**Software Versions Used**

**Oracle Database Version              :**               **11.2.0 XE (Express Edition)**

**Oracle Golden Gate Version       :               Oracle® GoldenGate 11g Release 2 (11.2.1.0.3**

**Oracle Golden Gate for Big Data               :               12.3.2.1.0**

**Configuring Golden Gate**

1. **Set path variable**
   1. set path=%path%;C:\Program Files\Java\jdk1.8.0\_141\jre\bin\server
2. **Create a User for Golden Gate**
3. Open SQL Command Line
4. Type “conn system/tiger” and press Enter.
5. CREATE TABLESPACE GGS\_OWNER DATAFILE 'diskb:tbs\_f5.dat' SIZE 500K REUSE AUTOEXTEND ON NEXT 500K MAXSIZE 100M; (creating a new tablespace ggs\_owner )
6. **create user ggs\_owner identified by ggs\_owner default tablespace ggs\_owner Temporary tablespace temp**; (creating a user ggs\_owner )
7. grant dba to ggs\_owner; (Giving dba privileges to the user)
8. Open to ggsci and Check connection using the below command
9. DBLOGIN USERID ggs\_owner , PASSWORD ggs\_owner
10. In ggsci execute the command “create SUBDIRS“
11. In ggsci execute the command “info all“ (which says all the process states…in which manager is stopped)
12. In ggsci execute the command  EDIT PARAMS MGR
13. Open mgr.prm file and add

PORT 7809

AUTOSTART ER \*

AUTOSTART EXTRACT \*

PURGEOLDEXTRACTS D:\JAVA\_STUDY\LIVESTREAMING\GG-11.2.1.0.3\ggtrails\\*, USECHECKPOINTS , MINKEEPDAYS 3

1. To start the manager “start mgr”
2. In ggsci execute the command “info all“ (It will show that the manager is running)
3. If there is any error, you can find it in ggserr.log file.

1. **Enable Archive Mode**

Conn sys as sysdba

<http://www.oracledistilled.com/oracle-database/backup-and-recovery/enabledisable-archive-log-mode-10g11g/>  follow this link to enable archive mode.

1. **Create a user/schema (Golden Gate will be replicating this schema).**
2. Conn sys as sysdba
3. CREATE USER scott IDENTIFIED BY tiger;
4. GRANT CREATE SESSION GRANT ANY PRIVILEGE TO scott;
5. GRANT CONNECT, RESOURCE, DBA TO scott;
6. Navigate to SQLPlus and connect using

conn scott/tiger

CREATE TABLE customers ( customer\_id number(10) NOT NULL, customer\_name varchar2(50) NOT NULL, city varchar2(50));

GRANT   SELECT, INSERT, UPDATE, DELETE ON   SCOTT.CUSTOMERS TO   scott;

Select \* from tab; (To view the tables accessible by the user)

1. **To Enable Supplemental Logging.**
   * 1. Conn sys as sysdba
     2. alter database add supplemental log data (primary key) columns;
     3. ALTER SYSTEM SWITCH LOGFILE;
     4. SELECT SUPPLEMENTAL\_LOG\_DATA\_MIN, SUPPLEMENTAL\_LOG\_DATA\_PK , SUPPLEMENTAL\_LOG\_DATA\_UI  from v$database; // To check whether the supplemental log is enabled
     5. Open ggsci
   1. Log into Database   “DBLOGIN USERID ggs\_owner, PASSWORD ggs\_owner”
   2. ADD TRANDATA SCOTT.CUSTOMERS
2. **To Add Extract**

In ggsci execute the command “EDIT PARAMS tmobile”. This will open a new file tmobile.prm.

Add the highlighted content in the file.

EXTRACT tmobile

USERID ggs\_owner, PASSWORD ggs\_owner

DISCARDFILE D:\JAVA\_STUDY\LIVESTREAMING\GG-11.2.1.0.3\ggtrails\SCOTT\sample\_srce.dsc, purge

EXTTRAIL D:\JAVA\_STUDY\LIVESTREAMING\GG-11.2.1.0.3\dirdat\1t

TRANLOGOPTIONS EXCLUDEUSER ggs\_owner

TABLE SCOTT.CUSTOMERS;

Once saved the above file Open ggsci

DBLOGIN USERID ggs\_owner , PASSWORD ggs\_owner

ADD EXTRACT tmobile, TRANLOG, BEGIN NOW

ADD EXTTRAIL D:\Softwares\GGS11g\V34342-01\dirdat\1t, EXTRACT tmobile

START EXTRACT tmobile

\*\*Golden Gate logs can be found here **D:\Softwares\GGS11g\V34342-01\dirrpt or**ggserr.log

**\*\*C:\oraclexe\app\oracle\product\11.2.0\server\network\ADMIN\tnsnames.ora**

1. **How to check whether the extract is working or not.**

Click logdump.exe in the Oracle GG home directory.

GHDR ON

DETAIL ON

DETAIL DATA

OPEN /u01/app/ogg/dirdat/lt000000000 (Adjust this location so that it points to your trial file location)

FILTER INCLUDE FILENAME SCOTT.CUSTOMERS

N(next)

Now you could see your updated records here.

1. **Configuring Replicat for Kafka Handler**

Install Golden Gage for BigData.

Open ggsci

GGSCI> CREATE SUBDIRS

Create a Manager parameter file:  
GGSCI> EDIT PARAM MGR

Add the below content to the mgr.prm file

PORT 7801

GGSCI>START MGR

GGSCI>INFO MGR

1. **How to configure Kafka**
2. **How to configure Zookeeper**

Copy and Rename “zoo\_sample.cfg” to “zoo.cfg” in C:\Tools\zookeeper-3.4.9\conf

Find & edit dataDir=/tmp/zookeeper to :\zookeeper-3.4.9\

Add in System Variables ZOOKEEPER\_HOME = C:\Tools\zookeeper-3.4.9

Edit System Variable named “Path” and append this in the last ;%ZOOKEEPER\_HOME%\bin;

Open command prompt and type zkserver.

1. **Configuring Kafka Adapter**

copy the configuration files from $OGG\_BD\_HOME/AdapterExamples/big-data/kafka directory to dirprm directory (1.custom\_kafka\_producer.properties  
2.kafka.props  
3.rkafka.prm)

Create a folder “lib” and copy all the jar files from KAFKA\_HOME/libs to this lib folder.

Open ggsci and execute the command.

add replicat rkafka, exttrail D:\JAVA\_STUDY\LIVESTREAMING\GG-11.2.1.0.3\dirdat\1t

Start Zookeeper

Start Kafka

(Open command prompt and go to your Apache Kafka directory and run following command.

**.\bin\windows\kafka-server-start.bat .\config\server.properties**

Data definitions are needed when the source and target tables have different definitions or the databases are of different types.

**Steps to generate DEFGEN File. This should be done in GG (not GG BD)**

1. In GGSCI, issue the following command to create a DEFGEN parameter file.

ggsci > Edit params defgen

defsfile ./dirsql/SOURCE\_DEFGEN.def

userid ggs\_owner password ggs\_owner

TABLE scott.customers;

1. Enter the parameters listed above in the order shown and invoke DEFGEN from GG\_HOME location

$ defgen paramfile ./dirprm/defgen.prm

1. Copy the generated SOURCE\_DEFGEN.def to the dirsql folder in GG-BD
2. Specify this copied SOURCE\_DEFGEN.def file location in REPLICAT param file(rkafka.prm)

For every DDL change happen to the source tables (only tables which are captured for replication) generate a NEW sourcedef file and copied to the target server

Start replicat by using

ggsci > start rkafka

**SMOKE Test**

1. Execute an update query in Customers table.
2. Check the kafka topics elastic\_search, mySchemaTopic are created by

**kafka-topics.bat --list --zookeeper localhost:2181**

1. To list the messages under topic elastic\_test

D:\Setups\kafka\_2.11-1.1.0\bin\windows>kafka-console-consumer.bat --zookeeper localhost:2181 --topic elastic\_test

1. Sample Queries

* insert into scott.customers (customer\_id,customer\_name,city) values (3,'Rajiv','tvm');
* update scott.customers set customer\_name='rahul';
* delete from scott.customers where customer\_id=3;

1. Kafka Messages created for the above operations

**Update Query**

{"table":"MYSCHEMATOPIC.CUSTOMERS","op\_type":"U","op\_ts":"2018-07-22 12:45:37.00

0000","current\_ts":"2018-07-22T12:45:41.489000","pos":"00000000030000001870","be

fore":{"CUSTOMER\_ID":1,"CUSTOMER\_NAME":"rahul","CITY":"kollam"},"after":{"CUSTOM

ER\_ID":1,"CUSTOMER\_NAME":"arun","CITY":"kollam"}}

**Insert Query**

{"table":"MYSCHEMATOPIC.CUSTOMERS","op\_type":"I","op\_ts":"2018-07-22 12:51:53.00

0000","current\_ts":"2018-07-22T12:51:55.751000","pos":"00000000030000002069","af

ter":{"CUSTOMER\_ID":3,"CUSTOMER\_NAME":"Rajiv","CITY":"tvm"}}

**Delete Query**

{"table":"MYSCHEMATOPIC.CUSTOMERS","op\_type":"D","op\_ts":"2018-07-22 12:53:59.00

0000","current\_ts":"2018-07-22T12:54:02.392000","pos":"00000000030000002223","be

fore":{"CUSTOMER\_ID":3,"CUSTOMER\_NAME":"Rajiv","CITY":"tvm"}}

1. **rkafka.prm**

REPLICAT rkafka

TARGETDB LIBFILE ggjava.dll SET property=dirprm/kafka.props

SOURCEDEFS ./dirsql/SOURCE\_DEFGEN.def

REPORTCOUNT EVERY 1 MINUTES, RATE

GROUPTRANSOPS 10000

MAP SCOTT.CUSTOMERS, TARGET mySchemaTopic.\*;

1. **kafka.props**

**gg.handlerlist = kafkahandler**

**gg.handler.kafkahandler.type=kafka**

**gg.handler.kafkahandler.KafkaProducerConfigFile=custom\_kafka\_producer.properties**

**#The following resolves the topic name using the short table name**

**gg.handler.kafkahandler.topicMappingTemplate=****elastic\_test**

**#The following selects the message key using the concatenated primary keys**

**gg.handler.kafkahandler.keyMappingTemplate=${primaryKeys}**

**gg.handler.kafkahandler.format=avro\_op**

**gg.handler.kafkahandler.SchemaTopicName=mySchemaTopic**

**gg.handler.kafkahandler.BlockingSend =false**

**gg.handler.kafkahandler.includeTokens=false**

**gg.handler.kafkahandler.mode=op**

**goldengate.userexit.writers=javawriter**

**javawriter.stats.display=TRUE**

**javawriter.stats.full=TRUE**

**gg.log=log4j**

**gg.log.level=INFO**

**gg.report.time=30sec**

**#Sample gg.classpath for Apache Kafka**

**gg.classpath=D:/JAVA\_STUDY/LIVESTREAMING/GG-BD-12.3.2.1.0/dirprm/libs/\***

**#Sample gg.classpath for HDP**

**#gg.classpath=/etc/kafka/conf:/usr/hdp/current/kafka-broker/libs/\***

**javawriter.bootoptions=-Xmx512m -Xms32m -Djava.class.path=ggjava/ggjava.jar**

1. **custom\_kafka\_producer.properties**

**bootstrap.servers=localhost:9092**

**acks=1**

**reconnect.backoff.ms=1000**

**value.serializer=org.apache.kafka.common.serialization.ByteArraySerializer**

**key.serializer=org.apache.kafka.common.serialization.ByteArraySerializer**

**# 100KB per partition**

**batch.size=16384**

**linger.ms=0**

<https://rmoff.net/2017/11/21/installing-oracle-goldengate-for-big-data-12-3-1-with-kafka-connect-and-confluent-platform/>

**To enable DDL Replication**

<http://www.vitalsofttech.com/configure-goldengate-ddl-replication/>

<https://ronniethedba.wordpress.com/2016/10/26/please-move-ggs_admin-to-its-own-tablespace/>

<https://sacoefrancis.blogspot.com/2017/11/oracle-goldengate-for-big-data-and.html> **(BEST TUTORIAL)**

<https://medium.com/@shaaslam/installing-apache-kafka-on-windows-495f6f2fd3c8> **(QUICK CONFIGURATION OF KAFKA)**

[**https://docs.oracle.com/goldengate/bd123110/gg-bd/GADBD/using-kafka-connect-handler.htm#GADBD-GUID-23F5CCE3-845C-43F0-A08E-42C2BD1824FB**](https://docs.oracle.com/goldengate/bd123110/gg-bd/GADBD/using-kafka-connect-handler.htm#GADBD-GUID-23F5CCE3-845C-43F0-A08E-42C2BD1824FB)**( KAFKA CONNECT HANDLER CONFIGURATION)**

**===================================================================================================================**

<https://sacoefrancis.blogspot.com/2017/11/oracle-goldengate-for-big-data-and.html>

<https://rmoff.net/2017/11/21/installing-oracle-goldengate-for-big-data-12-3-1-with-kafka-connect-and-confluent-platform/>

<https://zdatainc.com/2014/07/real-time-streaming-apache-storm-apache-kafka/>

<https://dzone.com/articles/big-data-trifecta-storm-kafka-0>

<https://endocode.com/blog/2015/04/22/building-a-stream-processing-pipeline-with-kafka-storm-and-cassandra-part-2-using-docker-containers/>

<https://serrate.net/2016/01/05/analysis-of-twitter-streams-with-kafka-and-storm/>

<https://blog.insightdatascience.com/building-a-forex-trading-platform-using-kafka-storm-and-cassandra-a48b262facc2>

<https://zdatainc.com/2014/07/real-time-streaming-apache-storm-apache-kafka/>

<http://www.ateam-oracle.com/oracle-goldengate-logdump/> **(LOG DUMP TUTORIAL)**

<http://mygoldengat.blogspot.com/2016/04/troubleshooting-oracle-goldengate.html>

<https://dbasolved.com/2013/11/27/oracle-golden-gate-12c-on-windows-2008r2-sql-server-testing/> **(To fix the MSVCP100.dll file missing issue)**

**To list kafka topics**

kafka-topics.bat --list --zookeeper localhost:2181

**To view the messages in a topic**

kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic **elastic\_test** --from-beginning

kafka-console-consumer.bat --bootstrap-server localhost:9092 --topic **mySchemaTopic** --from-beginning