



File located on s3

File read into file reader

Stream lines into producer

Lines broadcast to consumers

Consumer processes the lines in the EMR cluster

Spark Streaming processes the data stream

Persist the processes to the database

Database is consumed by the view