“REAL-TIME DATA STREAMING FROM RDBMS TO NON-RDBMS”

By

Md. Ashiqul Islam Shajal (ASC)

- Required tools and technology:
  - ✓ Apache Kafka.
  - ✓ Zookeeper.
  - ✓ Debezium Connector (Source).
  - ✓ Oracle Database Archive Log mode enable.
  - ✓ Ojdbc-8
  - ✓ Java/others language for consumer API dev.
  - ✓ ELK Stack.

- Configuration:
  - ✓ **Zookeeper Cluster Setup: (with 3 zK nodes):**

    Download Apache Zookeeper 3.8 tar file and create 3 different folder. Copy and paste this tar file into three different folders and extract tar file.
    Ex- folder-1: zkNode-1(server-1), folder-2: zkNode-2(server-2), folder-3: zkNode-3(server-3).

    Go to in path (../zkNode-1/apache-zookeeper-3.8.0-bin/conf) and create a configuration file named zoo.cfg and Do some configuration for all nodes. The configuration Code is given below –

    - ➢ For zkNode-1: (zk server-1)

```
tickTime=2000
initLimit=10
syncLimit=5
dataDir=/tmp/zookeeper-node-1
clientPort=2181
maxClientCnxns=60

4lw.commands.whitelist=*

server.1=10.11.200.109:2788:3788
server.2=10.11.200.109:2888:3888
server.3=10.11.200.109:2988:3988
```

- For zkNode-2: (zk server-2)

```
tickTime=2000
initLimit=10
syncLimit=5
dataDir=/tmp/zookeeper-node-2
clientPort=2182
maxClientCnxns=60

4lw.commands.whitelist=*

server.1=10.11.200.109:2788:3788
server.2=10.11.200.109:2888:3888
server.3=10.11.200.109:2988:3988
```

- For zkNode-3: (zk server-3)

```
tickTime=2000
initLimit=10
syncLimit=5
dataDir=/tmp/zookeeper-node-3
clientPort=2183
maxClientCnxns=60

4lw.commands.whitelist=*

server.1=10.11.200.109:2788:3788
server.2=10.11.200.109:2888:3888
server.3=10.11.200.109:2988:3988
```

✓ **Zookeeper Cluster Setup: (with 3 zK nodes):** Create temporary log data file and define those file using specific id. After shut down these file will be deleted so if you want, you can store this log data in different location.

```
mkdir /tmp/zookeeper-node-1
mkdir /tmp/zookeeper-node-2
mkdir /tmp/zookeeper-node-3
```

```
echo 1 >> /tmp/zookeeper-node-1/myid
echo 2 >> /tmp/zookeeper-node-2/myid
echo 3 >> /tmp/zookeeper-node-3/myid

cat /tmp/zookeeper-node-1/myid
cat /tmp/zookeeper-node-2/myid
cat /tmp/zookeeper-node-3/myid
```

- ✓ **Start Zookeeper Clusters:**
  - Start in foreground: bin/zkServer.sh start-foreground
  - Start in Background: bin/zkServer.sh start
  - Stop in Background: bin/zkServer.sh stop
- ✓ **Check Zookeeper Cluster Status:**
  - fardaus@era-ai:~$ echo stat | nc 10.11.200.109 2181

    Zookeeper version: 3.8.0-5a02a05eddb59aee6ac762f7ea82e92a68eb9c0f, built on 2022-02-25 08:49 UTC
    Clients:
    /10.11.200.109:54454[0](queued=0,recved=1,sent=0)

    Latency min/avg/max: 0/0.0/0
    Received: 1
    Sent: 0
    Connections: 1
    Outstanding: 0
    Zxid: 0x0
    Mode: follower
    Node count: 5

  - fardaus@era-ai:~$ echo stat | nc 10.11.200.109 2182

    Zookeeper version: 3.8.0-5a02a05eddb59aee6ac762f7ea82e92a68eb9c0f, built on 2022-02-25 08:49 UTC
    Clients:
    /10.11.200.109:43658[0](queued=0,recved=1,sent=0)

Latency min/avg/max: 0/0.0/0
Received: 1
Sent: 0
Connections: 1
Outstanding: 0
Zxid: 0x100000000
Mode: leader
Node count: 5
Proposal sizes last/min/max: -1/-1/-1

- fardaus@era-ai:~$ echo stat | nc 10.11.200.109 2183

  Zookeeper version: 3.8.0-5a02a05eddb59aee6ac762f7ea82e92a68eb9c0f, built on 2022-02-25 08:49 UTC
  Clients:
  /10.11.200.109:53444[0](queued=0,recved=1,sent=0)
  Latency min/avg/max: 0/0.0/0
  Received: 1
  Sent: 0
  Connections: 1
  Outstanding: 0
  Zxid: 0x100000000
  Mode: follower
  Node count: 5

✓ **Kafka Cluster Setup: (with 3 Kafka nodes):**

- System Prerequisites:
  1) Java 8 or Higher.
  2) RAM size minimum 1GB.

Download kafka_2.13-3.3.2.tgz file and create 3 different folder. Copy and paste this tar file into three different folders and extract tar file.

Ex- folder-1: kafkaNode-1(server-1), folder-2: kafkaNode-2(server-2), folder-3: kafkaNode-3(server-3).

Go to in path (../kafkaNode-1/kafka_2.13-3.3.1/config/) and edit a configuration file named server.properties and Do some configuration for all nodes. The configuration Code is given below –

- For kafkaNode-1: (kafka server-1)

```
############ Server Basics #################
broker.id=1
############ Socket Server Settings #########
listeners=PLAINTEXT://10.11.200.109:9092
############ Log Basics ####################
log.dirs=/tmp/kafka-logs-node-1
############ Zookeeper #####################
zookeeper.connect=10.11.200.109:2181,10.11.200.109:2182,10.11.200.109:218
3
```

- For kafkaNode-2: (kafka server-2)

```
############ Server Basics #################
broker.id=2
############ Socket Server Settings #########
listeners=PLAINTEXT://10.11.200.109:9093
############ Log Basics ####################
log.dirs=/tmp/kafka-logs-node-2
############ Zookeeper #####################
zookeeper.connect=10.11.200.109:2181,10.11.200.109:2182,10.11.200.109:218
3
```

- For kafkaNode-3: (kafka server-3)

```
############ Server Basics #################
broker.id=3
############ Socket Server Settings #########
listeners=PLAINTEXT://10.11.200.109:9094
############ Log Basics ####################
log.dirs=/tmp/kafka-logs-node-3
############ Zookeeper #####################
zookeeper.connect=10.11.200.109:2181,10.11.200.109:2182,10.11.200.109:218
3
```

- ✓ **Start Kafka Clusters:**
  - Start in foreground: bin/kafka-server-start.sh config/server.properties
  - Stop in foreground: bin/kafka-server-stop.sh config/server.properties
  - Start in Background: bin/kafka-server-start.sh -daemon config/server.properties
  - Stop in Background: bin/kafka-server-stop.sh -daemon config/server.properties

- ✓ **Check Status of Kafka Brokers:**
  - fardaus@era-ai:~$ echo dump | nc localhost 2181

SessionTracker dump:

Global Sessions(3):

0x100008709140000 18000ms

0x100008709140001 18000ms

0x30000875a0f0000 18000ms

ephemeral nodes dump:

Sessions with Ephemerals (3):

0x100008709140001:

/brokers/ids/3

0x100008709140000:

/brokers/ids/2

0x30000875a0f0000:

/controller

/brokers/ids/1

Connections dump:

Connections Sets (2)/(3):

0 expire at Tue Apr 11 12:32:48 BDT 2023:

3 expire at Tue Apr 11 12:32:58 BDT 2023:

ip: /127.0.0.1:47724 sessionId: 0x0

ip: /10.11.200.109:56062 sessionId: 0x100008709140001

ip: /10.11.200.109:51350 sessionId: 0x100008709140000

✓ Enable archive log mode in Oracle Database:

- Step-1: Enter in SYS user and check archive is enable or not.

```
SQL> SELECT  LOG_MODE  FROM  v$database;
     LOG_MODE
     ------------------
     NOARCHIVELOG
```

- Step-2: Shutdown Database.

```
SQL> shutdown immediate;
```

- Step-3: After shutdown connect to sys user.

```
SQL> conn sys/sys123 as sysdba;
```

- Step-4: Start database on mount stage.

```
SQL> startup mount;
```

- Step-5: Alter database on archive log mode.

```
SQL> alter database archivelog;
```

- Step-6: Alter database on archive log mode and check status.

```
SQL> alter database archivelog;
SQL> select log_mode from v$database;
     LOG_MODE
     ------------------
     ARCHIVELOG
SQL> archive log list;
```

- Step-7: Now open database.

```
SQL> alter database open;
```

✓ Debezium source connector setup (Oracle CDC):

- Download debezium source connector:
  Link: https://debezium.io/releases/1.9/

- Unzip the debezium file and paste in location path:
  Path: /kafkaNode-1/kafka_2.13-3.3.1/plugins

- Download jdbc driver:
  Type: ojdbc8
  Version: ojdbc8-21.1.0.0.jar (based on installed database version i.e – oracle 21c)

- Copy and Paste this ojdbc8-21.1.0.0.jar file in location path:
  Path: /kafkaNode-1/kafka_2.13-3.3.1/plugins/debezium-connector-oracle

✓ Connector starter file setup:

- Go to location path "/kafkaNode-1/kafka_2.13-3.3.1/config" and edit connector starter file named "connect-standalone.properties"

- Configuration –
  bootstrap.servers=10.11.200.109:9092,10.11.200.109:9093,10.11.200.109:9094
  key.converter=org.apache.kafka.connect.json.JsonConverter
  value.converter=org.apache.kafka.connect.json.JsonConverter
  key.converter.schemas.enable=true
  value.converter.schemas.enable=true
  plugin.path=/home/fardaus/elk_abs/kafkaEnv/kafkaNode-1/kafka_2.13-3.3.1/plugins/debezium-connector-oracle (provide your path based on your configuration)

- ✓ **Connection configuration file setup:** Create a file on path "/kafkaNode-1/kafka_2.13-3.3.1/config" named "ora-connector.properties"

ora-connector.properties

```
name=oracle-source-connector
connector.class=io.debezium.connector.oracle.OracleConnector
tasks.max=1
topic.prefix=oracle_topic
database.hostname=10.11.200.117
database.port=1521
database.dbname=XE
database.user=sys as sysdba
database.password=sys123
database.pdb.name =XEPDB1
database.server.name=DESKTOP-P7C7D44
database.history=io.debezium.relational.history.FileDatabaseHistory
database.history.file.filename=/home/fardaus/elk_abs/kafkaEnv/kafkaNode-1/kafka_2.13-3.3.1/config/history/file
include.schema.changes=false
table.whitelist=elk_user.employees
table.include.list =elk_user.employees
schema.history.internal.kafka.topic=debezium_schema_history
bootstrap.servers=10.11.200.109:9092
schema.history.internal.kafka.bootstrap.servers=10.11.200.109:9092,10.11.200.109:9093,10.11.200.109:9094
```

- ✓ **Start source connector:**
  - bin/connect-standalone.sh config/connect-standalone.properties config/ora-connector.properties