Azure Databricks

Helping data teams solve the world's toughest problems

Presenter name
Title
Company

Agenda

- Qualification criteria and Azure Databricks Sales Pitch
- Industry Focus
- Hadoop Migrations
- Discount Plans (RI) and Big Data Migration Promotion
- Driving consumption with Databricks

Qualification criteria and ADB sales pitch

Unified data analytics platform for accelerating innovation across data science, data engineering, and business analytics

Global company with 5,000 customers and 450+ partners

Original creators of popular data and machine learning open source projects

Azure Databricks

First party service

Enterprise-ready: secure & compliant

Native integration with Azure services

Partner ecosystem ready

Azure Databricks

Amazing growth in the first 2 years

Thousands

Azure Databricks customers globally

Millions

Server-hours spinning up every day

2 Exabytes

Data processed per month

30+

Regions available worldwide

Azure Databricks

DELTA LAKE

Open-source storage layer that brings ACID transactions to Apache Spark™ and big data workloads.

75%

of the data processed in Azure Databricks is Delta

Azure Databricks

mlflow™

More than 200 contributors and

2.5M+

monthly downloads

Thousands of customers: Data-driven innovation across industries

Healthcare and Life Sciences

Financial Services

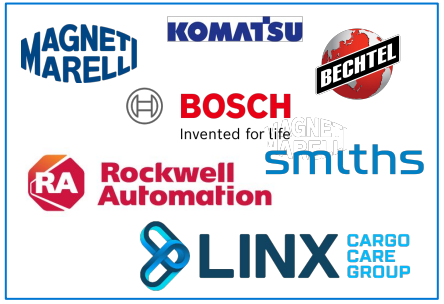
Media and Entertainment

Retail, CPG, and eCommerce

Energy & Utilities

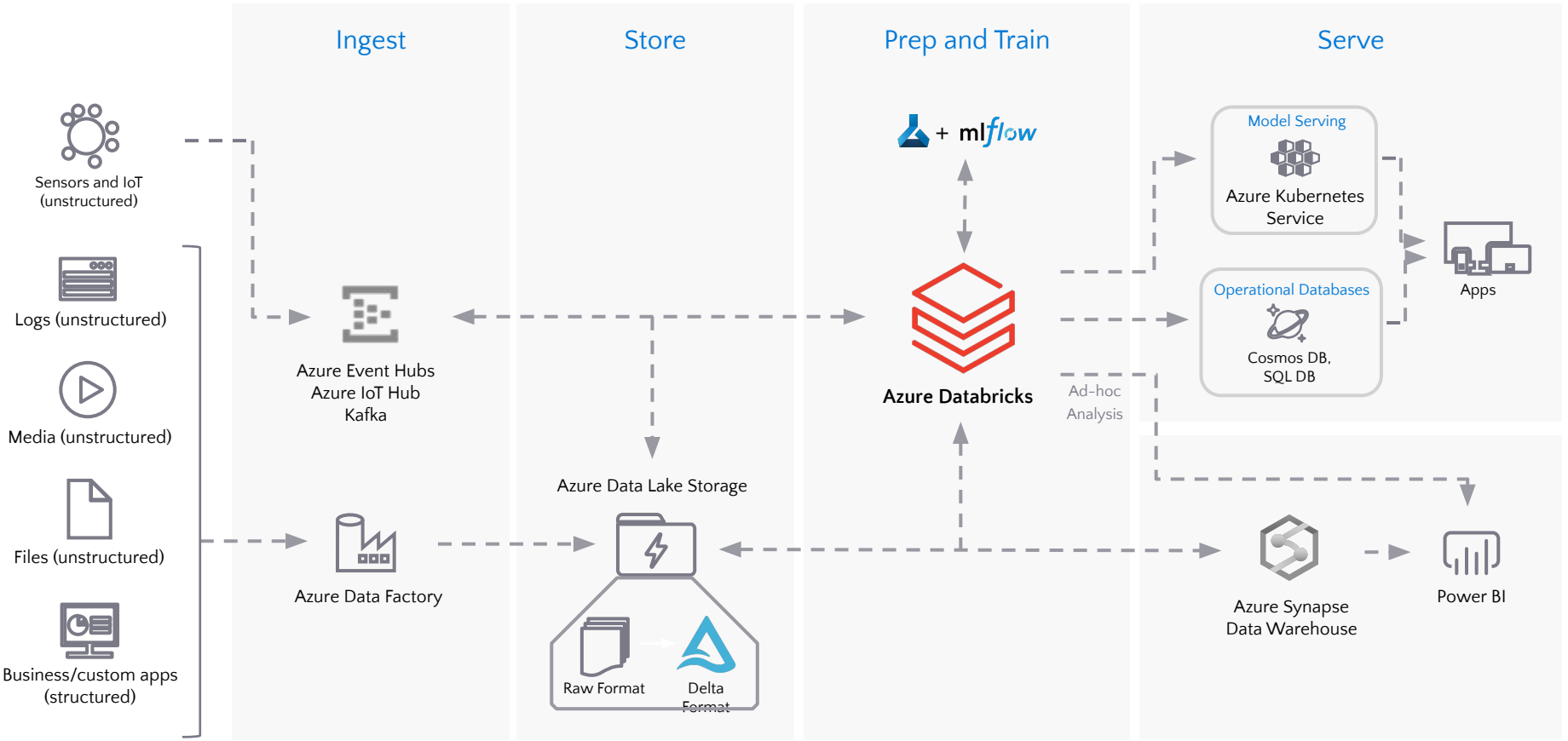
Enterprise Technology

Manufacturing & Automotive



Telecom

Azure Databricks And Azure Data Services



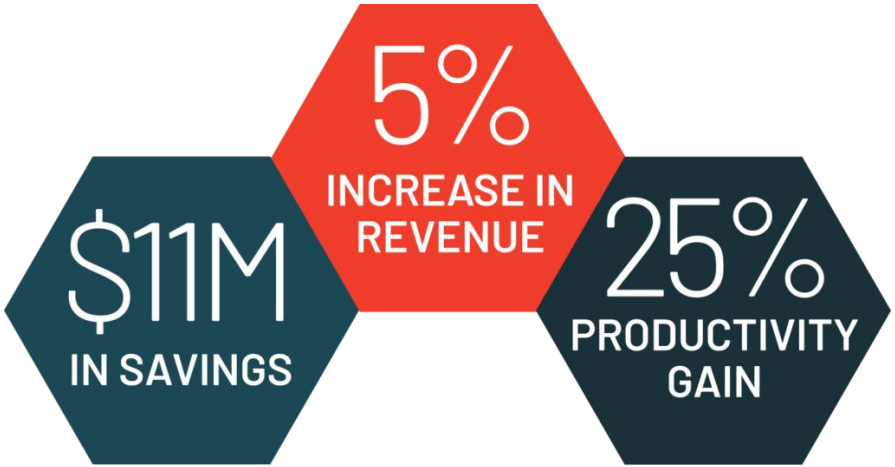
ADB Qualification criteria

Prospect	Azure Ready	Data	Big Data Expertise	User Personas	Timeline to Prod
Great	All Project Data in Azure	10TB +	In Production with Spark/Hadoop	Combination of "true" Data Engineers, Scientist and Advanced Analysts	ASAP
Good	Most Project Data in Azure	1-10TB	Production Hadoop / Moving to Spark	Single Team of "true" Data Engineers or Data Scientists	1-3 months
Okay	Some Project Data in Azure	100GB - 1TB	Learning/New to Hadoop or Spark. MPPs Only (SAP Hana, Teradata, etc)	Tradition GUI based ETL Users (Informatica, Talend, SSIS, etc.) , "Aspirational" Data Scientists, SQL only Analyst	3-12 months
Effort	No Project Data in Azure	<100GB	No Big Data Tools	Undetermined	12+ months

Azure Databricks

Databricks delivers nearly $29 million in economic benefits and pays for itself in less than six months

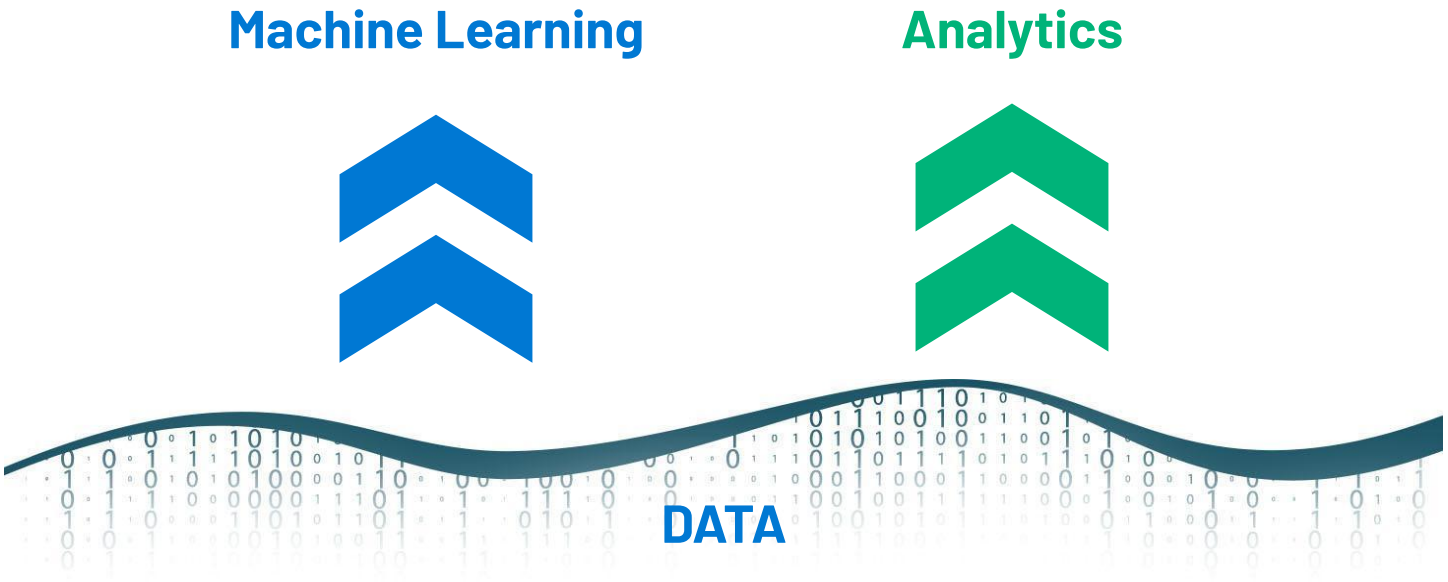
Databricks customers experience revenue acceleration, improved data team productivity and infrastructure savings resulting in a 417% ROI on their data analytics and AI projects.



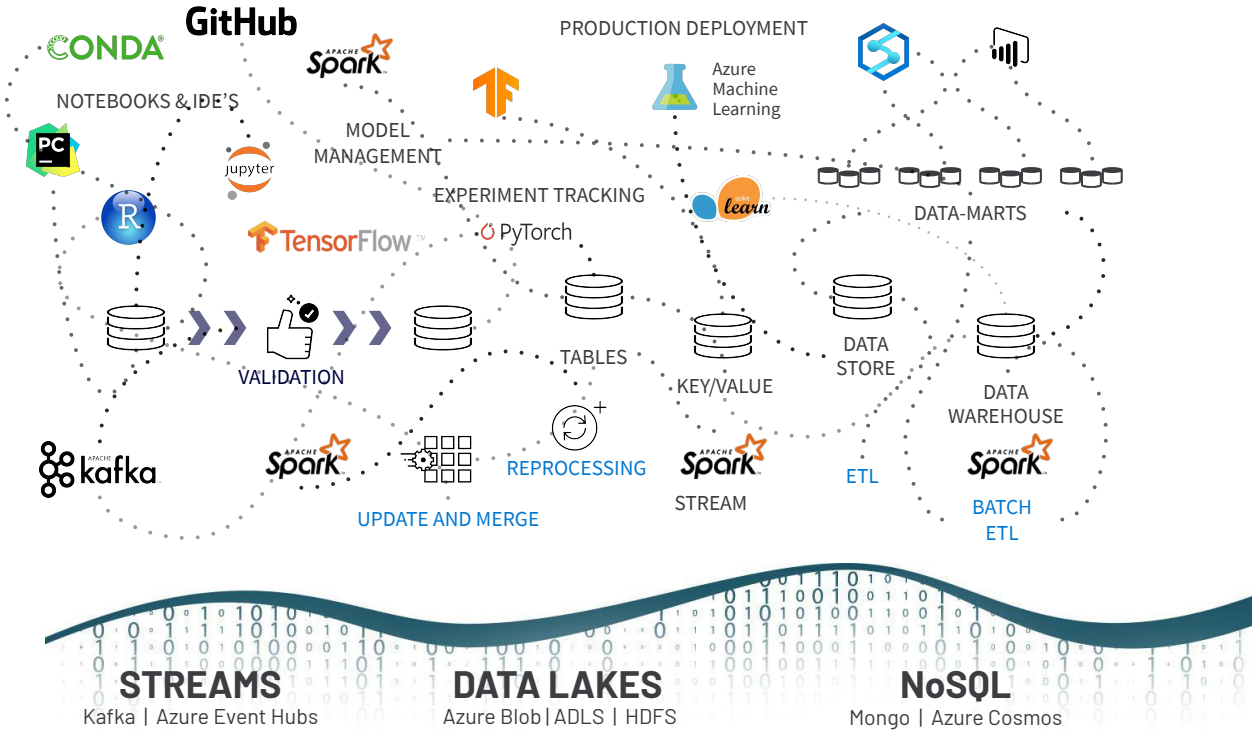
Unlocking business value

Azure Databricks

These companies combined their massive data with machine learning and analytics capabilities



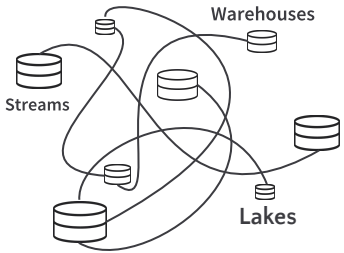
Most organizations fail to unlock business value due to data, technology and people silos



Unlocking business value: Four challenges

1

Data is messy, siloed and slow



2

ML is hard, Production is harder

Data Engineers in IT

Data Scientists in Business

3

BI is limited to a fraction of data

4

Lack of enterprise readiness

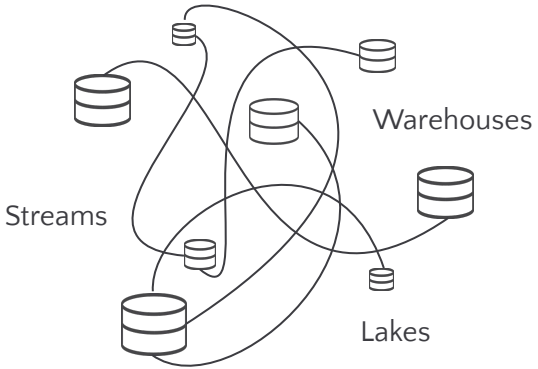
Fragmented security

Poor reliability

Disjointed governance

Make all your data ready for analytics and ML

1 Data is messy, siloed and slow



Unified Data Service

Build open, reliable, fast data lakes with all your data

Big Data

Business Data

Applications

DELTA LAKE™

Open High Quality Fast

Your Existing Data Lake

Azure Data Lake Storage

hadoop

Azure Blob Storage

APACHE Spark

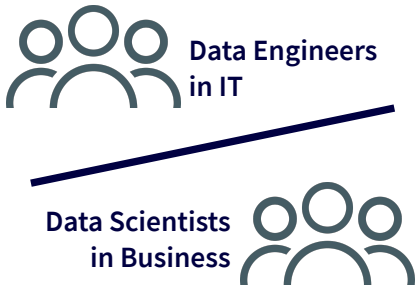
Unified Engine

BI Reporting

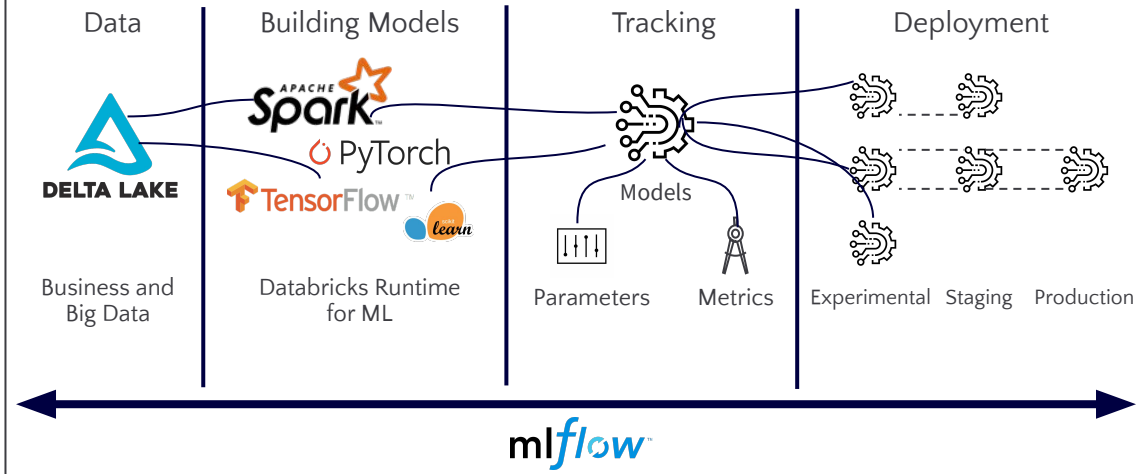
Machine Learning

Unify data and ML across the full lifecycle

2 ML is hard, Production is harder



Data Science and ML Workspace



Standardize ML lifecycle from experimentation to production

Enable analytics directly on all your source data

3 BI is limited to a fraction of data

BI Integrations for data lake

Reports | Dashboards | Ad hoc data science

DELTA LAKE

APACHE Spark

All your data with high quality and great performance

Applications: salesforce, Marketo, zendesk, Google Analytics

Data stores: mongoDB, ORACLE, MySQL, redis, kafka, cassandra

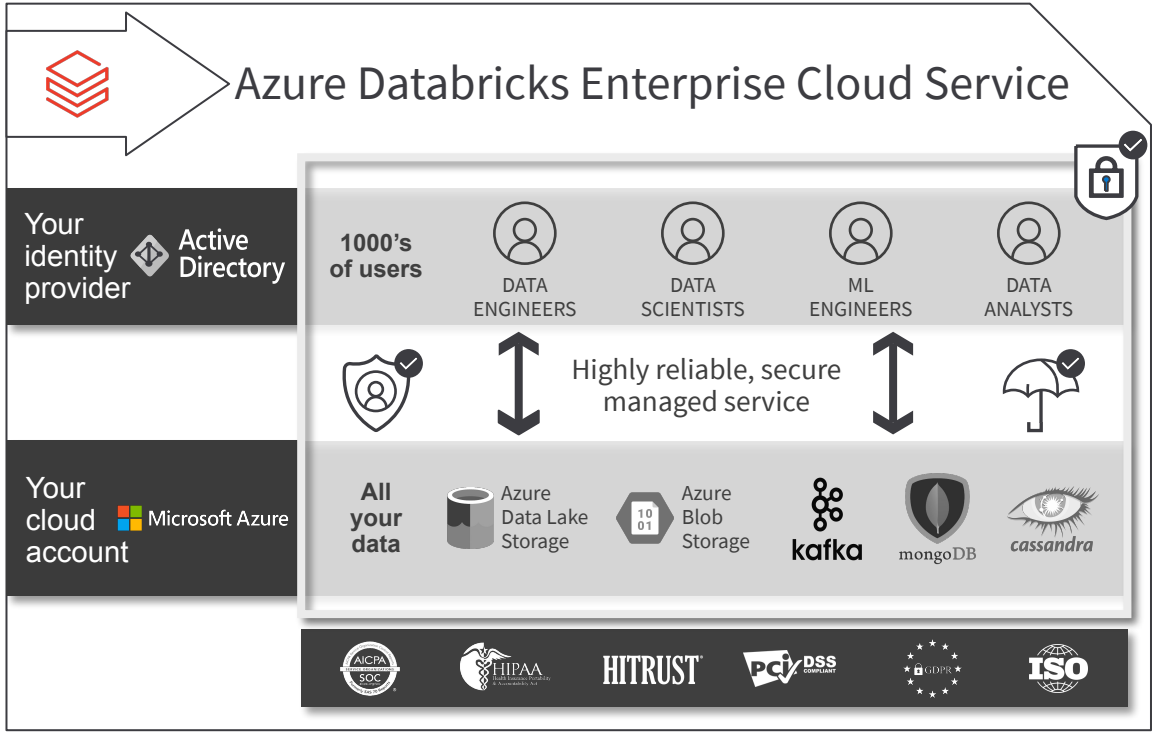
Files

Azure Databricks

Leverage cloud native platform for enterprise grade solution

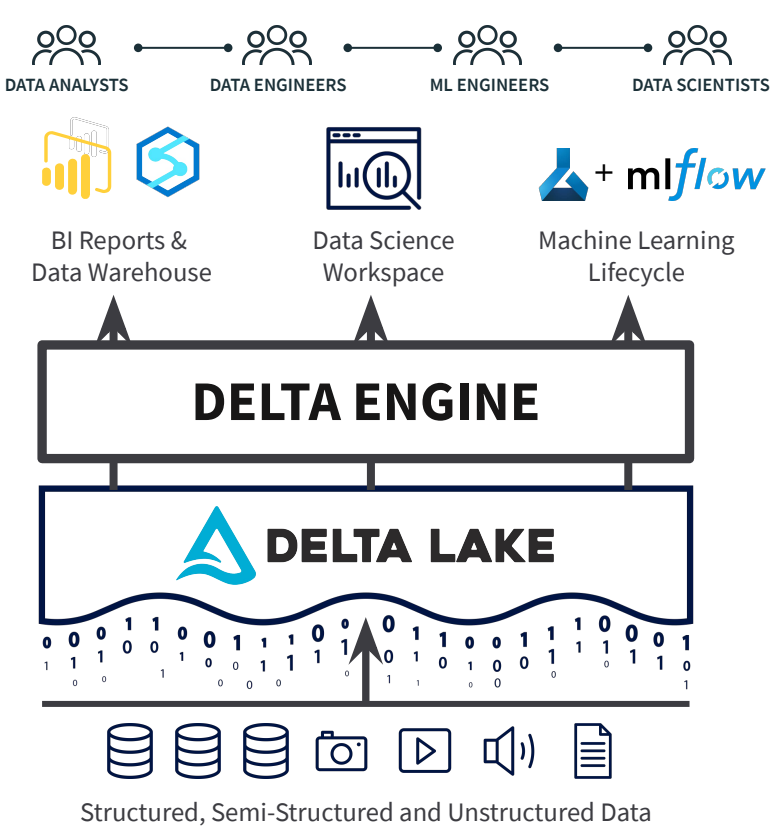
4 Lack of enterprise readiness

- Fragmented security
- Poor reliability
- Disjointed governance



Unifying Data Analytics

Azure Databricks



Unifying data teams

One architecture for every use case

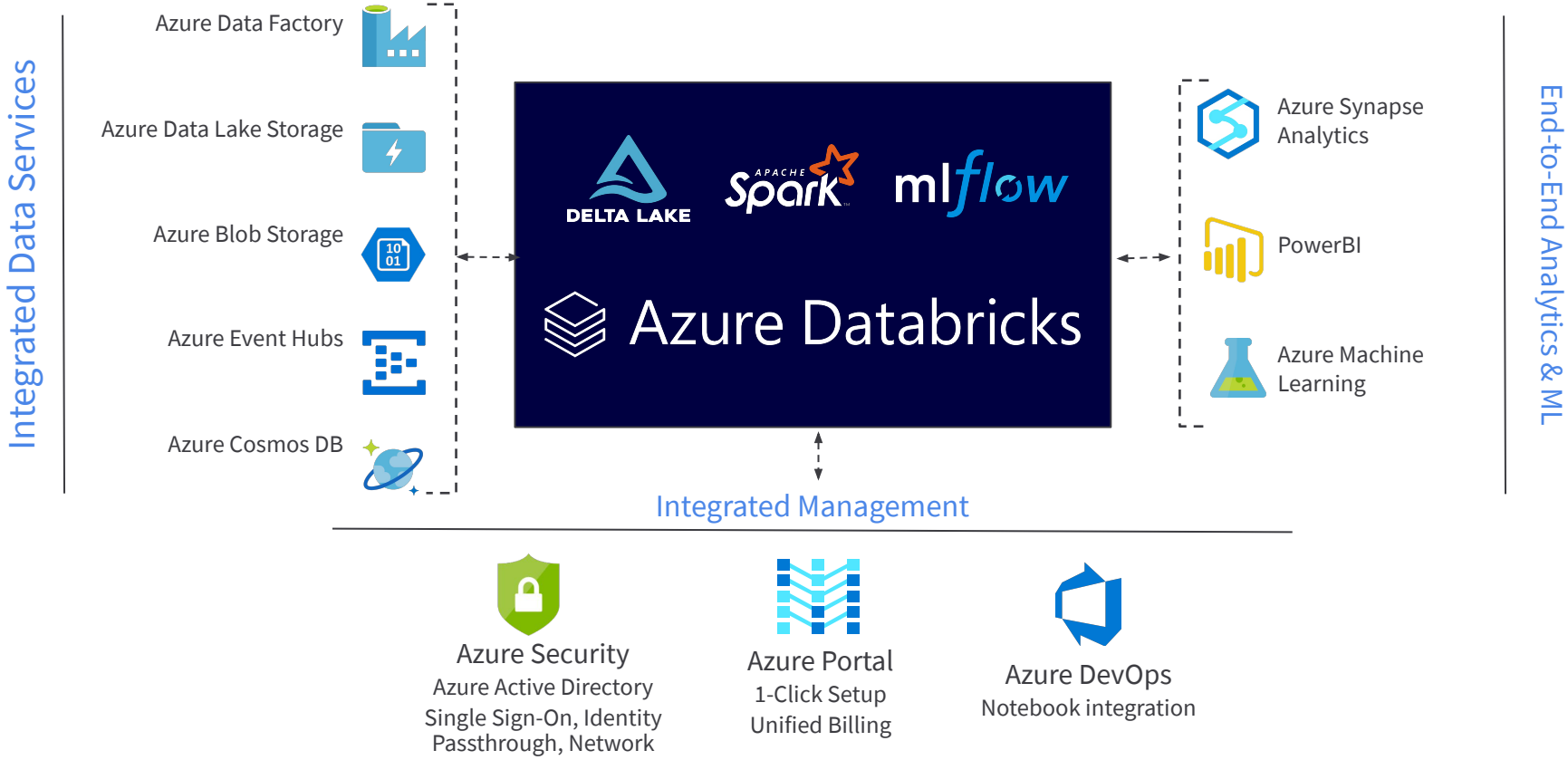
High performance query engine

Structured transactional layer

Data Lake for all your data

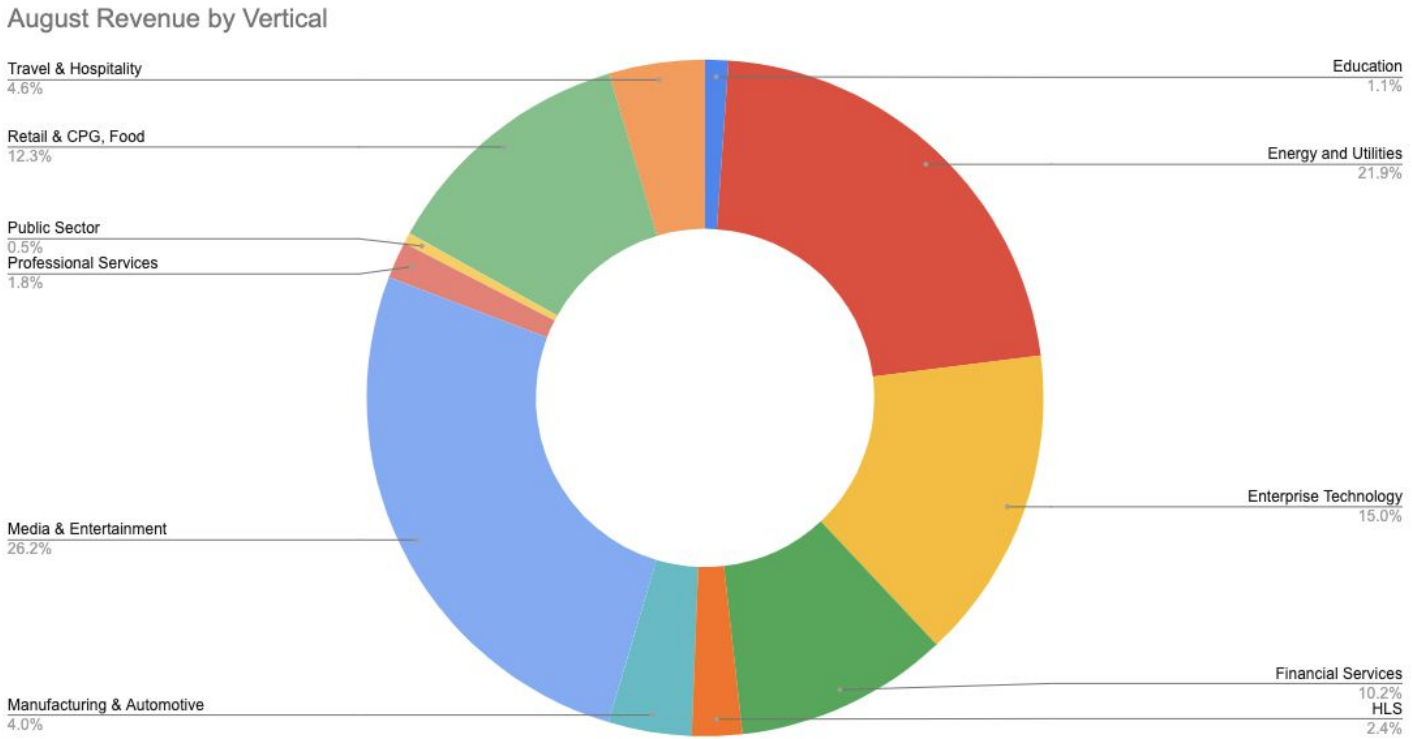
Azure Databricks

First-Party Service, Natively Integrated into Azure

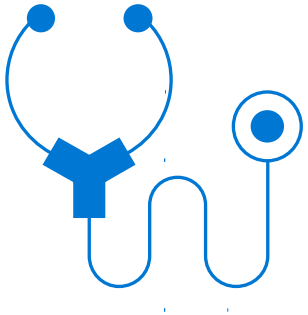


Industry Focus

ADB Revenue by vertical Asia Time zone



Innovative industry use cases



Healthcare

Genomics & DNA sequencing

Accelerated drug discovery

Smarter trial design

Efficient Inventory optimization

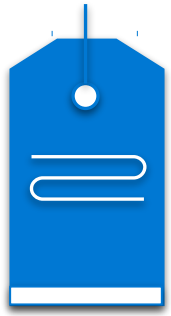
Financial services

Fraud detection and prevention

Customer 360 & personalization

Underwriting

Asset management



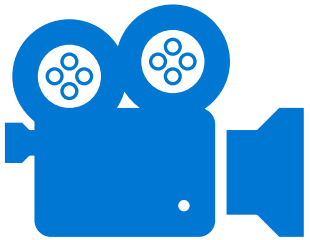
Retail/CPG

Personalize customer experience

Optimize merchandising & pricing

Inventory optimization

Last-mile delivery



Media/entertainment

Ad targeting

Content personalization

Ad inventory optimization

Sentiment analysis

Leading to transformational changes

Azure Databricks

Product recommendation

The average size of a single cart has decreased

Provide personalized digital content to shoppers

Increase cart size

ASOS delivers 15.4 million personalized experiences with 33 orders per second

Predictive maintenance

Unplanned downtime results in cost overruns

Predict when maintenance should be performed

Minimize downtime

Hybrid solution predicts onboard water usage, saving $200k/ship/year

Demand forecasting

Solar energy production is inconsistent

Align energy supply with the optimal markets

Maximize revenue

Distributed power generation increases revenue by over €100M

Databricks Industry Solution Accelerators

Retail & CPG	Banking & Fin Serv	Health & Life Sciences	Media & Comms
Demand Forecasting (Causal)	Risk / VAR Calculation	Genomics: DNASeq	Quality Of Service
Demand Forecasting (Timeseries)	Computer Vision - Claims Damage Estimation	Genomics: Live-updating QC Pipelines	Foot Traffic Ad Campaign Analysis
Safety Stock	Alternative Data	Genomics: Joint Genotyping	Content Recommendations
Customer Lifetime Value	Environment Social Governance (ESG)	Genomics: Common-variant GWAS	Addressable Advertising
Consumer Segmentation	Credit Risk and Regulations	Genomics: Rare-variant GWAS	Audience Personalization
Customer Churn Analysis	Credit Expected Loss	Genomics: Normalization/Cohort Merging	Churn Propensity Score
Next Best Offer		Medical IoT: EEG event detection	Next Best Offer

Hadoop migration

Customer Qualification:

Target Market and On-Prem Spark Migration

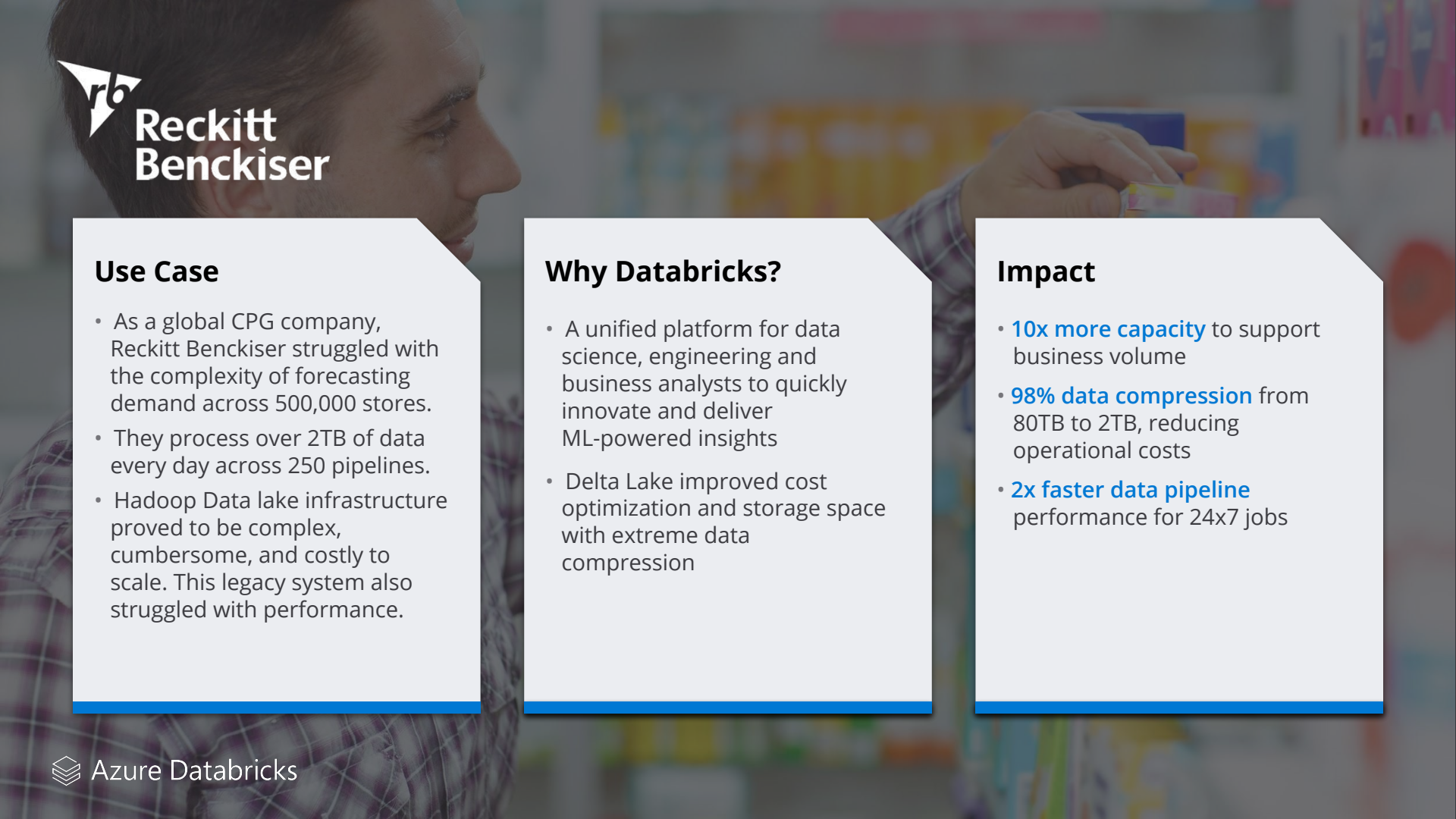
Azure Databricks

Opportunity Strength	Decision Timeline (Driven by upcoming renewal or other factors)	On-prem Big Data Cluster: Size and Config	Leadership Mandate
High Priority	Need to make cloud technology choice in 6 months OR Account is S500	20+ production nodes Hive, Spark, data engineering & analytics	Clear mandate for cloud migration w/ named sponsor
Priority	Need to make cloud technology choice in 12 months	20+ production nodes Hive, Spark, data engineering & analytics	Emerging mandate for cloud migration.
Case-by-Case	No clear decision timeline (but can be influenced to be within 12 months)	20+ production nodes Hive, Spark, data engineering & analytics	Leadership is considering cloud options Evaluations/POCs in process
Not a Good Candidate	No cloud migration within 12 months	Predominantly HBase and/or Storm usage (without any data engineering)	No exec support No cloud migration sponsor

Powering Innovation with Modern Data Analytics

Customers that migrated from On-premises Big Data Lake

Azure Databricks



Reckitt Benckiser

Use Case

- As a global CPG company, Reckitt Benckiser struggled with the complexity of forecasting demand across 500,000 stores.
- They process over 2TB of data every day across 250 pipelines.
- Hadoop Data lake infrastructure proved to be complex, cumbersome, and costly to scale. This legacy system also struggled with performance.

Why Databricks?

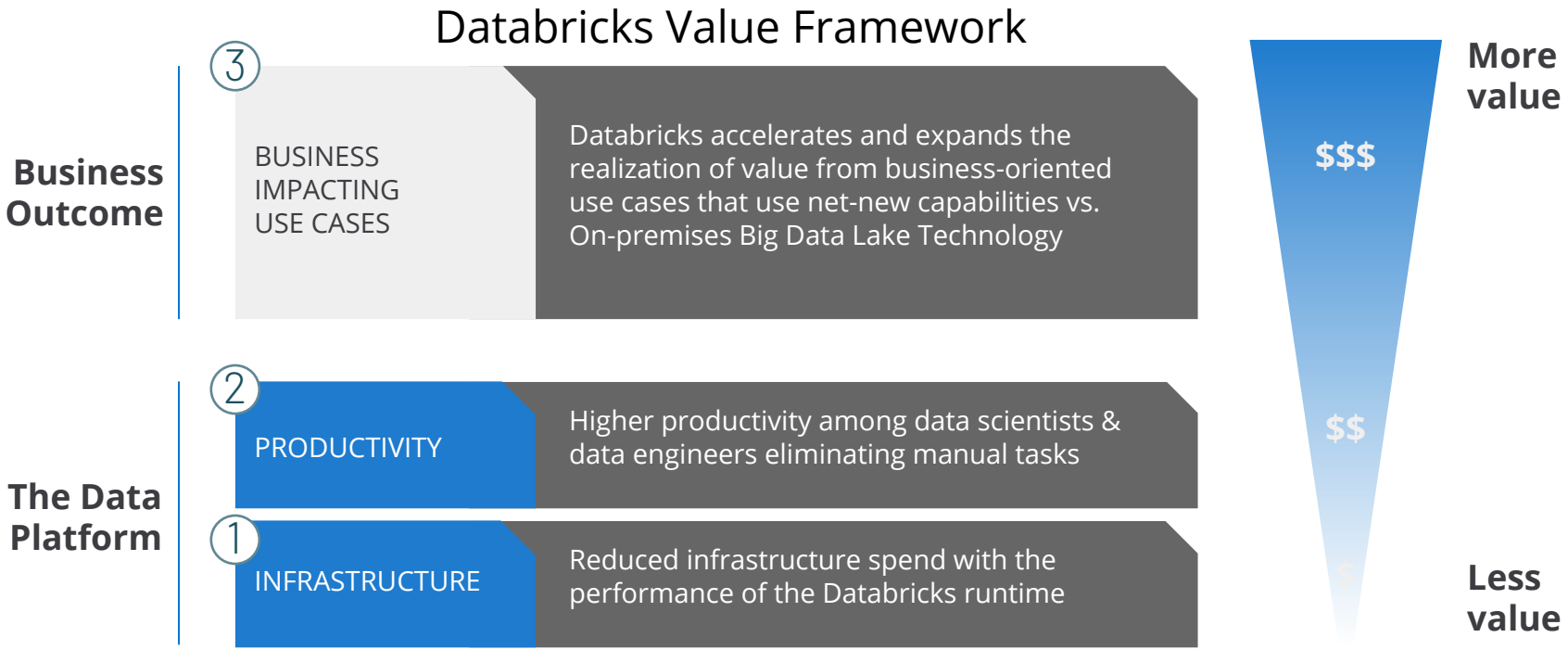
- A unified platform for data science, engineering and business analysts to quickly innovate and deliver ML-powered insights
- Delta Lake improved cost optimization and storage space with extreme data compression

Impact

- **10x more capacity** to support business volume
- **98% data compression** from 80TB to 2TB, reducing operational costs
- **2x faster data pipeline** performance for 24x7 jobs

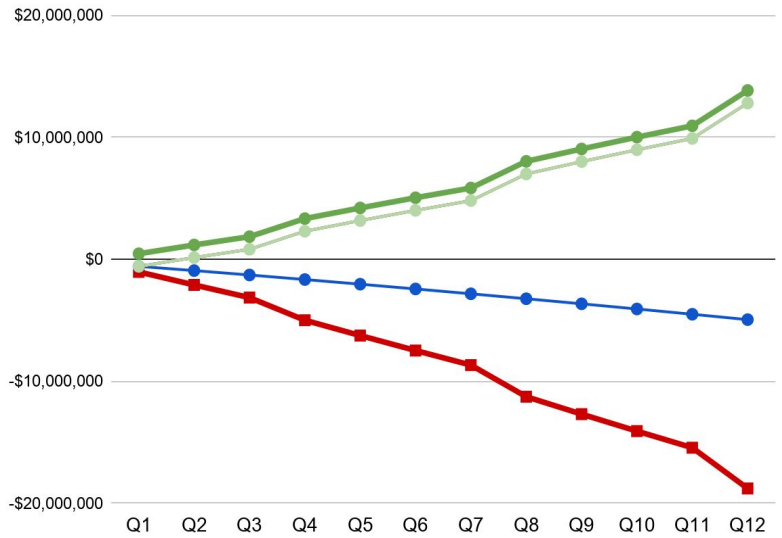
Azure Databricks

Databricks Drives New Business Value at 3 Levels



$17.7M in value delivered with Databricks

On-premises **costs** vs. Databricks **value** & **investment**
Units: $ Cumulative PV over 3 years



Source: Databricks value model

$13.8 M Potential value with Databricks

$12.8 M **Net impact** Includes cost of both solutions during migration

-$4.9 M Investment - Databricks, migration & cloud

-$18.8 M Current technology - Cost of inaction

Databricks customer example:
Large U.S. Telco,
156 node cluster

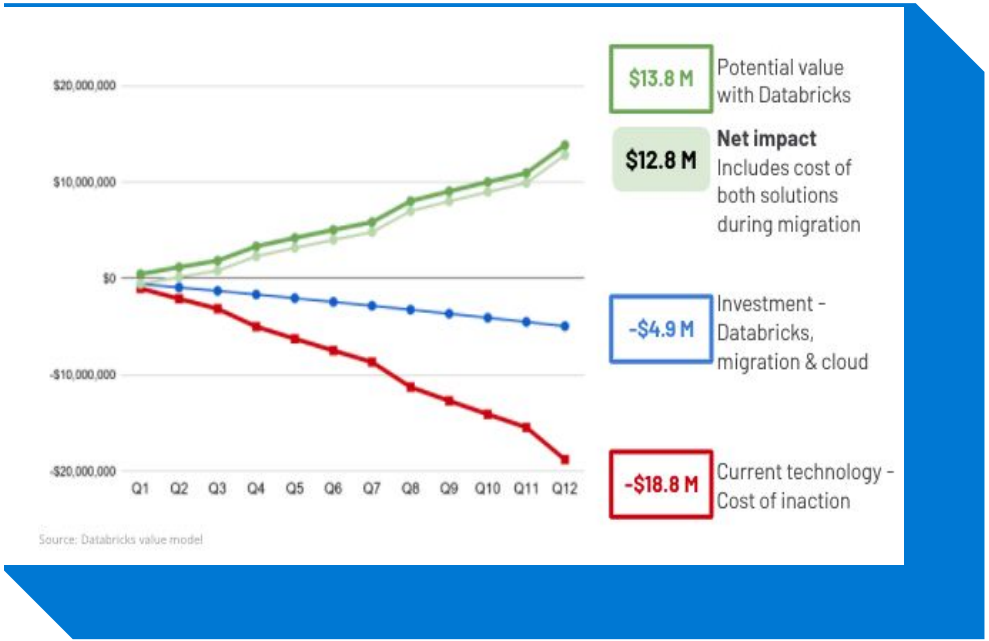
Value of Databricks
- Avoided Data lake licensing
- No need to add expensive new hardware for additional capacity
- Avoided data center costs
- Avoided Data-lake admin costs

Databricks investment
- Databricks usage & support
- Migration
- Cloud compute

On-premises Data-lake costs
- Data center, Cluster administration, new hardware, licensing

Work with us for a Tailored Value Case for Your Migration

Tailored Financial Analysis

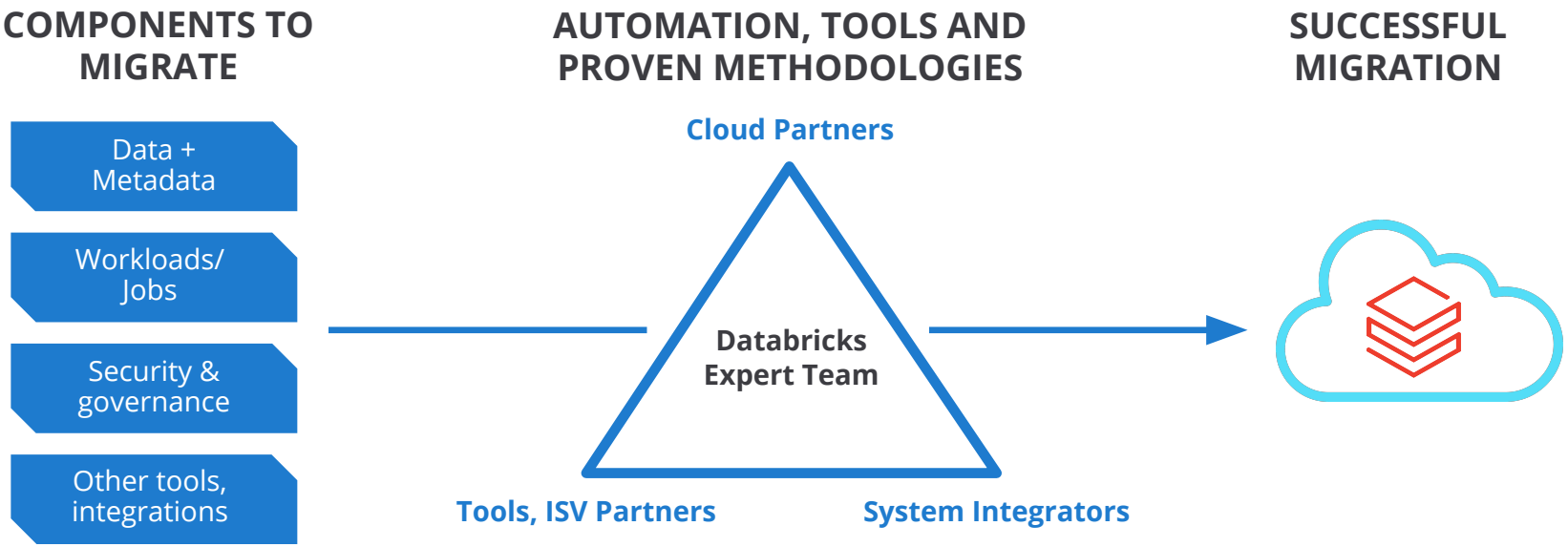


Customer example

Tailored business case to be produced by answering 4 core questions:

1. How many **nodes** in your on-premises Big Data lake cluster ?
2. How many **people** support your On-premises environment?
3. When is your **renewal** with the On-premises Data lake vendor ?
4. How do you expect your **data needs to grow** over time?

Proven Migration Strategy: Reduce Risk, Costs



Strategy Options: **Lift & shift** (faster, automatable) **Transformation** (higher impact)

Save up to 52% when migrating to Azure Databricks

Microsoft Azure

Contact

Overview Solutions Products Documentation Pricing Training Marketplace Partners Support Blog

Now save up to 52 percent when migrating to Azure Databricks

Posted on July 15, 2020

Bharat Sandhu, Director, Azure Data & AI

More than ever before, companies are relying on their big data and artificial intelligence (AI) systems to find new ways to reduce costs and accelerate decision-making. However, customers using on-premises systems struggle to realize these benefits due to administrative complexity, inability to scale their fixed infrastructure cost-effectively, and lack of a shared collaborative environment for data engineers, data scientists and developers.

To make it easier for customers to modernize their on-premises Spark and big data workloads to the cloud, we're announcing a new migration offer with Azure Databricks. The offer includes:

- **Up to a 52 percent discount** over the pay-as-you-go pricing when using the Azure Databricks Unit pre-purchase plans. This means that customers can free themselves from the complexities and constraints of their on-premises solutions and realize the benefits of the fully managed Azure Databricks service at a significant discount.
- **Free migration assessment** for qualified customers.

Azure Databricks is a fast, easy, and collaborative Apache Spark-based service that simplifies building big data and AI solutions. Since its debut two years ago, Azure Databricks has experienced significant adoption from customers, such as Shell, Cerner, Advocate Aurora Health, and Bosch, which are using it to run mission-critical big data and AI workloads.

We've also seen several customers accelerating their migration of on-premises systems to Azure Databricks for the following reasons:

Offer summary:

- **Up to a 52 percent discount** over the pay-as-you-go pricing when using the Azure Databricks Unit pre-purchase plans.
- **Free migration assessment** for qualified customers.

Learn more at:

dbricks.co/SaveUpTo52

On-premises Big Data Migration Technical Assessment
2 Day Workshop

Identify

Existing Workloads

Infrastructure and Security Requirements

Prioritize

Identify Priority Workloads

Dependent Schedules

Roadmap & Value

Assess New Workloads

Business Value

Target State

Cloud Native Reference Architecture

Migration Schedule

Prescriptive recommendations

Customized On-premises Big Data Lake Migration Success Plan with a Free Expert Assessment

1 Pre-questionnaire + 2 day workshop led by experts

- Learn about how Databricks works and how your current workloads, tools and processes map and transform in the future state in cloud

2 Technical, Use-case and Business Value analysis

- High level current and future state architecture, discuss use-cases and prioritize them, understand how $$ value is driven with the migration

3 Proposal and Recommendations for path forward

- The expert team will summarize all the findings and walk through the proposed costs, business value summary and recommended migration plan

Our Partner Ecosystem will Accelerate Migrations

ISV Partners and Migration Tools

Security Governance

Cloud

Azure Databricks

Migration SWAT team +

CS Packaged Services For Migration

Infoworks

StreamSets

Knowledge Lens
Accelerating Actionable Insights

precisely

WANDISCO

talend

Qlik

Informatica

STREAMANALYTIX

PRIVACERA

IMMUTA

collibra

Alation

Azure

Cognizant

accenture

Capgemini

Infosys

avanade

BLUEGRANITE

Celebaltech

Tech Mahindra

TATA CONSULTANCY SERVICES

Mindtree

wipro

TRACE3

phData

slalom

PARIVEDA

IMPETUS

Consulting & SI Partners

Driving consumption

BUSINESS IMPACT

Company Goal:

Text

IMPROVE REVENUE GROWTH

Use Case
With enhanced customer insights, enterprises can deliver relevant offers, helping both to lower the cost of customer acquisition and drive revenue growth.

Metric (Positive)

Provide text of expected outcomes

COST SAVINGS & PRODUCTIVITY

Use Case
With enhanced customer insights, enterprises can deliver relevant offers, helping both to lower the cost of customer acquisition and drive revenue growth.

Metric (Positive)

Provide text of expected outcomes

RISK MITIGATION

Use Case
With enhanced customer insights, enterprises can deliver relevant offers, helping both to lower the cost of customer acquisition and drive revenue growth.

Metric (Positive)

Provide text of expected outcomes

Business challenges

Azure Databricks

Challenges	Use case	Desired outcomes/metrics
TEXT	TEXT	Details. Describe the success/results of the solution.
TEXT	TEXT	Details. Describe the success/results of the solution.
TEXT	TEXT	Details. Describe the success/results of the solution.
TEXT	TEXT	Details. Describe the success/results of the solution.

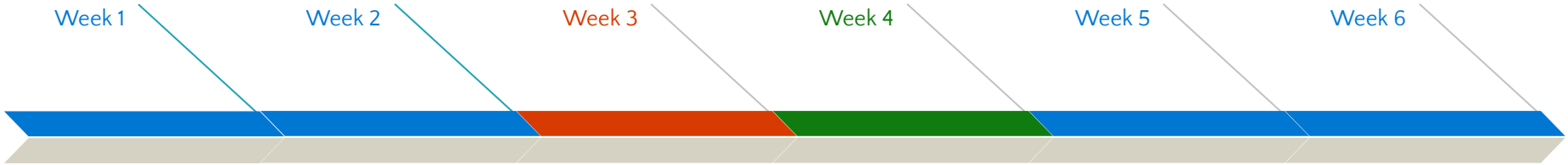
Initial use cases

Azure Databricks

Use case	How do we support the use case
Use case	How do we support the use case
Use case	How do we support the use case
Use case	How do we support the use case

Success Plan/Deployment Plan

Azure Databricks



Planning

Identify Key Use Cases

Designate Project Lead

Select Partner(s)

Create Test Plan (1–2 hours)

Prerequisites

Security Review

Executive Sync

Partner Sync

Installation

Slack Channel Setup

Pre-testing

Data Loading

Confirm User Setup

Test to Confirm Data Source Connectivity

Testing

Day 1—Working Session (3–4 Hours)

Day 4—Working Session (1 Hour)

Day 7—Working Session (1 Hour)

With Partner(s)

Review

POC Readout

Proposal Review

Executive Summary and Proposal

Negotiation/ Procurement

Onboarding

Introduction with CS/PS teams

Partner Delivery Assurance

New use case development and delivery

Technical validation on Azure Databricks

Azure Databricks

1

Pick a use case(s)

Identify a current initiative the will drive business value
and the success metrics to be measured by the business

2

Enable Databricks platform with secure data

Ensure datasets are in a secure within cloud environment
or can be accessed securely from the cloud

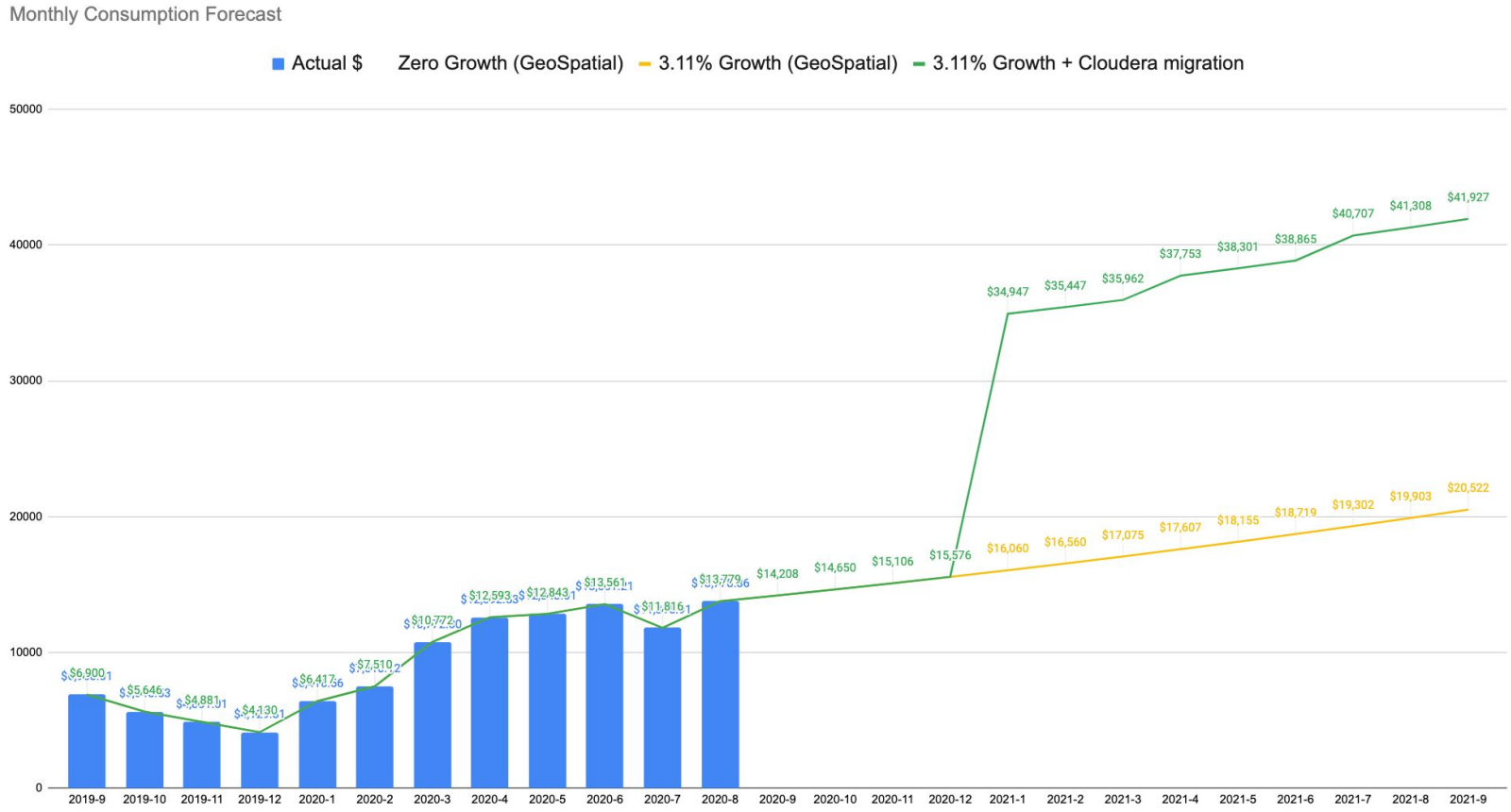
3

Co-develop what success looks like

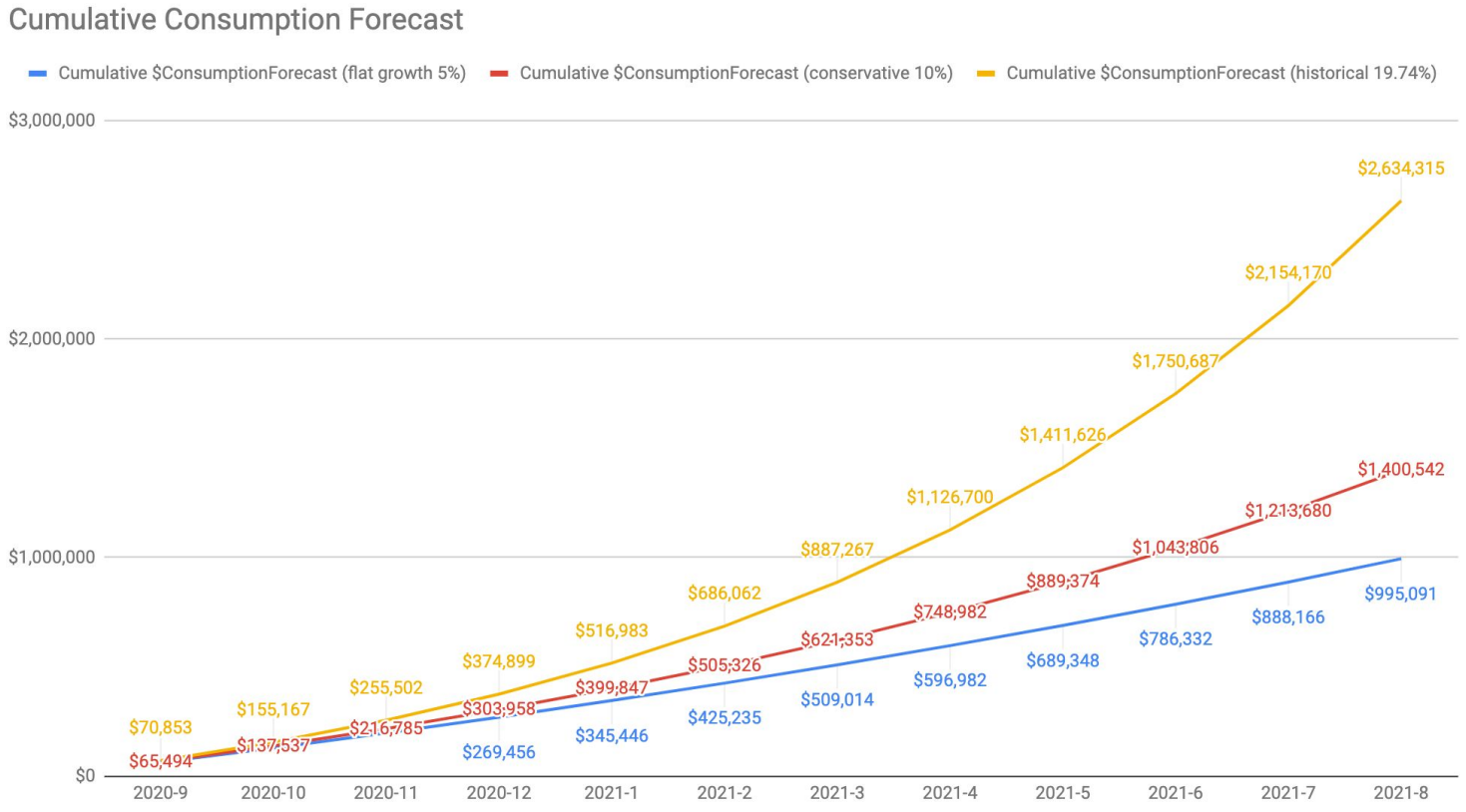
Have your teams work with Databricks and partners to
show value in weeks and months; expand out to other use cases

ADB Pre Purchase plans and Big Data migration promotion

Geospatial Use Cases monthly consumption

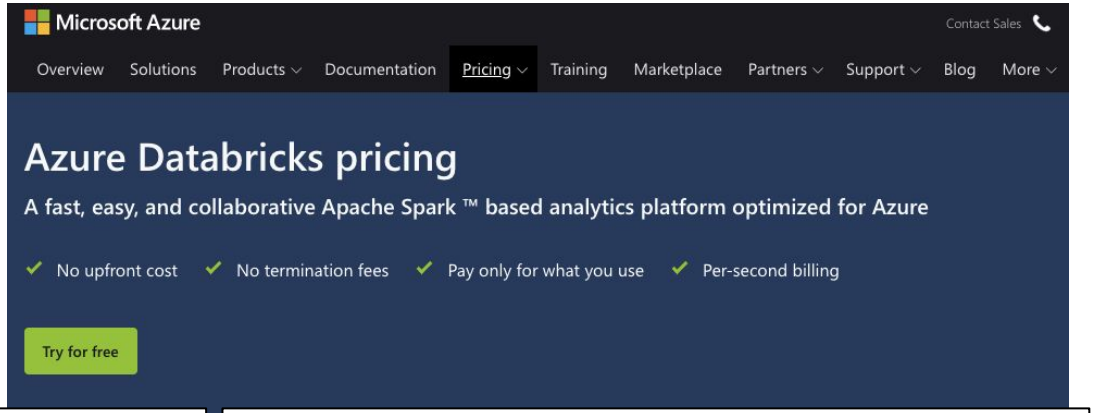


Expected growth over next 12 months



ADB Pre-Purchase plan

[Big Data migration to ADB link]



1-year pre-purchase plan

DATABRICKS COMMIT UNIT (DBCU)	PRICE (WITH DISCOUNT)	DISCOUNT
25,000	$23,500	6%
50,000	$46,000	8%
100,000	$89,000	11%
200,000	$172,000	14%
350,000	$287,000	18%
500,000	$400,000	20%
750,000	$578,000	22%
1,000,000	$730,000	27%
1,500,000	$1,050,000	30%
2,000,000	$1,340,000	33%

3-year pre-purchase plan

DATABRICKS COMMIT UNIT (DBCU)	PRICE (WITH DISCOUNT)	DISCOUNT
75,000	$69,000	8%
150,000	$135,000	10%
300,000	$261,000	13%
600,000	$504,000	16%
1,050,000	$819,000	22%
1,500,000	$1,140,000	24%
2,250,000	$1,640,000	27%
3,000,000	$2,070,000	31%
4,500,000	$2,970,000	34%
6,000,000	$3,780,000	37%

Strategic Customer Program at-a-Glance

When meeting the criteria, value can be accelerated even further

Executive Sponsorship	Strategic customers have a **close relationship with our** business, product, delivery, Spark, ML and AI **leadership**.
Spark Expertise at Scale	Building advanced skill sets for Data Engineers, Data Scientists, and Analysts. **Support for conferences and hackathons**
Customer Success Engineers	**Technical and project guidance** for Customer Use Cases to drive business value
Resident Solutions Architect	Dedicated access to named highly technical advisory and **hands-on resources**
Technical Support	Tech Support available via a **multiple channels**
Prioritize Product Management	Early visibility to product roadmaps, private previews, product advisory boards to **influence product roadmap direction**

PRICING OPTIONS

OPTION 1

 # DBU

Interactive DBU: 800,000 (20%)
Automated DBU: 200,000 (80%)

$ PRICE

Optional
Text

Text

OPTION 2

 # DBU

Interactive DBU: 800,000 (20%)
Automated DBU: 200,000 (80%)

$ PRICE

Optional
Text

Text

OPTION 3

 # DBU

Interactive DBU: 800,000 (20%)
Automated DBU: 200,000 (80%)

$ PRICE

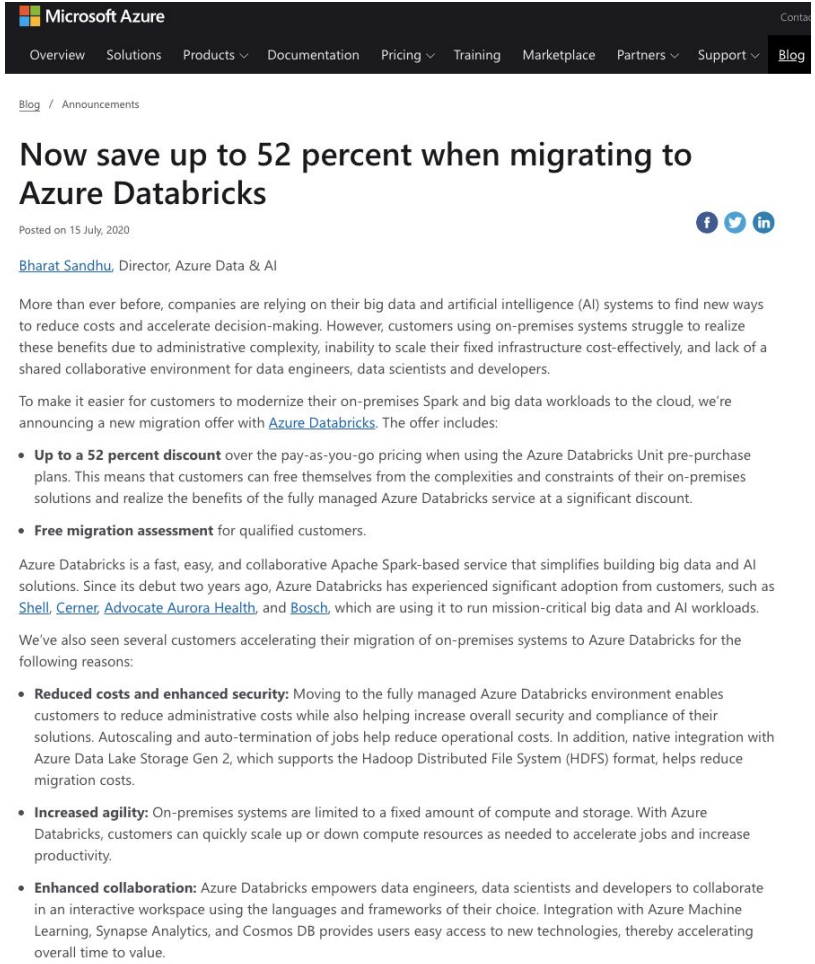
Optional
Text

Text

Big Data Migration to ADB

[Big Data migration to ADB link]

Not Available to CSP customers



Microsoft Azure

Overview Solutions Products Documentation Pricing Training Marketplace Partners Support Blog

Blog / Announcements

Now save up to 52 percent when migrating to Azure Databricks

Posted on 15 July, 2020

Bharat Sandhu, Director, Azure Data & AI

More than ever before, companies are relying on their big data and artificial intelligence (AI) systems to find new ways to reduce costs and accelerate decision-making. However, customers using on-premises systems struggle to realize these benefits due to administrative complexity, inability to scale their fixed infrastructure cost-effectively, and lack of a shared collaborative environment for data engineers, data scientists and developers.

To make it easier for customers to modernize their on-premises Spark and big data workloads to the cloud, we're announcing a new migration offer with Azure Databricks. The offer includes:

- **Up to a 52 percent discount** over the pay-as-you-go pricing when using the Azure Databricks Unit pre-purchase plans. This means that customers can free themselves from the complexities and constraints of their on-premises solutions and realize the benefits of the fully managed Azure Databricks service at a significant discount.
- **Free migration assessment** for qualified customers.

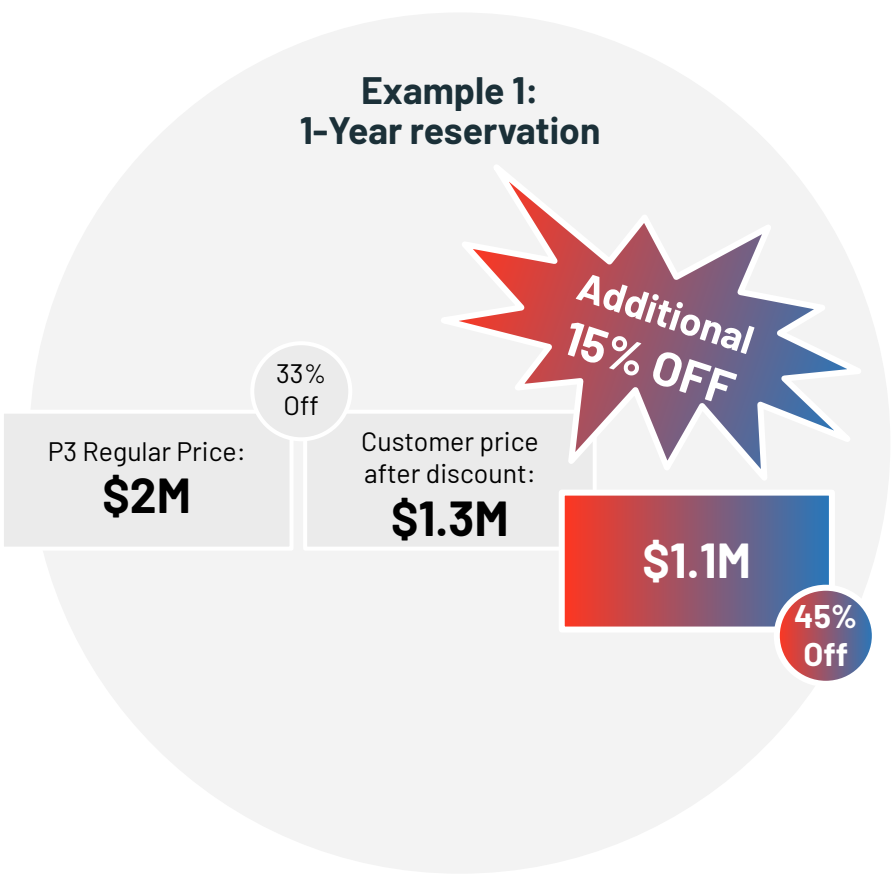
Azure Databricks is a fast, easy, and collaborative Apache Spark-based service that simplifies building big data and AI solutions. Since its debut two years ago, Azure Databricks has experienced significant adoption from customers, such as Shell, Cerner, Advocate Aurora Health, and Bosch, which are using it to run mission-critical big data and AI workloads.

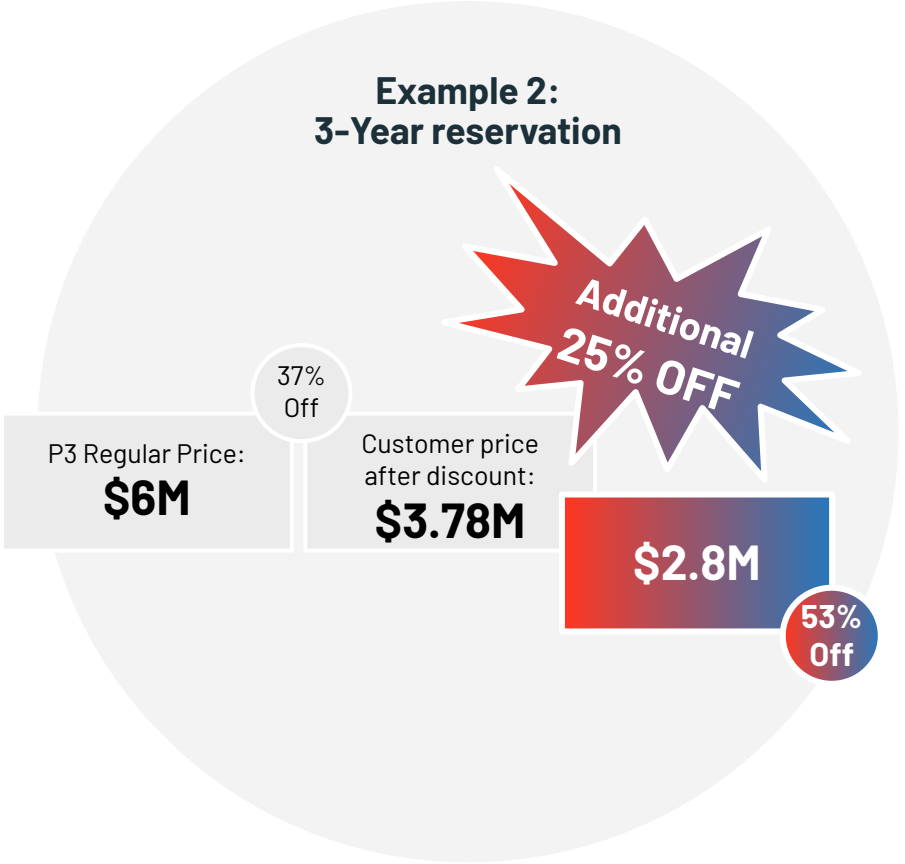
We've also seen several customers accelerating their migration of on-premises systems to Azure Databricks for the following reasons:

- **Reduced costs and enhanced security:** Moving to the fully managed Azure Databricks environment enables customers to reduce administrative costs while also helping increase overall security and compliance of their solutions. Autoscaling and auto-termination of jobs help reduce operational costs. In addition, native integration with Azure Data Lake Storage Gen 2, which supports the Hadoop Distributed File System (HDFS) format, helps reduce migration costs.
- **Increased agility:** On-premises systems are limited to a fixed amount of compute and storage. With Azure Databricks, customers can quickly scale up or down compute resources as needed to accelerate jobs and increase productivity.
- **Enhanced collaboration:** Azure Databricks empowers data engineers, data scientists and developers to collaborate in an interactive workspace using the languages and frameworks of their choice. Integration with Azure Machine Learning, Synapse Analytics, and Cosmos DB provides users easy access to new technologies, thereby accelerating overall time to value.

Offer Example

Azure Databricks





DBU pre-purchase pricing

Azure Databricks

1-Year			Migration Offer		
DBU list price	Discount	Customer price after discount	Additional Promo Discount	Final Customer Price	Overall discount
$25 K	6%	$24 K	N/A	N/A	N/A
$50 K	8%	$46 K	N/A	N/A	N/A
$100 K	11%	$89 K	15% Off	**$75.7 K**	**24%**
$200 K	14%	$172 K		**$146.2 K**	**27%**
$350 K	18%	$287 K		**$244 K**	**30%**
$500 K	20%	$400 K		**$340 K**	**32%**
$750 K	22%	$578 K		**$497 K**	**34%**
$1 M	27%	$730 K		**$620.5 K**	**38%**
$1.5 M	30%	$1.05 M		**$900 K**	**41%**
$2 M	33%	$1.34 M		**$1.1 M**	**43%**

3-Year			Migration Offer		
DBU list price	Discount	Customer price after discount	Additional Promo Discount	Final Customer Price	Overall discount
$75 K	8%	$69 K	N/A	N/A	N/A
$150 K	10%	$135 K	25% Off	**$101 K**	**33%**
$300 K	13%	$261 K		**$196 K**	**35%**
$600 K	16%	$504 K		**$378 K**	**37%**
$1.05 M	22%	$819 K		**$614 K**	**42%**
$1.5 M	24%	$1.14 M		**$855 K**	**43%**
$2.25 M	27%	$1.64 M		**$1.2 M**	**45%**
$3 M	31%	$2.07 M		**$1.6 M**	**48%**
$4.5 M	34%	$2.97 M		**$2.2 M**	**51%**
$6 M	37%	$3.78 M		**$2.8 M**	**52%**

Azure Databricks

Fast, easy, and collaborative Apache Spark–based analytics platform optimized for Azure.

Microsoft 1st party service.

Native integration with Azure.

Enterprise Ready.

Fastest growing Data & AI service on Azure.
Drives consumption & EA burndown.
Retires ACR.

1. Engage your Databricks seller in account planning to identify and drive opportunities.
2. Learn more about the new DBU pre-purchase plan (faster & easier to use!)
3. Invite customers to Azure Databricks Virtual Workshops and Webinars
4. Sell through use cases and business value.