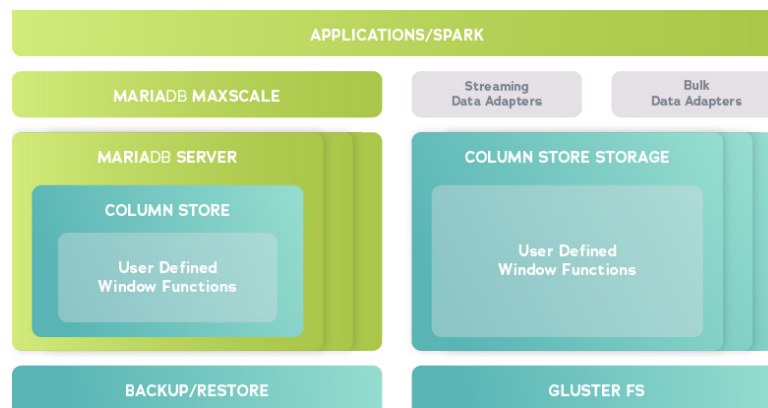# MARIADB AX + SPARK

Artificial intelligence and machine learning (ML) adoption among various industries opens up new opportunities. For example, in retail, data scientists apply machine learning to customer data to uncover common patterns and better predict shopping behavior. In manufacturing, machine learning is applied to predict machine failure which can reduce unplanned production downtime, improve cost-efficiency and increase performance and production.

With the mainstream use of machine learning analytics, Apache Spark has become one of the most popular platforms for machine learning libraries. Spark provides machine learning algorithms such as classification, regression, clustering and collaborative filtering in Scala, Python, R and Java so that data scientists can easily apply those algorithms in large datasets with in-memory performance. The challenge is that many data scientists find it difficult to operationalize their insights.

To make valuable insights available faster and easier, MariaDB AX introduces Spark Connector to stream the data insights found in Spark to the data warehouse, MariaDB AX. MariaDB Spark Connector takes the result set in Spark DataFrame and sends it to MariaDB ColumnStore via the bulk write API. By leveraging ColumnStore's streaming data adapter, data can be ingested at high speed – 100x faster than generic JDBC.

Spark ML can also be applied directly on the data in ColumnStore using standard JDBC.

MARIADB AX ARCHITECTURE

* See mariadb.com/subscription-services-policies for details.

In addition to Spark Connector, MariaDB AX provides various ways to ingest data – via the bulk data adaptor and streaming data adaptor using Kafka. Those data can be further analyzed with a scalable, highperformance columnar database without the need for complex, time-consuming batch processes, schemas and indexes optimized for a small number of predefined queries or expensive, proprietary appliances.

mariadb.com