Data Source

Data Processing

Data Manipulation

Data Analysis

Web/App Server logs

GitHub Repository
(Python scripts, Logs,
Shell Scripts etc.)

Amazon **EC2**

ubuntu

Kafka

Kafka Producer

Kafka
Connect
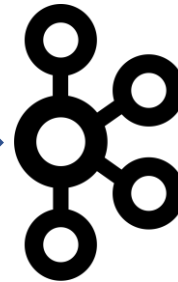
Kafka
Connect

Kafka Consumer

databricks

APACHE **Spark**

Scala

amazon QuickSight

+ableau

Real Time Data Pipeline using Spark and Kafka on AWS EC2

Sushant Burde