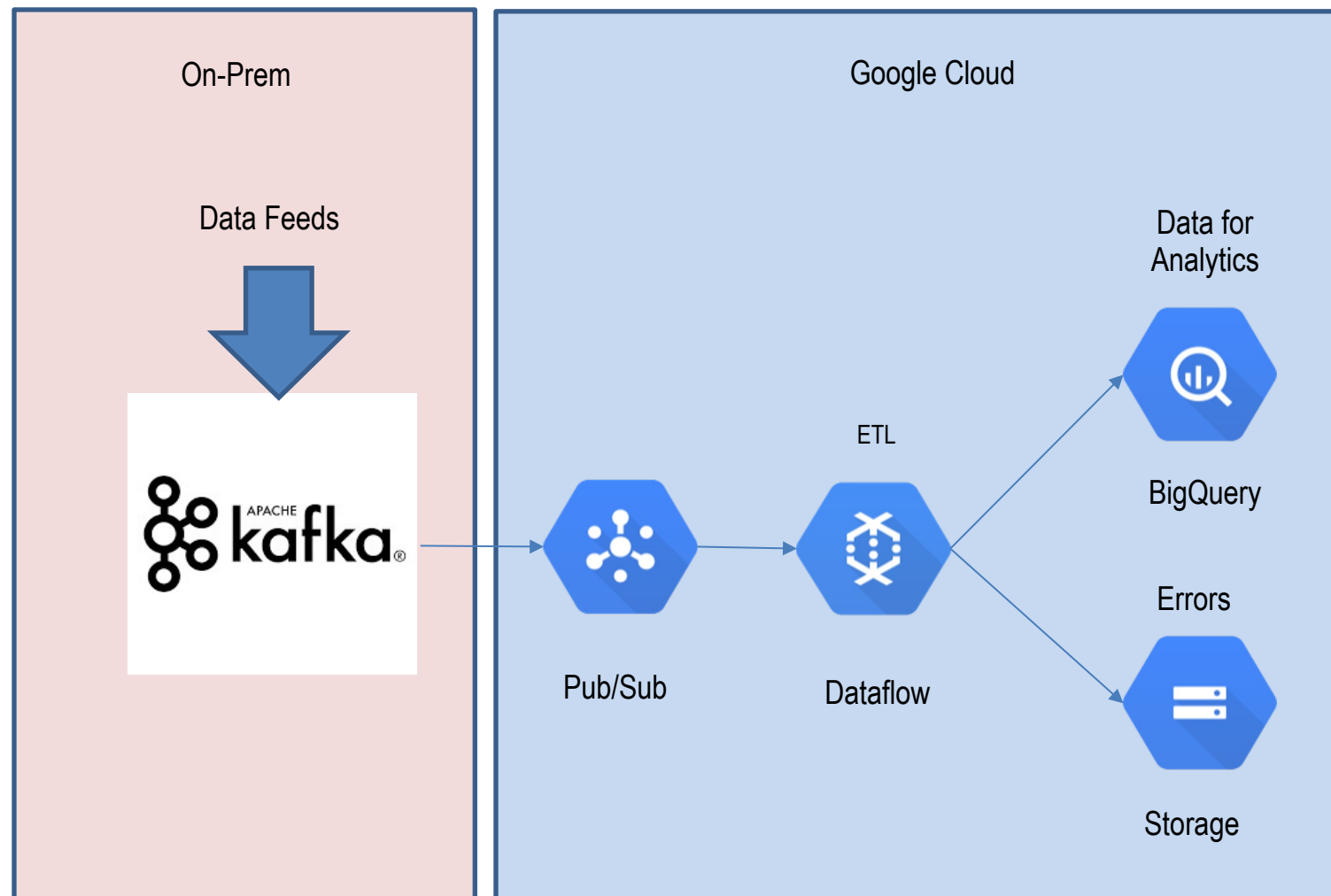


Automobile Dealer – Data Management & Analytics

UK based automobile dealer needed a system for ingestion, transformation and analytics of their streaming data feeds



Real-time streaming data ingestion and analytics

Need

- Creating managed infrastructure on cloud for streaming data.
- Ingestion of thousands of records per second
- Building a serverless analytics system to reduce cost
- Automating data pipeline

Solution

- Built a hybrid system with managed services on the Google cloud
- Automated transfer of data from various sources to analytics system
- Streaming data feeds to Confluent Kafka
- Kafka to cloud data transfer through Confluent Kafka connect
- ETL using Google cloud dataflow
- Serverless analytics using Google BigQuery
- Intermediate storage through Google cloud storage

Results

- Fully managed system reducing infrastructure management
- End to end automation, reducing human intervention to few minutes a day
- Serverless analytics reducing infrastructure cost by 80%
- Auto scaling to cater to seasonal demand for processing
- Automated reconciliation of message counts to ensure no data loss