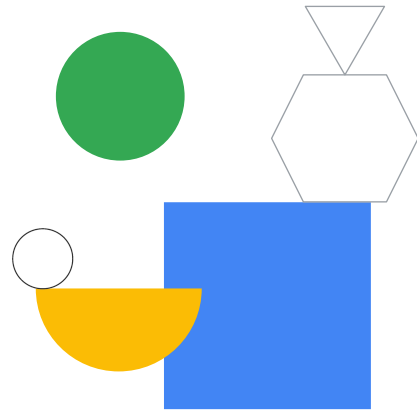


Building Resilient Streaming Analytics Systems on Google Cloud

Summary



Let's review some key concepts we covered on building resilient streaming analytics systems on Google Cloud.

Summary

- Streaming is data processing for unbounded data sets.
- Pub/Sub is a messaging service used as a starting point to process streaming data.
- Dataflow provides a serverless service for processing batch and streaming data.
- BigQuery allows you to stream records into a table.
- Cloud BigTable can handle NoSQL queries over large datasets (petabytes) in milliseconds.

We started off with what streaming data is and the challenges associated with processing streaming data. We said it was important to be able to ingest varying amounts of data because you could have spikes in your data. It is important to be able to deal with unexpected delays because latency is a fact of life.

We want to be able to derive real-time insights from the data even as the data is streaming in. In order to do that, we look at the architecture that consisted of ingesting the data with Pub/Sub, processing the data in stream using Dataflow, and streaming it into BigQuery for durable storage and interactive analysis. We also spent time talking about how Bigtable is a better solution when a much higher throughput is desired.

Summary

- Use BigQuery's analytic window functions for advanced analysis.
- BigQuery has built in GIS functionality.
- Use BigQuery's WITH clause to help modularize your queries.
- There are many ways to optimize your query performance.

And finally, we went back to BigQuery to look at some of its advanced analysis capabilities with window functions and GIS functionalities, as well as reviewed ways to optimize query performance.

Data Engineering on Google Cloud learning path



Modernizing Data Lakes and Data Warehouses with Google Cloud



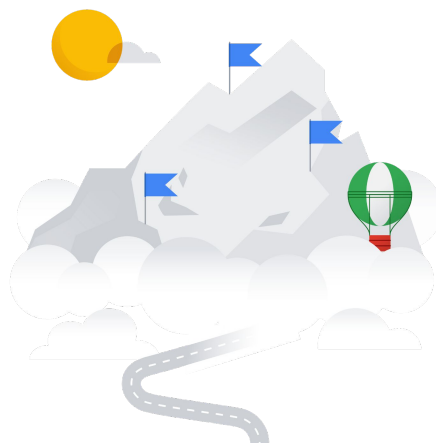
Building Batch Data Pipelines on Google Cloud



Building Resilient Streaming Analytics Systems on Google Cloud



Smart Analytics, Machine Learning and AI on Google Cloud



Congratulations on completing **Building Resilient Streaming Analytics Systems on Google Cloud**.

Smart Analytics, Machine Learning, and AI on Google Cloud is the fourth and final part of the Data Engineering on Google Cloud course, which we will cover on day 4.