

11/15/2017

Norwood Smith: software engineer

BigQueryS: application name

BigQueryS is a Scala application that queries
Google BigQuery for commit message lengths from
The Github repo (bigquery-public-data:github_repos.commits).

The application runs continuously until n queries have been completed
Where n is determined by the command line and if no parameter is present
Will default to 3.

The application is written in Scala, backed by an Akka Actor.
Implemented by an Akka Scheduler, queries are made once every hour
Until n queries are performed.

A running average is kept by way of a list of length, sum pairs.
The average pair is calculated by :

Sum of map keys, sum of (map key * map key value)

The median is calculated from the distribution map. The median position(s) are
Calculated from the mid positions of the sum of the map values and these positions
Are then correlated to the values that would occupy those positions.
A list is kept for each of the running average values and a list for each
Of the running median values.