

GUIA

DE

KAFKA + FLINK



Descarregar e extrair Flink

https://downloads.apache.org/flink/flink-1.19.0/flink-1.19.0-bin-scala_2.12.tgz



Iniciar Flink + SQL Client

Dentro da pasta kafka (cd flink-1.19.0)

Para inicializar o server + interface SQL:

```
bin/start-cluster.sh
```

```
bin/sql-client.sh
```



PASTA FLINK – SQL Connectors

Cria uma pasta, dentro da pasta lib, com o nome SQL-Connectors.

Descarrega o connector KAFKA + JSON e coloca-os dentro da pasta SQL-Connectors do Flink.

<https://repo1.maven.org/maven2/org/apache/flink/flink-sql-connector-kafka/3.2.0-1.19/flink-sql-connector-kafka-3.2.0-1.19.jar>

<https://repo1.maven.org/maven2/org/apache/flink/flink-json/1.19.0/flink-json-1.19.0.jar>



CLI Kafka – Produzir mensagens via JSON

Como input, cria um ficheiro JSON com as temperaturas registadas.

Exemplo: `{"id": "sensor_2", "TS": "2025-06-21T10:01:00", "temperatura": 23.0}`

De seguida, utilizando comandos de producer, insere o ficheiro json como input para o tópico criado.

Após as mensagens serem introduzidas no tópico, acede ao SQL Client do FLINK e cria uma tabela temperaturas com as seguintes configurações:

id String; temperatura DOUBLE, TS String;

```
('connector' = 'kafka',
```

```
'topic' = 'temperaturas',
```

```
'properties.bootstrap.servers' = 'localhost:9092',
```

```
'format' = 'json',
```

```
'scan.startup.mode' = 'earliest-offset');
```

Agora a tua tabela temperaturas no FLINK está a ser embutida pelo KAFKA. Podes consultar os dados recorrendo ao *select*.



DESAFIO 1

CENÁRIO: Cria um sistema de sensores de temperatura onde a informação é providenciada via JSON. Executa as seguintes operações:

1. No FLINK, faz uma consulta à tabela e atribui “Temperaturas altas” para temperaturas > 25°C;
2. Filtra a tabela sensores para temperaturas acima dos 20°C e insere a nova informação num novo tópico KAFKA. De seguida, valida as mensagens no KAFKA.



DESAFIO 2

CENÁRIO: Com base na tabela temperaturas, cria uma nova tabela onde armaneza a média das temperaturas numa janela periódica de 10min.



DESAFIO 3

CENÁRIO: Através do SQL Flink, adiciona num tópico novo no KAFKA informações sobre a média das temperaturas assim como a localidade dos sensores.

NOTA:

1. Utiliza o ficheiro localizacao.csv para mapeares as localizações;
2. O conector no flink a utilizar é filesystem.