Surge Pricing Ad. Demo Installation

This document guides you through the steps you need to follow to go through the Surge Pricing Ad. Demo. Please follow the accompanying video as well present on the platform.

Kafka

- Run the following 2 commands to run the zookeeper and kafka server. Please remember to go to the kafka directory first which will be present in the downloads directory in your ec2-instance.

```
cd downloads/kafka_2.12-2.3.0
nohup bin/zookeeper-server-start.sh config/zookeeper.properties &
nohup bin/kafka-server-start.sh config/server.properties &
```

- After this go to the directory where you have downloaded the Surge pricing code files and then run the following commands to execute the file_to_kafka.py file

```
chmod 700 file_to_kafka.py
python file_to_kafka.py
```

- You can now check whether the kafka topic has been created correctly or not by running the following command. Remember to go to the kafka binary directory first before running the following command.

```
bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic
surge_pricing_demo --from-beginning
```

Elasticsearch

- For Elasticsearch, you need to run the following prerequisite commands first before starting elasticsearch -

```
sudo su
ulimit -n 65536
sudo sysctl -w vm.max_map_count=262144
su ec2-user
```

- Now run the following command to run elasticsearch

```
cd downloads/elasticsearch-6.4.1/
bin/elasticsearch
```

- Now you need to run the following command to make the mapping in elasticsearch for the data that will come after the Spark Streaming process. Please note that you will have to open the EC2 instance in a new window for this step since elasticsearch will run on the previous window. Also, the IP mentioned below is the elastic IP for your instance. Please replace it in this command accordingly. -

```
curl -X PUT  http://23.20.131.88:9200/rides_demo/  -H 'Content-Type:
application/json'  -d '{
"mappings": {
"surge": {
"properties": {
"geohash": {
"type": "geo_point"
},
"surge": {
"type": "float"
},
"ts": {
"type": "date"
}
}
}
}
}'
```

- You can now run the following command to just check whether the mapping has been created properly or not -

```
curl -X GET http://23.20.131.88:9200/rides_demo/_mapping -H 'Content-Type:
application/json' -H 'Postman-Token:
c94fcf37-2984-418e-be77-3d40fff64cf8,6f3741c3-2b07-4b7e-b0e0-a36be4e4f15b'
-H 'cache-control: no-cache,no-cache'
```

Spark Streaming

- Now, before running the spark streaming job, you need to run the following command. Make sure that you are currently in the directory where you have downloaded the advanced demo code scripts -

```
export SPARK_KAFKA_VERSION=0.10
```

- Now you can run the following command to run the Spark Submit job

```
spark-submit --packages org.apache.spark:spark-sql-kafka-0-10_2.12:3.0.1
spark_test.py
```

- Now you can finally check whether the new Kafka topic has been created by running the Kafka consumer in another window of the EC2 instance.

```
bin/kafka-console-consumer.sh --bootstrap-server localhost:9092 --topic
surge_pricing_demo_output --from-beginning
```

Writing Data to Elasticsearch

- You can now run the es.py script to write the data to Elasticsearch from the topic. Make sure that you have edited the **name of the output kafka topic** and the **Elasticsearch mapping** correctly in this code file -

```
chmod 700 es.py
python es.py
```

- You can now also run the following command to check whether the data is coming to Elasticsearch

```
curl -X GET 'http://23.20.131.88:9200/rides_demo/_search?size=10'
```

Kibana Visualisation

- Now, you need to first start Kibana -

```
cd downloads/kibana-6.4.1-linux-x86_64/
bin/kibana
```

- Open Kibana on your web browser by going to **<public ip of ec2>:5601** and go to **Management > Index Patterns > Create Index Pattern > type your index name > next > select timestamp field > Create Index Pattern**. Also, make sure that you have changed the time range in Kibana by going to **Discover** and then changing the time range to **Last 5 years**
- Once the index pattern is created go to **Visualize > Coordinate Map (Maps) > Select the index pattern you've created > Select the surge column in Metrics and geo point in bucket > Choose heatmap in options tab > go**.