



# Cloud Data Analytics Company

Enabling enterprises with data driven decision making

CloudOps

Data Analytics

Modern Apps



# Anblicks at a glance

---



**2004**

Established in Texas, USA



**500+**

Technology Professionals



**200+**

Customers Served



**3 Countries**

USA, Australia, India



**16 Books**

Authored by Employees



**Accelerators**

Data Engineering, Data Science

# Anblicks CoEs and Services

## CloudOps



**Cloud Strategy & Assessment**  
Infra, Apps, DevOps, Security



**Infrastructure as Code (IaC)**  
Terraform, Azure ARM & Blueprints



**DevOps Automation**  
Azure DevOps



**Containerization & Orchestration**  
Kubernetes, Azure Kubernetes



**Hybrid, Multi-cloud Management**  
Azure Arc



**24x7 Cloud Managed Support**  
Infrastructure Management

## Data Analytics



**Fully Managed ETL/ELT**  
Azure Data Factory



**Lakehouse for Analytics**  
Azure Databricks



**AI/ML and Advanced Analytics**  
Azure ML, Databricks MLflow



**Scalable Managed Database**  
Azure SQL, Synapse



**Modern Cloud Data Warehouse**  
Snowflake



**Actionable Business Intelligence**  
Power BI

## Modern Apps



**Application Modernization**  
Azure Webapps, Serverless



**APIs & Microservices**  
Azure APIs, Containers, Functions



**Cloud-native Applications**  
Azure fully managed services



**Web / Mobile Applications**  
Portals, Dashboards



**UI / UX Design**  
Design Thinking, Ideation



**QA Automation**  
App, API, Data Testing

Microsoft  
Partner



Gold Data Analytics

# Hadoop to Azure Databricks Migration



# Hadoop to Azure Databricks Engagement Process

Discovery & Planning

Target Delta Architecture

Migration

Validation

STAGE 1:  
DISCOVERY

**ANALYZE:** Catalog and Understand different workloads running in the current Hadoop environment. Analyze current Hadoop usage and auto migration percent utilizing the Hadoop to Azure Databricks self service analyzer.

**DESIGN:** Size and Design the new architecture to support both data and applications including CDC

**PLAN:** Detailed migration plan to ensure minimal exceptions and risks. Provision and Configure Databricks Account, Workspaces and Clusters

STAGE 2:  
IMPLEMENTATION

**MIGRATE SCHEMA:** H2D tool takes care of this. DDL and scripts will be translated to be valid in Databricks. UDFs can be manually migrated through Anblicks and workload migration based on the migration plan defined.

**MIGRATE DATA:** Hadoop to Azure Databricks tool takes care of this. The tool first uploads data to a cloud storage (ADLS Gen2/S3 bucket) before loading into Databricks. Alternatively Azure Data Box or AWS Snowball could be used.

**DATA INGESTION (ONGOING):** Connect Databricks directly with the sources of data that are currently writing into Hadoop. Hadoop to Azure Databricks tool allows to perform file level change data capture (CDC) until all the ETL changes are complete.

**APIs, REPORTING Tools:** Integrations for apis and reporting tools will be modified and reconnected

STAGE 3:  
VALIDATION

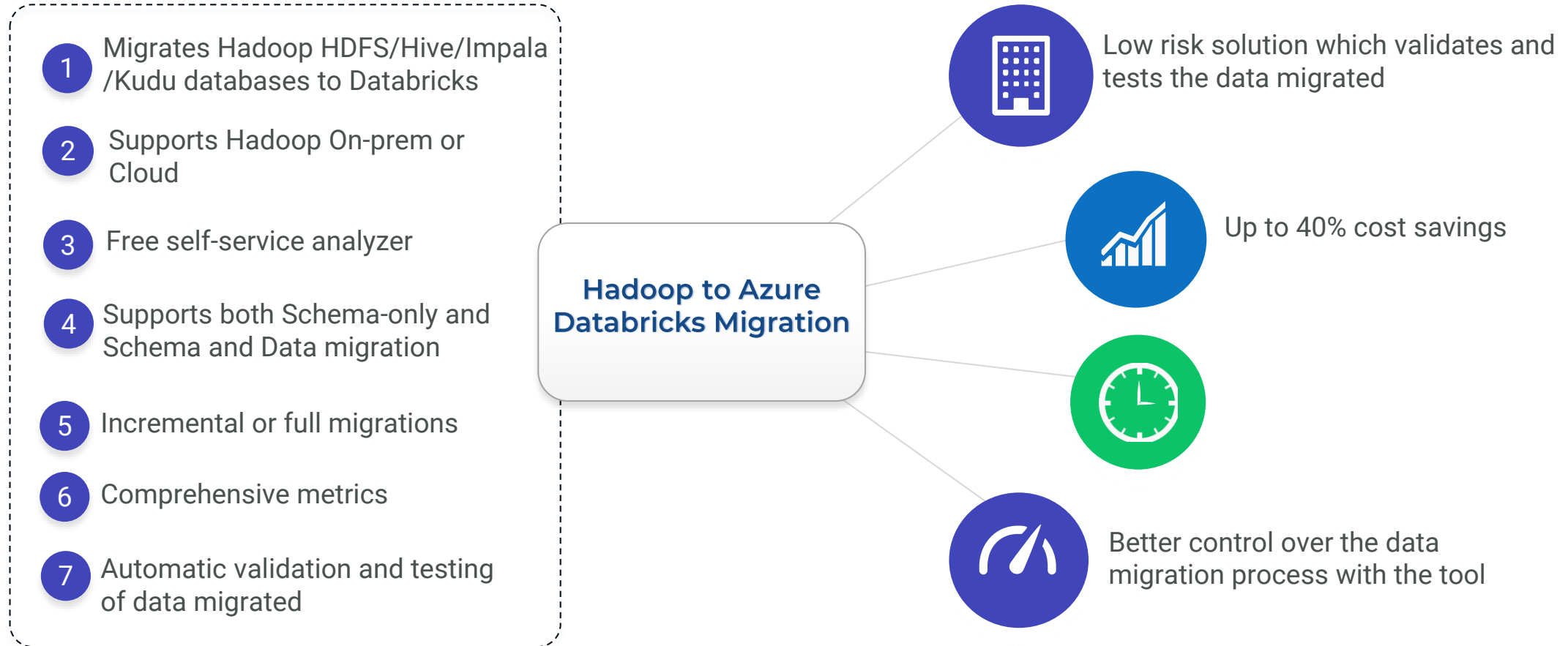
**DATA VALIDATION:** Side by side comparison of the legacy data results and Azure Databricks results. Verify data was moved successfully and that any translated logic is giving the intended results.

**DECOMMISSION HADOOP:** Inform all the Hadoop users about their new Azure Databricks accounts. Turn off all Hadoop load processes and revoke access to Hadoop.

**ROADMAP:** Establish plan for future and for the next phases

# Hadoop to Azure Databricks: Accelerate the Migration

Low-cost solution leveraging open-source technologies to perform the big data migration to Azure Databricks



# Key Takeaways

---

## Hadoop to Azure Databricks Features



Push button solution to migrate Hadoop HDFS/Hive/Impala/Kudu databases to Azure Databricks



Supports both Schema-only and Schema and Data migration



Saves up to 60% of time and 40% of costs

## Benefits of migration from Hadoop to Azure Databricks



No hardware lock-in



Make better, quicker business decisions



Reduce the cost of scaling data management and analytics





# Thank You



[www.anblicks.com](http://www.anblicks.com)



[sales@anblicks.com](mailto:sales@anblicks.com)