Performing SQL Queries on Streaming Data

Janani Ravi
CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

Tracking pipelines using metrics

Counter, distribution, and gauge metrics

Using SqlTransform to create a transform from a SQL query

Using Metrics in Beam

Metrics

Track number of processed elements

Track errors during processing

Track requests made to external APIs

Metrics

Every metric is associated with a namespace and a name

Each metric is reported against a specific step in the pipeline

Metrics are created dynamically at runtime

If a runner does not support a metric, metric updates are dropped gracefully

Metrics

Counter metrics: Report a single value that can be incremented or decremented

Distribution metrics: Hold the distribution of reported values

Gauge metrics: Hold the last seen value of all values reported

Demo

Using counter metrics in Beam pipelines

Demo

Using distribution metrics in Beam pipelines

Demo

Using gauge metrics in Beam pipelines

Demo

Executing SQL queries on input data

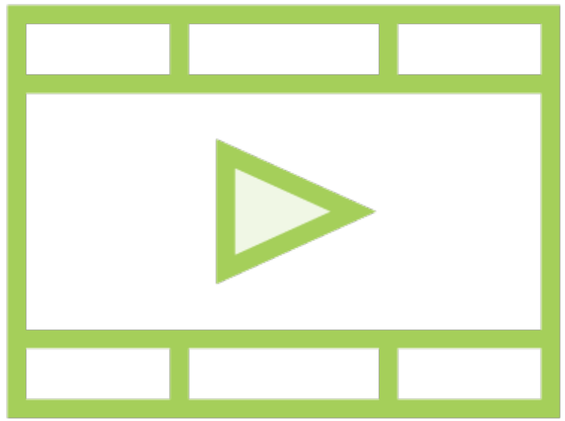
Summary

Tracking pipelines using metrics

Counter, distribution, and gauge metrics

Using SqlTransform to create a transform from a SQL query

Related Courses



Conceptualizing the Processing Model for Apache Spark Structured Streaming

Conceptualizing the Processing Model for the GCP Dataflow Service