

# Uber Real-Time Streaming Platform

---

This project is a full real-time data streaming platform built on Azure to simulate and process Uber-like ride events.

It uses:

- Apache Kafka for real-time ingestion
  - Apache Spark Structured Streaming
  - Azure Databricks for analytics
  - Azure Blob Storage (Data Lake)
  - Terraform for Infrastructure as Code
- 

## Architecture

Producers → Kafka → Spark Streaming → Azure Databricks → Azure Data Lake → Analytics

---

## Cloud Infrastructure

All infrastructure is deployed using Terraform on Microsoft Azure:

- Virtual Network & Subnet
  - Kafka Virtual Machine
  - Azure Databricks Workspace
  - Azure Storage Account (Data Lake)
  - Network & Security configuration
- 

## Project Structure

terraform/ → Infrastructure as Code

kafka/ → Kafka setup & producers

spark/ → Spark streaming & batch jobs

databricks/ → Databricks notebooks

---

## How to Deploy

```
cd terraform  
terraform init  
terraform apply
```