

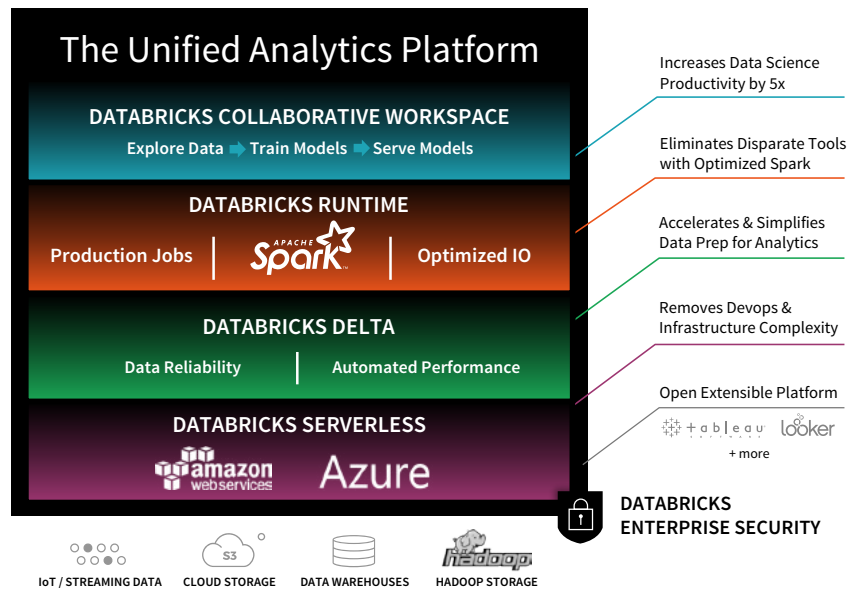
# Databricks for Data Engineering

## Build Fast and Reliable Data Pipelines

As companies set their sights on making data-driven decisions or automating business processes with cutting edge technologies such as machine learning and artificial intelligence, mastering data engineering is an essential step to ensure the infrastructure is in place to operationalize data pipelines needed to perform analytics against a growing volume of data from multiple sources. The key to success for a data engineer is to be armed with the right technologies and tools to perform mission-critical data cleansing, transformations, and manipulations, to make business use cases such as real-time dashboards or fraud detection possible.

## Better Data Engineering with Databricks

Founded by the team who created Apache Spark™, Databricks provides a Unified Analytics Platform that accelerates innovation by unifying data science, engineering, and business. With Databricks, data engineers can securely and reliably deploy production data pipelines with ease.



## Accelerate Performance with Databricks Runtime, Built on Apache Spark

### DATABRICKS IO

Leverages a vertically integrated stack to optimize the I/O layer and processing layer to significantly improve the performance of Spark in the cloud.

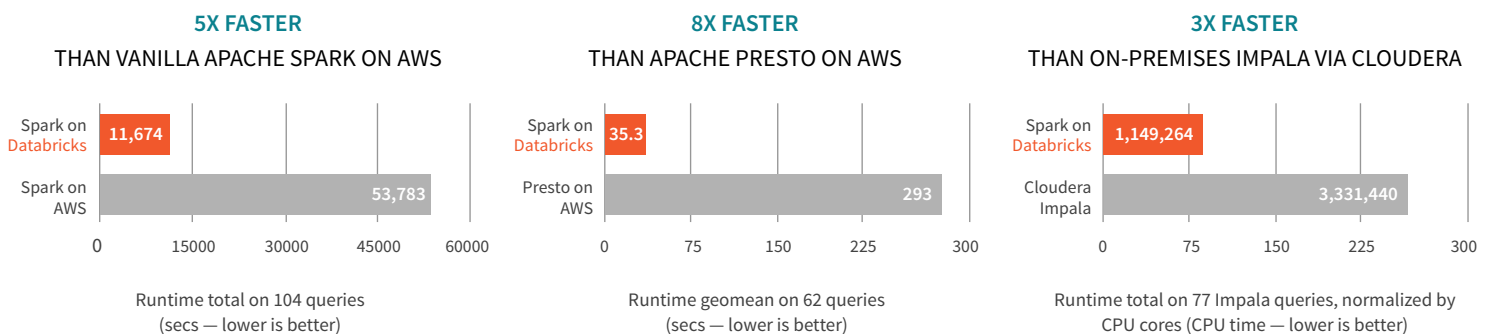
### DATABRICKS SERVERLESS

A serverless architecture that democratizes infrastructure through the auto-configuration and scaling of compute resources — enabling best-in-class performance at dramatically lower costs.

### FULLY MANAGED IN THE CLOUD

A cloud-native platform that abstracts the complexities of big data infrastructure, resulting in a highly elastic, reliable and performant platform to build innovative products.

## The Fastest Big Data Platform in the Cloud



Read the blog: [databricks.com/cloud-benchmarks](https://databricks.com/cloud-benchmarks)

# Streamline Processes from ETL to Production

## PRODUCTION WORKFLOWS

A unified platform that streamlines end-to-end workflows from data ingest and ETL, to data exploration and model building, to productionizing models and data products.

## UNIFYING ALL ANALYTICS

Move seamlessly across various types of analytics including batch, ad hoc, machine learning, deep learning, stream processing, and graph.

## ROBUST INTEGRATIONS

Plug into a wide variety of AWS tools and data stores with built-in connectors and integrate with other data engineering services to facilitate CI/CD with comprehensive APIs.

*We were able to reduce data processing time from **48 hours** to **45 minutes** with Databricks.*

*– Dennis Vallinga, Business Analyst, Shell*

# Our Spark Expertise is our Edge

## SUPPORT

Unparalleled Apache Spark support by the creators of Apache Spark.

## SERVICES

Faster innovations with Databricks and Spark with solution architecting and workload optimization services.

## ALWAYS AVAILABLE

Around-the-clock coverage to ensure problems are resolved quickly, with response times as fast as one hour for production tier support.

## ENGINEER RESOURCES

Online library of documentation, best practices, user guides, and other technical resources.

# Databricks Enterprise Security

## STRONG DATA ENCRYPTION

Benefit from best-in-class data protection at rest and in motion.

## INTEGRATED IDENTITY MANAGEMENT

Seamless integration with enterprise identity providers via SAML 2.0 and Active Directory.

## ROLE-BASED ACCESS CONTROLS

Fine-grained management access to every component of the enterprise data infrastructure, including files, clusters, code, application deployments, and dashboards.

## MONITOR AND AUDITING

Tap into comprehensive audit logs to monitor and troubleshoot issues.

## COMPLIANCE STANDARDS

Databricks has successfully completed SOC 2 Type 2 certification and can offer a HIPAA-compliant solution.

# Lower TCO

## BETTER PERFORMANCE

Performance-tuned clusters allow you to complete jobs in a shorter time, reducing cloud compute costs.

## FULLY-MANAGED CLUSTERS

Further reduce costs by avoiding the time-consuming tasks to build, configure, and maintain complex Spark infrastructure.

## PAY FOR ONLY WHAT YOU USE

Billing up to the nearest second keeps your costs down.

## PRICED FOR DATA ENGINEERING

Lower price point for data engineering production workloads.

# Data Engineering, Simplified

Databricks' Unified Analytics Platform removes the complexity of data engineering while accelerating performance of data engineering tasks from data access to ETL, allowing engineers to build fast and reliable data pipelines more easily to support the business.

VIACOM



Hotels.com

overstock



Mc Graw Hill Education

Get started with Databricks for data engineering today with a **free trial**.

