Introduction

For performance improvement in indexer service reindexing jobs, kafka-connect is getting used to do part of pushing records from kafka-topic to elastic search. The creation of reindexing jobs will be through indexer service only as ealier, but the portion where data is pushed to elastic search would be handled through kafka-connect and not through indexer as it was before. So for reindexing, kafka connect should be run after intiating reindexing job through indexer service. Please follow steps mentioned next to run kafka-connect.

Steps

- Prepare request for kafka connect
  - Replace following fields in the fields in request mentioned at end of this document
    - {{kafka-topic}}   -> whatever index name is defined in legacy index mapping of indexer config with suffix "-encriched". Ex:- index name:- property-services
                    Then kafka topic:- "property-services-enriched"
    - {{elastic search-index}} -> name of elastic search index defined in index config
    - {{connector-uniquename}} -> some unique identifier for connector (would be used for deletion)
- Run legacy index reindexing job in indexer, then go inside playground pod and run command mentioned in next section.
- Wait till all records are indexed
- Then run curl -X DELETE http://kafka-connect.kafka-cluster:8083/connectors/<connector-name> to delete connector

Command

```
curl -X POST \

  http://kafka-connect.kafka-cluster:8083/connectors/ \

H 'Content-Type: application/json' \

H 'Cookie: SESSIONID=f1349448-761e-4ebc-a8bb-f6799e756185' \

H 'Postman-Token: adabf0e8-0599-4ac9-a591-920586ff4d50' \

H 'cache-control: no-cache' \

d '{

  "name": "{{connector-uniquename}}",

  "config": {

    "connector.class": "io.confluent.connect.elasticsearch.ElasticsearchSinkConnector",

    "connection.url": "http://elasticsearch-data-v1.es-cluster:9200",
```

```
    "type.name": "general",

    "topics": "{{kafka-topic}}",

    "key.ignore": "false",

    "schema.ignore": true,

    "value.converter.schemas.enable": false,

    "key.converter": "org.apache.kafka.connect.storage.StringConverter",

    "value.converter": "org.apache.kafka.connect.json.JsonConverter",

    "transforms": "TopicNameRouter",

    "transforms.TopicNameRouter.type": "org.apache.kafka.connect.transforms.RegexRouter",

    "transforms.TopicNameRouter.regex": ".*",

    "transforms.TopicNameRouter.replacement": "{{elastic search-index}}",

    "batch.size": 10,

    "max.buffered.records": 500,

    "flush.timeout.ms": 600000,

    "retry.backoff.ms": 5000,

    "read.timout.ms": 10000,

    "linger.ms": 100,

    "max.in.flight.requests": 2,

    "errors.log.enable": true,

    "errors.deadletterqueue.topic.name": "{{kafka-topic}}-es-failed",

    "tasks.max": 1
  }
}'
```

Ex:-

```
curl -X POST
http://kafka-connect.kafka-cluster:8083/connectors/
-H 'Content-Type: application/json'
-H 'Cookie: SESSIONID=f1349448-761e-4ebc-a8bb-f6799e756185'
-H 'Postman-Token: adabf0e8-0599-4ac9-a591-920586ff4d50'
-H 'cache-control: no-cache'
-d '{

"name": "cms-case-es-sink28",

"config": {

"connector.class": "io.confluent.connect.elasticsearch.ElasticsearchSinkConnector",

"connection.url": "<http://elasticsearch-data-v1.es-cluster:9200",>

"type.name": "general",

"topics": "property-services-enriched",

"key.ignore": "false",

"schema.ignore": true,

"value.converter.schemas.enable": false,

"key.converter": "org.apache.kafka.connect.storage.StringConverter",

"value.converter": "org.apache.kafka.connect.json.JsonConverter",

"transforms": "TopicNameRouter",

"transforms.TopicNameRouter.type": "org.apache.kafka.connect.transforms.RegexRouter",

"transforms.TopicNameRouter.regex": ".*",

"transforms.TopicNameRouter.replacement": "property-services",

"batch.size": 10,

"max.buffered.records": 500,

"flush.timeout.ms": 600000,

"retry.backoff.ms": 5000,
```

"read.timout.ms": 10000,

"linger.ms": 100,

"max.in.flight.requests": 2,

"errors.log.enable": true,

"errors.deadletterqueue.topic.name": "property-services-es-failed",

"tasks.max": 1

}

}'