Kafka Streams

Integration here means integrating with database and with message broker

Reference links

Root url

<https://spring.io/projects/spring-cloud-stream/#overview>

<https://docs.spring.io/spring-cloud/docs/current/reference/html/>

Intro to spring cloud stream

Page by page🡪 <https://docs.spring.io/spring-cloud-stream/reference/spring-cloud-stream.html>

Single page🡪 <https://docs.spring.io/spring-cloud-stream/docs/4.0.4/reference/html/spring-cloud-stream-binder-kafka.html#_reference_guide>

Kafka stream binder

<https://docs.spring.io/spring-cloud-stream/reference/kafka/kafka-streams-binder/usage.html>

Spring documentation links

<https://docs.spring.io/spring-cloud-stream/reference/kafka/kafka-streams.html>

Sample code url’s

<https://github.com/spring-cloud/spring-cloud-stream-samples/>

<https://github.com/confluentinc/kafka-streams-examples>

<https://github.com/spring-cloud/spring-cloud-stream-samples/>

<https://github.com/LearningJournal/Kafka-Streams-with-Spring-Cloud>

Kafka streams is an API or JAR

Its an api for building real time streams processing apps

For this earlier people used apache spark, apache flink

What is apache Kafka

Its an event streaming platform

As it is distributed (it is a cluster of servers) it is scalable, elastic and fault tolerant,

Data will be replicated across the storage nodes(brokers)

Features

1. High level DSL- domain specific language, functional api
2. Low level processor api for specific use cases

These will read a stream of events from inbound topic and do some processing & transformation and write them to outbound topic

Stream-

What is stream: -

Stream is a continuous flow of data and it is unbounded

You don’t need to explicitly request them, you will keep on receiving them

What Is event stream processing

ESP means taking action on series of data that is originating from a system which continuously creates data

Source – is nothing but origin of events – java 8 supplier

Processor – Function- it has both input and output

Sink- data will be saved by this is the end of processing –consumer

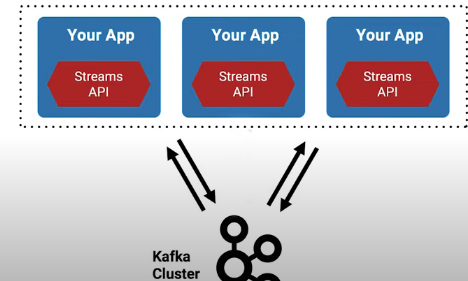
What are Kafka Streams API features?

1. This API is used to process the stream of data where input and output is stored on kafka cluster
2. It enables to consume the stream of data from 1 topic & perform analytical or transformation work and send data to other topic system
3. The Kafka Streams API transforms and enriches data
4. Stream processing is done by application code **not done by any broker**, this is also same like java streams

Ex:- filtering, transforming..

Proc

Jars or starters needed – cloud stream, apache kafka streams



- Supports per-record stream processing with millisecond latency

Here also We can encrypt the data over wire

■ The Kafka Streams API transforms and enriches data

- Supports per-record stream processing with millisecond latency

- Supports stateless processing, stateful processing, windowing operations

■ Write standard Java applications and microservices to process your data in real time

No separate processing cluster required

Elastic, highly scalable, fault-tolerant

Deploy to containers, VMs, bare metal, cloud, on prem

- Equally viable for small, medium, and large use cases

- Fully integrated with Kafka security

- Supports exactly-once semantics as of 0.11.0

■ The Kafka Streams API is part of the open-source Apache Kafka project

Source processor is a component which will read data from kafka topic and converts to a stream

Sink processor will convert that stream and sinks to other topic

Stateless Transformations:- map, mapValues, filter

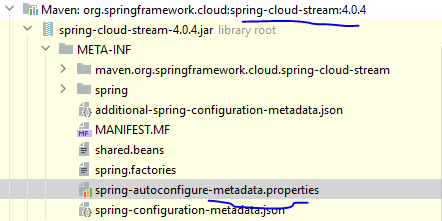
Stateful Transformations:- aggregations(count, reduce), joins, windowing

Built in Abstractions:- Stream builder, KStream, KTable, GlobalKTable

Auto configuration- Spring cloud stream

For this auto configuration see the “spring-cloud-stream”jar

If u want to see which properties are binded to which class then open the below properties file



org.springframework.cloud.stream.config.BindersHealthIndicatorAutoConfiguration.AutoConfigureAfter=org.springframework.cloud.stream.config.BindingServiceConfiguration

@EnableConfigurationProperties({BindingServiceProperties.class, SpringIntegrationProperties.class})

Here if u see all the properties in properties file will binded to this BindingServiceProperties.class

Creating sample project

Add jars like spring cloud stream, **Spring for Apache Kafka Streams, Lombok in start.spring.io**