

# CDN Log Analytics on Microsoft Azure



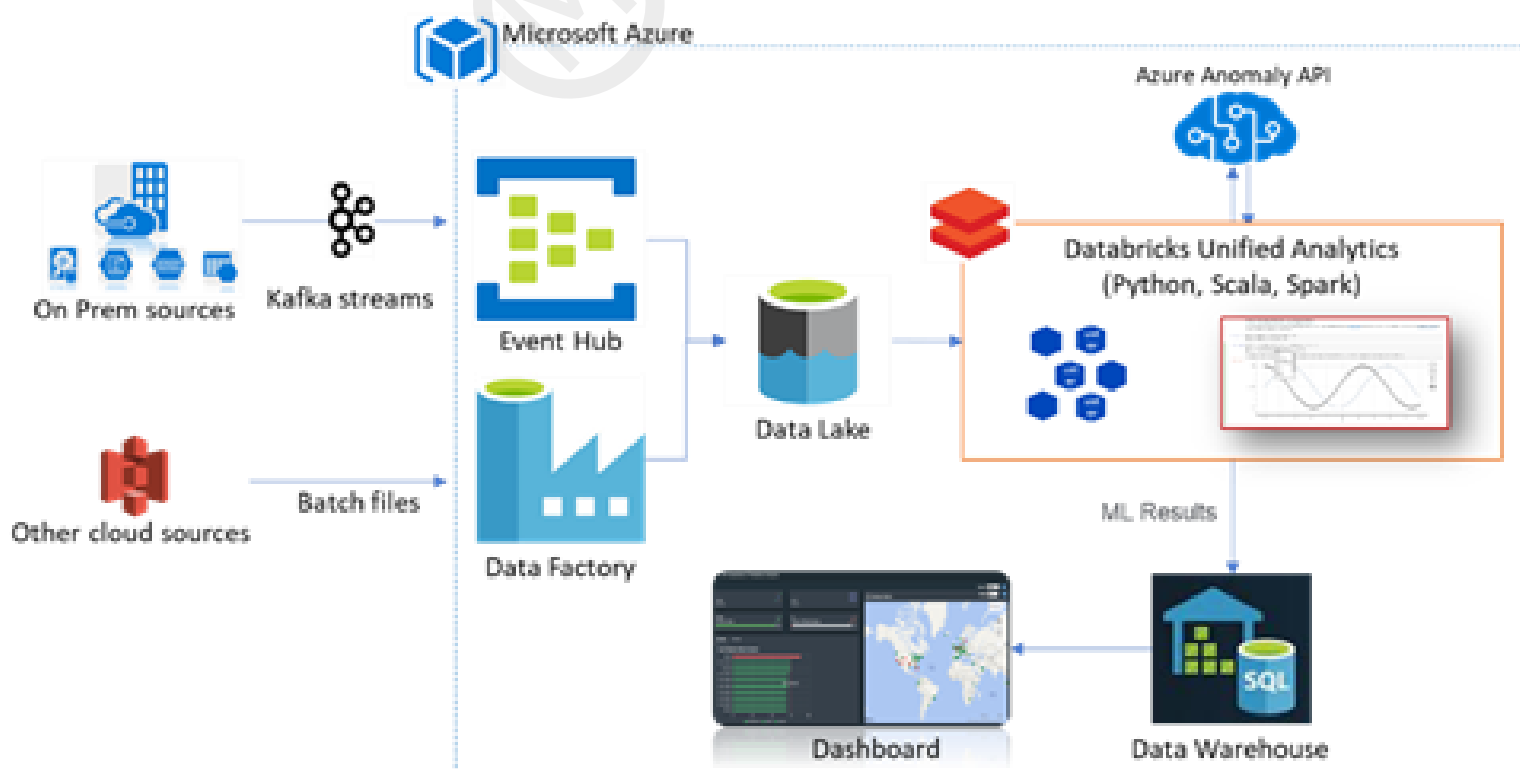
## Business challenge

Content Delivery Network generate logs as they stream videos across the internet to our homes and Mobile devices. These logs which runs into terabytes of volume contain vital information about CDN server performance and video streaming quality. Analyzing this data in real time to understand customer experience and network issues has its own challenges.

## Approach and Solution

The approach was to build a scalable platform using Azure cloud solution to analyze 10TB of data per day.

- The data was ingested through Kafka streams with dedicated Eventhub and moved to Azure Data Lake using Eventhub capture.
- The ingested data was preprocessed using an 800 core Azure Databricks cluster and send to Anomaly Detector API for real time anomaly detection.
- A Machine learning model using LightGBM was built to predict video buffering in end customer devices.
- The processed data and prediction outputs were pushed to Azure Datawarehouse for surfacing on a custom-made UI



## Benefits

- Real Time Network Anomaly detection, resulting in quick action improving video streaming quality
- Ability to analyze historical trends and patterns from 200TB of archived data
- Improved customer experience and business revenues

