Data Engineering Project

Melis HARMANTEPE

Producer

{"timestamp":"Mon May 29 22:49:38 CEST 2023","droneId":2,"latitude":-156.21837528653572,"longitude":-70.32894993964109,"citizens":
[{"name":"Dennis","surname":"FITZGERALD","score":5,"words":"exhausted,joyous,sad,hello,joyous,depressed"},
{"name":"Hannah","surname":"JONES","score":8,"words":"exhausted,joyous,sad,hello,joyous,depressed"},
{"name":"Jeremy","surname":"DAVIES","score":2,"words":"exhausted,joyous,sad,hello,joyous,depressed"}]}

```
{"timestamp":"Mon May 29 22:48:36 CEST 2023","droneId":4,"latitude":72.54274733426436,"longitude":9.213236255445054,"citizens":[{"name":"Jeremy","
[kafka-producer-network-thread | producer-1] INFO org.apache.kafka.clients.Metadata - [Producer clientId=producer-1] Resetting the last seen epoch
{"timestamp":"Mon May 29 22:49:38 CEST 2023","droneId":2,"latitude":-156.21837528653572,"longitude":-70.32894993964109,"citizens":[{"name":"Dennis
{"timestamp":"Mon May 29 22:50:38 CEST 2023","droneId":1,"latitude":139.85127917806017,"longitude":-4.18917521331656,"citizens":[{"name":"Tom","su
{"timestamp":"Mon May 29 22:51:38 CEST 2023","droneId":4,"latitude":-137.17299364885892,"longitude":15.706220350658072,"citizens":[{"name":"Sophia
{"timestamp":"Mon May 29 22:52:38 CEST 2023","droneId":3,"latitude":31.626873993543484,"longitude":-21.922976817049204,"citizens":[{"name":"Tom","
{"timestamp":"Mon May 29 22:53:38 CEST 2023","droneId":3,"latitude":2.1977572094345135,"longitude":-72.9024942379593,"citizens":[{"name":"Jeremy",
```
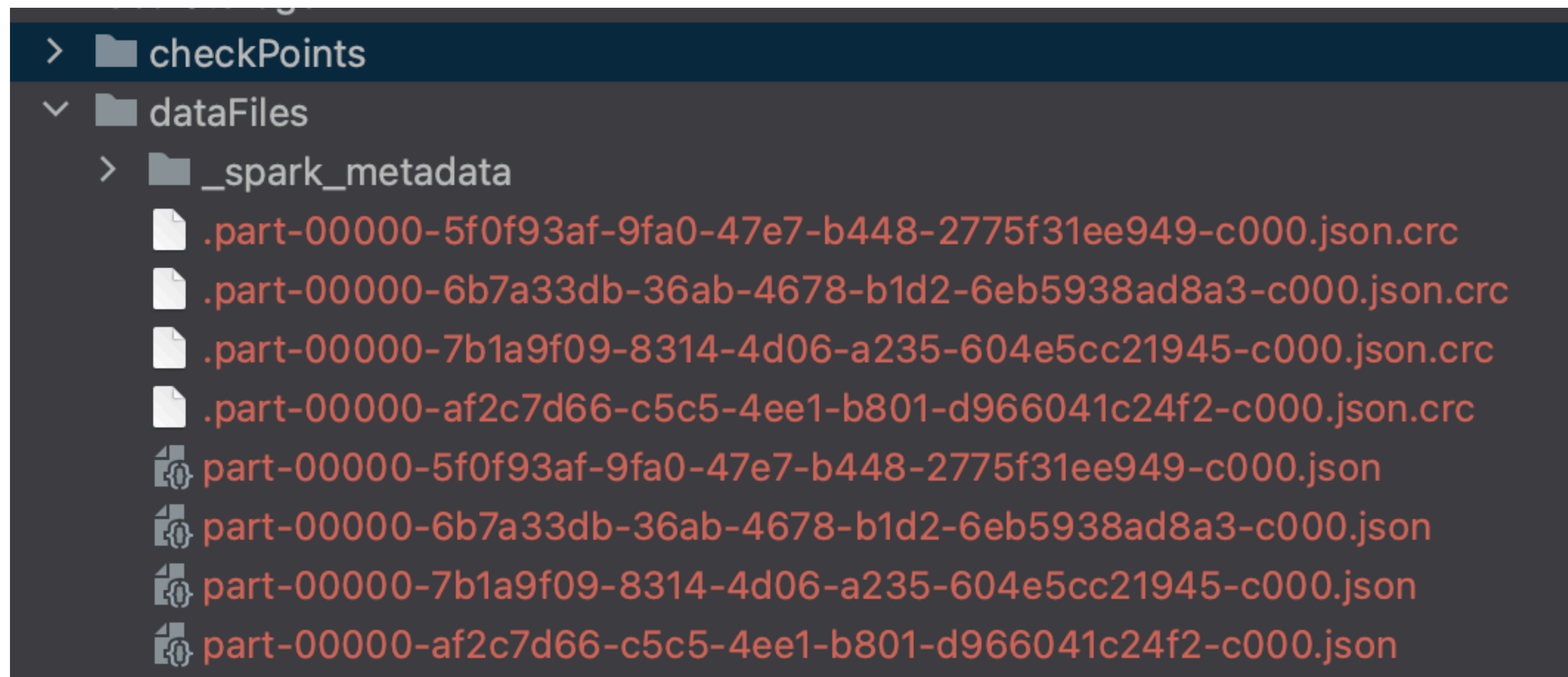
- Kafka Producer

- Ecrit dans le topic "reports"

- Envoie un object json (valeur), droneId en string (clé)

- Sérialize en strings

Alerts Consumer

```
Alert!! Agitated citizens detected with score less than 5 at latitude 72.54274733426436 and longitude 9.213236255445054
Name: Jeremy, Surname: DAVIES, Score: 1
Name: John, Surname: JONES, Score: 3
Information :
Drone ID :  4 Timestamp : Mon May 29 22:48:36 CEST 2023
-------------------
Alert!! Agitated citizens detected with score less than 5 at latitude -156.21837528653572 and longitude -70.32894993964109
Name: Jeremy, Surname: DAVIES, Score: 2
Information :
Drone ID :  2 Timestamp : Mon May 29 22:49:38 CEST 2023
-------------------
```

- Kafka consumer (Group Id : alerts)

- Ecoute sur le topic "reports"

- Déserialise en string

- Retourne les alertes sur la console

Storage

- Consumer (Group Id : storage)

- **Spark** Structured **Streaming** Application (micro-batch)

- Lis les données depuis Kafka Stream, depuis le topic "reports"

- Ecrit les données en **partition** de json sur un dossier en local (stockage distribué)

- Peut s'inscrire sur plusieurs topics et écrire dans plusieurs data storage (scalable)

- Fault tolerance grâce à checkPoints

Analytics

- Spark Structured Streaming

- Lis chaque partition json depuis le répertoire en local

- Retourne un DataFrame

- Appliquer les actions et transformation sur les DF

10 most peaceful citizens

```
+------+----------+-------------+
|  name|   surname|avgPeaceScore|
+------+----------+-------------+
|  John|     BROWN|          8.0|
|Sophia|     LEWIS|          7.0|
|   Tom|     JONES|          6.0|
|Sophia|FITZGERALD|          5.0|
|Jeremy|  WILLIAMS|          5.0|
|Sophia|    WRIGHT|          3.0|
| Clara|FITZGERALD|          2.0|
| Chloe|    WRIGHT|          1.0|
| Chloe|    DAVIES|          0.0|
|Sophia|  WILLIAMS|          0.0|
+------+----------+-------------+
```

Top 10 most heard words

```
+---------+-----+
|    words|count|
+---------+-----+
|exhausted|   10|
|    happy|   10|
+---------+-----+
```

Analytics

The list of citizens with most alerts

```
+------+----------+-----+
| name|   surname|count|
+------+----------+-----+
| Chloe|    DAVIES|    1|
|Sophia|  WILLIAMS|    1|
|Sophia|    WRIGHT|    1|
| Chloe|    WRIGHT|    1|
| Clara|FITZGERALD|    1|
+------+----------+-----+
```

Citizens who have never created a riot

```
+------+----------+-----+
|  name|   surname|score|
+------+----------+-----+
|Sophia|FITZGERALD|    5|
|  John|     BROWN|    8|
|Jeremy|  WILLIAMS|    5|
|   Tom|     JONES|    6|
|Sophia|     LEWIS|    7|
+------+----------+-----+
```