Kafka Setup

POC VM Details:

ssh userid: azureuser

pem_key : wmqs-poc-avm-sea-kafka-prvkey.pem

wmqs-poc-avm-sea-kafka: 10.152.202.158

Installation steps:

We can run kafka as docker containers or in Distributed environment but due to permission issue currently running in standalone environment.

Standalone Approach:

1) Download Confluent packages

curl -O http://packages.confluent.io/archive/5.5/confluent-5.5.2-2.12.tar.gz

untar: tar -xf confluent-5.5.2-2.12.tar.gz

cd confluent-5.5.2/share

mkdir confluent-hub-components

2) Download ibm jars [confluentinc-kafka-connect-ibmmq-11.0.2] and copy to **/data1/confluent-5.5.2/share/confluent-hub-components**

Goto the link to download **https://www.confluent.io/hub/confluentinc/kafka-connect-ibmmq**

3) Download all client ibm jar: **com.ibm.mq.allclient-9.1.0.1.jar** and copy to **/data1/confluent-5.5.2/share/confluent-hub-components /confluentinc-kafka-connect-ibmmq-11.0.1/lib/**

Artifactory location : ext-release-local/com/ibm/mq.allclient/9.1.0.1/mq.allclient-9.1.0.1.jar

http://artifactory.asia.essilor.group/artifactory/webapp/#/artifacts/browse/tree/General/ext-release-local/com/ibm/mq.allclient/9.1.0.1/mq.allclient-9.1.0.1.jar

4) Create connector1.properties file : **/data1/confluent-5.5.2/connector1.properties**

File: connector1.properties

5) Update /data1/confluent-5.5.2/etc/kafka/connect-standalone.properties file configuration

Updated File: connect-standalone.properties

6) Finally start the service [Added init scripts]

systemctl start mq-kafka [start/stop/restart/status] or bin/connect-standalone etc/kafka/connect-standalone.properties connector1.properties

Postgres worker node setup:

Added new postgres cdc connector jars in default plugin path:

confluent-5.5.2/share/confluent-hub-components/debezium-debezium-connector-postgresql-1.4.1

Added new properties files:

1) /data1/confluent-5.5.2/etc/kafka/postgres.connect-standalone.properties

2) /data1/confluent-5.5.2/postgres.connector1.properties

Startup commands:

systemctl start postgres-kafka [start/stop/restart/status] or bin/connect-standalone etc/kafka/postgres.connect-standalone.properties postgres.connector1.properties

Logs:

/data1/confluent-5.5.2/logs/

-rw-r--r--. 1 root root 3008 May 11 05:13 postgres-kafka_connect.err
-rw-r--r--. 1 root root 136974 May 11 05:14 postgres-kafka_connect.out

Prometheus connector setup:

Download prometheus connector plugin jars "confluentinc-kafka-connect-prometheus-metrics-1.1.6-preview"

https://www.confluent.io/hub/confluentinc/kafka-connect-prometheus-metrics

Add to location /data1/confluent-5.5.2/share/confluent-hub-components/confluentinc-kafka-connect-prometheus-metrics-1.1.6-preview

Add prometheus.connect-standalone.properties to "/data1/confluent-5.5.2/etc/kafka"

Add prometheus-connector2.properties to "/data1/confluent-5.5.2"

Startup commands:

systemctl status prometheus-kafka [start/stop/status] (or) bin/connect-standalone etc/kafka/prometheus.connect-standalone.properties prometheus-connector2.properties

/data1/confluent-5.5.2/logs/prometheus_kafka_connect.out

Once connector is up and running go to Confluent cloud , go to cluster select topics

click on KsqlDB – > Run below queries to create streams

KsqlDB Queries:

```
create stream machine_flag with (kafka_topic='machine-flag', value_format='avro');

set 'auto.offset.reset'='earliest';


create table machine_flag_agg
with (kafka_topic='machine_flag_agg', key_format='avro', value_format='avro', partitions=1)
as
select
lab,
machine,
flag,
substring(timestamptostring(timestamp, 'yyyy-MM-dd HH:mm:ss'), 12, 2) as hour,
count(machine) as doubleValue
from machine_flag
window tumbling (size 5 seconds)
group by lab, machine, flag, substring(timestamptostring(timestamp, 'yyyy-MM-dd HH:mm:ss'), 12, 2)
emit changes;


create stream machine_flag_agg_prometheus with (kafka_topic='machine_flag_agg', key_format='avro', value_format='avro');

create stream prometheus_sink with (kafka_topic='prometheus_sink', value_format='avro', partitions=1)
as
select rowkey->machine as machine, 'machine_flag' as `name`, 'gauge' as `type`, rowtime as `timestamp`, struct(`doubleValue` := doublevalue)
as `values`, struct(`lab` := rowkey->lab, `machine` := rowkey->machine, `flag` :=rowkey->flag, `hour` :=rowkey->hour, `method` := 'update') as
`dimensions` from machine_flag_agg_prometheus partition by rowkey->machine emit changes;


create stream prometheus_sink_counter_1 with (kafka_topic='prometheus_sink_counter_1', value_format='avro', partitions=1)
as
select rowkey->machine as machine, 'machine_flag_counter' as `name`, 'counter' as `type`, rowtime as `timestamp`, struct(`doubleValue` :=
doublevalue) as `values`, struct(`lab` := rowkey->lab, `machine` := rowkey->machine, `flag` :=rowkey->flag, `hour` :=rowkey->hour, `method` :=
'update') as `dimensions` from machine_flag_agg_prometheus partition by rowkey->machine emit changes;
```

Login to Confluent Cloud:

1) Login to Confluent cloud and check events in the topics:

https://confluent.cloud/environments/env-1drxz/clusters

2) Issue with Login:

Contact: mariacarmel.caballa@essilor.com ,

vishnu.n@essilor.com.sg

DEBUG:

To Check logs : **cd /data1/confluent-5.5.2/logs**

To Enable Debug mode:

"curl -s -X PUT -H "Content-Type:application/json" \ http://localhost:8083/admin/loggers/org.apache.kafka.connect.runtime.WorkerSourceTask \ -d '{"level": "TRACE"}' | jq '."