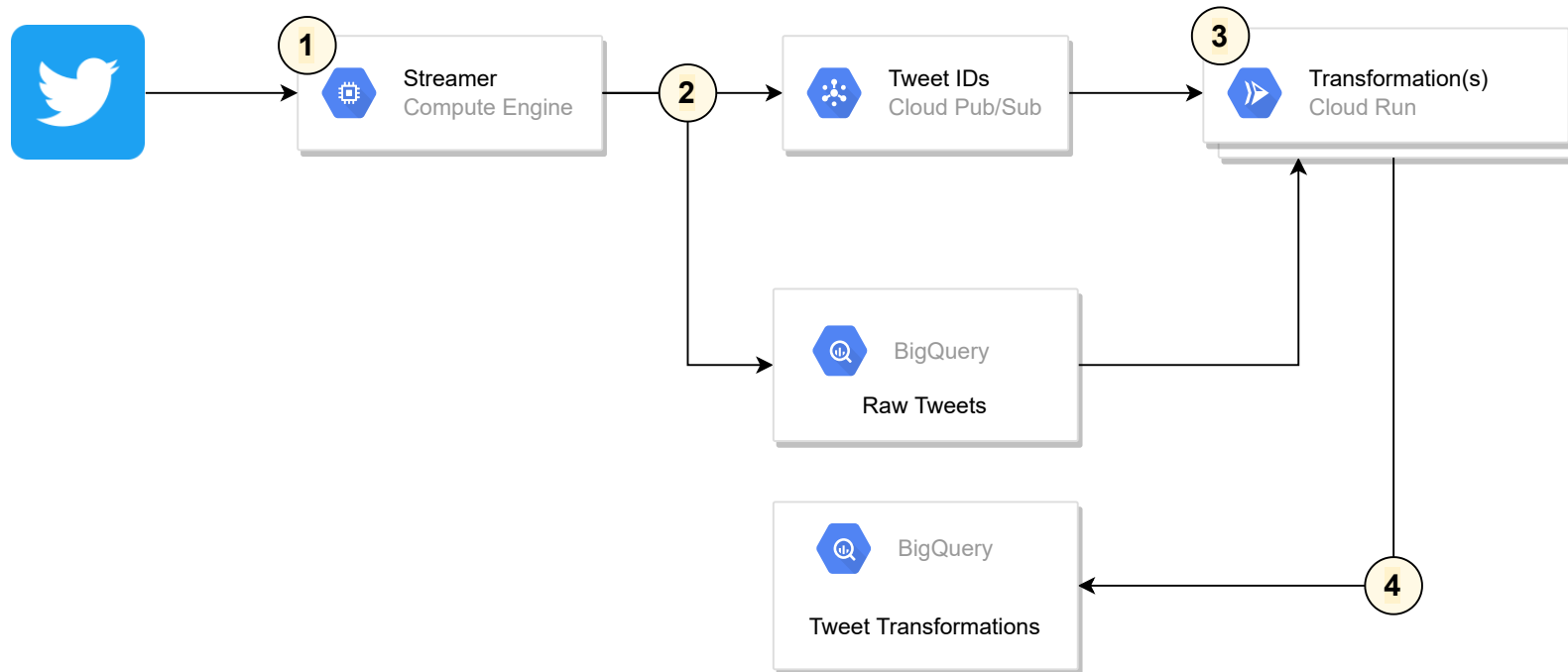


Streamalyzer

Google Cloud Platform - Architecture



Data Flow

1. Streamer continuously collects tweets from Twitter API
2. Streamer stores raw tweet data in BigQuery and publishes a PubSub message with the tweet's ID
3. Transformation services listen for new tweet messages and apply transforms (e.g. add weather info, determine lat/lon, sentiment analysis, etc). They may retrieve more information from the raw tweet data in BigQuery.
4. Transformations are stored in BigQuery in a separate dataset, but with the tweet's ID