Elastically Scalable and Fault Tolerant Stream Processing in the Cloud - Video Stream Processing using OpenCV

William Parker[gg18045]

University of Bristol, Bristol, UK gg18045@bristol.ac.uk
https://github.com/ccdb-uob/ccdb-gg18045-scalable-app

Abstract. This paper outlines the architectural design and decisions, implementation and construction of an elastically scalable and fault-tolerant cloud application which performs the parallelisable task of object tracking in streamed video data. The object tracking computation is implemented on serialised video frames preprocessed in a Python Lambda function on upload to S3; after serialisation the frames are streamed to a Kafka topic and the topic is consumed by Spark (pyspark) and processed using OpenCV (cv2).

The cloud is deployed using a Docker image which executes a fabfile (fabric) script; this script builds horizontally auto-scaling EC2 instances on which a Kubernetes cluster is deployed. This cluster is partitioned into four node groups; web frontend, web backend, ZooKeeper service and Kafka producers/consumers/brokers, and a Spark cluster.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

References

1.

A Resources and Links

1. https://github.com/ccdb-uob/stream_event_generator
2. https://github.com/ccdb-uob/ccdb-gg18045-scalable-app