditional on-premises data systems present several common challenges that impact scalability, cost, and performance:

- A. High maintenance costs** – Legacy systems require ongoing investments in hardware and infrastructure management. These expenses often divert resources from innovation and limit IT budgets for strategic priorities.
- B. Limited scalability** – On-premises data warehouses lack the flexibility of cloud-based platforms, making it difficult and costly to scale as data volumes grow. Fixed capacity can prevent organizations from adapting quickly to fluctuating data demands.
- C. Performance bottlenecks** – Slow data retrieval, limited support for concurrent users, and inefficient processing lead to delays and performance issues. As data volumes increase, these limitations become more restrictive, impacting decision-making.
- D. Data silos** – Integrating data from different sources and departments can be complex within legacy environments, leading to fragmented data and limited visibility across the organization. Siloed data restricts the scope and accuracy of analytics, further limiting potential insights.



Modernizing with Azure Databricks

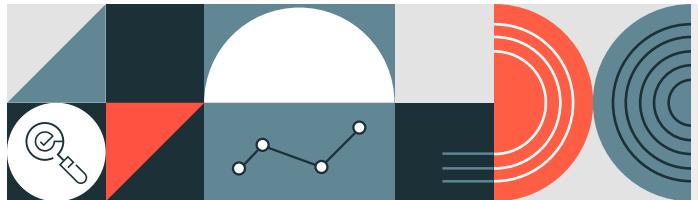
Azure Databricks offers a cloud-native, AI-powered Data Intelligence Platform co-developed with Microsoft that addresses the challenges of legacy systems. This first-party platform provides the scalability and flexibility needed to meet the demands of modern analytics, including machine learning (ML), real-time data processing, data warehousing, and generative AI applications. Migrating to Azure Databricks can deliver:

- **Seamless integration within the Azure ecosystem** – Azure Databricks allows organizations to leverage the full Azure suite, ensuring compatibility across various business functions, from data engineering to AI workflows.
- **Reduced infrastructure management** – By moving to a managed, cloud-based solution, companies can shift IT resources from maintenance to more strategic, value-added activities. This shift supports agility, operational efficiency, and cost savings.
- **Enhanced support for real-time and AI-driven applications** – The platform's capabilities in handling high-demand analytics make it possible for organizations to act on real-time data insights. With AI/Business Intelligence (BI) Genie integration, users can leverage natural language queries and generative AI features to democratize data access and accelerate insight discovery, helping drive responsive and proactive decision-making.

Strategic partnership for future growth

Microsoft and Databricks unite leading technology and innovation to support organizations in their data transformation efforts. As the first-party service within Azure, Azure Databricks is optimized for security, governance, and scalability in the cloud. Microsoft and Databricks work collaboratively to ensure that Azure Databricks evolves alongside industry needs, offering organizations a platform that is continuously updated and aligned with enterprise-grade requirements for security and compliance.

Organizations that adopt Azure Databricks can confidently transition to the cloud, knowing that the platform's features and updates are designed to support both current and future needs in analytics, data science, and machine learning.



DID YOU KNOW?

Azure Databricks provides Unity Catalog, an integrated data governance solution that allows businesses to manage data access, track usage, and enforce compliance across data environments. Unity Catalog supports transparency and security for organizations with complex data governance requirements.

SECTION 2:

The new data reality – Why change is necessary

The growing need for modern data infrastructure

Data is one of the most valuable assets for any business, and modernizing data infrastructure is critical to leverage it fully. Many companies still rely on legacy systems that limit scalability, add operational costs, and slow access to real-time insights, all of which impact competitiveness. By migrating to Azure Databricks, organizations can integrate both structured and unstructured data in real time, making it feasible to implement agile, data-driven decisions and optimize operations. Starting with a phased migration plan helps test and validate the new system, reducing transition risks and maximizing efficiency improvements.

As data demands rise, organizations are finding that traditional data warehouse systems cannot keep up. This section covers the impact of today's data needs, the challenges of outdated systems, and why a cloud-native solution like Azure Databricks is essential.

Rising data demands require new approaches

The rate of data generation continues to accelerate, fueled by a mix of digital interactions, IoT devices, and data sources that extend beyond structured formats. Businesses now need systems that can manage both structured and unstructured data — such as customer feedback from social media and real-time sensor data. Traditional systems often cannot handle this level of diversity and scale.

Key aspects of these growing demands include:

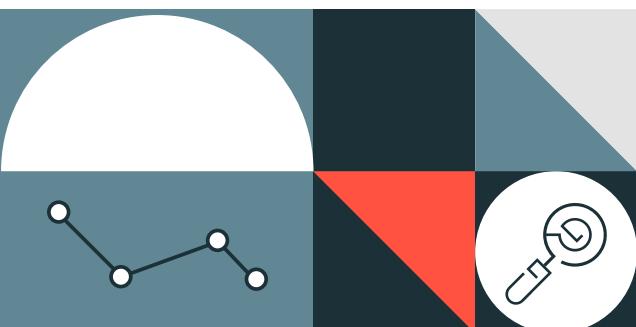
EXPANDING DATA SOURCES	REAL-TIME DECISION-MAKING	ADVANCED ANALYTICS
<p>Data now originates from numerous channels, including mobile applications, connected devices, web analytics, and transaction systems. Organizations require a unified architecture that can efficiently integrate and process these various data types.</p>	<p>Timely insights are increasingly critical for tracking customer behavior, optimizing operations, and adapting to market changes. Fast data processing enables businesses to act quickly and make informed decisions.</p>	<p>AI and machine learning are becoming standard tools for uncovering predictive insights, automating processes, and enhancing customer experiences. These complex workflows require infrastructure capable of high-volume data processing and sophisticated data transformations.</p>

Legacy infrastructure often cannot support these modern demands, creating barriers to effective data use. Azure Databricks offers the scalability, speed, and adaptability needed to fully leverage data as a strategic resource.

Limitations of traditional data warehouses

Traditional data warehouse appliances were designed for stable, structured reporting, making them ill-suited for the dynamic data landscape of today. The following limitations are commonly encountered by organizations relying on outdated systems:

- **Scalability constraints** – Physical hardware restrictions make it challenging to scale on-premises systems to meet increased data demands. Scaling often involves significant investments in new hardware, which are both costly and time-intensive, limiting an organization's ability to manage demand spikes effectively.
- **High operational costs** – Maintaining legacy data warehouses involves substantial expenses, including costs for hardware, manual maintenance, and regular system upgrades. As data needs continue to grow, these costs can quickly outpace the value provided by the system.
- **Slow response times and bottlenecks** – Growing data volumes create performance issues in traditional systems, leading to slower queries, limited support for concurrent access, and complex data retrieval processes. These bottlenecks delay insights, impacting the organization's ability to respond swiftly to data-driven needs.
- **Data silos and limited integration** – On-premises systems often operate in isolation, making it difficult to consolidate data across departments. This siloed data limits the scope of analysis and prevents businesses from accessing a full, integrated view of their operations.



DID YOU KNOW?

Organizations that migrate to Azure Databricks from legacy systems often see significant reductions in total cost of ownership, ranging from 30–50%, as they shift from maintaining costly on-premises infrastructure to a fully managed, scalable cloud environment. This allows teams to redirect resources from system management to data analysis and innovation.

The benefits of a modern cloud-native platform

Transitioning to a cloud-native platform like Azure Databricks allows organizations to replace outdated infrastructure with a flexible, scalable environment that supports a broad range of data science and analytics needs.

Key advantages of migrating to Azure Databricks include:

ELASTIC SCALABILITY	Azure Databricks allows independent scaling of compute and storage resources, enabling organizations to meet periods of high demand without heavy hardware investments. This elasticity supports cost-effective scaling, as businesses can adjust based on real-time needs.
AUTOMATED MAINTENANCE AND UPDATES	With Azure Databricks, infrastructure maintenance and software updates are handled within the Azure ecosystem. Automated cluster management and system optimizations allow data teams to focus on analytics and strategy rather than upkeep.
ENHANCED ACCESSIBILITY AND COLLABORATION	As a cloud-based solution, Azure Databricks enables secure and performant data access for global teams, providing shared workspaces that allow data engineers, scientists, and analysts to collaborate in real time. This environment fosters greater alignment and faster insights across the organization.

CASE STUDY 1 (FINANCIAL SERVICES):

Belfius Bank – Enhanced insights for decision-making with Azure Machine Learning

Belfius Bank, a major financial institution in Belgium, utilized Azure Machine Learning to improve its credit risk assessments and deliver real-time financial insights. By transitioning its data operations to the Azure platform, Belfius gained the ability to analyze massive datasets, driving faster and more accurate insights. This transition not only optimized risk analysis but also improved customer experience by delivering personalized banking solutions efficiently.



SECTION 3:

Azure Databricks as a catalyst for transformation

Streamlining data operations

As the first-party solution with Microsoft, Azure Databricks is a powerful transformation tool that goes beyond simply replacing legacy systems. This native integration ensures seamless compatibility with Azure services, enterprise-grade security, and optimal performance within the Azure ecosystem. With a single platform for data engineering, analytics, and machine learning, it enables organizations to move from static, resource-heavy infrastructure to a cloud-native environment that supports integrated data operations and delivers advanced insights.

Azure Databricks brings data engineering, analytics, and AI into one unified environment. This approach minimizes silos, streamlines complex workflows, and enhances collaboration, enabling teams to manage the entire data lifecycle – from ingestion to insight – on a single platform. By centralizing these operations, businesses can access real-time insights and conduct batch processing in ways that are not feasible with legacy systems.

Key features that drive this transformation include:



Real-time and batch processing – Azure Databricks allows for both real-time streaming and batch data processing, letting organizations react quickly to business changes and handle large-scale analytics efficiently.



Scalable AI and machine learning – Azure Databricks facilitates the deployment of AI into production, supporting both classical and generative AI applications. Integrated with widely used machine learning tools like MLflow, PyTorch, and TensorFlow, the platform enables organizations to efficiently manage the full AI lifecycle, from experimentation and model training to deployment. This end-to-end support allows teams to seamlessly operationalize AI across business functions, driving insights and automation at scale.

Accelerating time to insight with real-time processing

Real-time processing capabilities in Azure Databricks allow businesses to act on data as events unfold, providing a strategic advantage in industries where timing is critical. The platform's structured streaming functionality supports real-time monitoring, making it ideal for applications in finance, retail, and manufacturing where rapid response can impact outcomes significantly.

For example:



Streaming analytics for live insights – With structured streaming and Delta Live Tables (DLT), Azure Databricks enables continuous monitoring and analysis, allowing teams to respond to customer needs, operational shifts, and market changes in real time. DLT automates the setup and maintenance of reliable, high-quality data pipelines, ensuring data accuracy and consistency for real-time analytics.



Low-latency processing – By reducing latency from data collection to insight, Azure Databricks supports fast decision-making for applications like fraud detection, supply chain optimization, and customer experience management.

Driving innovation with AI and machine learning

Azure Databricks provides a comprehensive toolkit for AI and machine learning, designed to empower organizations to innovate with minimal complexity:



End-to-end MLOps and AI support – Azure Databricks provides comprehensive MLOps capabilities that extend beyond traditional lifecycle management. With MLflow for experiment tracking and deployment, as well as integrated support for Vector Search, GPU model serving, and hosted LLMs, the platform helps organizations put both classical and generative AI into production. Additional tools like the feature store, agent framework for autonomous workflows, and model evaluation and fine-tuning capabilities ensure seamless scaling and governance for AI applications.



Comprehensive model deployment and monitoring – The platform's robust model management features, including model monitoring, AutoML for rapid prototyping, and support for continuous improvement, enable teams to deploy and monitor AI models effectively. Azure Databricks also offers advanced capabilities like the feature store and model tracking across diverse use cases, driving AI and ML insights with minimal manual intervention.