

Understanding and Using the Apache Beam APIs



Janani Ravi

CO-FOUNDER, LOONYCORN

www.loonycorn.com

Overview

**Pipelines for data transformation,
runners for execution**

**Build and execute pipelines in Java
and Python**

Visualize and debug pipeline stages

Monitor logs in Dataflow

Write results to BigQuery

Demo

Implementing Apache Beam pipeline in Java

Execute pipeline on the local machine and on the Google Cloud

Demo

Build and execute a Python pipeline to find top-selling products

Monitor logs generated by pipeline executing

Demo

Configure the number of workers on Dataflow

Monitor slow stages using graph visualization

Demo

**Integrate Dataflow with BigQuery to
write results**

Summary

**Pipelines for data transformation,
runners for execution**

**Build and execute pipelines in Java
and Python**

Visualize and debug pipeline stages

Monitor logs in Dataflow

Write results to BigQuery