Session 23:

More Kafka

Assignment 1

Task 1:

Create a java program MyKafkaProducer.java that takes a file name and delimiter as input arguments.

It should read the content of file line by line.

Fields in the file are in following order

1. Kafka Topic Name

2. Key

3. value

For every line, insert the key and value to the repsective Kafka broker in a fire and forget mode.

After record is sent, it should print appropriate message on screen.

Pass dataset_producer.txt as the input file and -as delimiter.

LINK: https://drive.google.com/file/d/0B_Qjau8wv1KoSnR5eHpKOF9rTFU/view?usp=sharing

Solution:

Imports required for the program is given below:

```
import
org.apache.kafka.clients.producer.KafkaProd
ucer; import
org.apache.kafka.clients.producer.ProducerR
ecord; import java.io.BufferedReader;
import java.io.FileReader; import java.io.IOException; import
java.util.Properties
```

We configure the properties for KafkaProducer:

- *We create a new instance of Properties called props*
- *Using this instance we add properties to kafkaProducer like,*

bootstrapserver/meta-data-brokerlist, key and value serializers

```
Properties props = new Properties();
props.put("bootstrap.servers",
"localhost:9092");
props.put("key.serializer",
"org.apache.kafka.common.serialization.StringSerializer");
props.put("value.serializer",
"org.apache.kafka.common.serialization.StringSerializer");
```

We then instantiate the KafkaProducer class called producer, we have mentioned string in <> because both key and value are String

We add the properties instance (props)to KafkaProducer instance

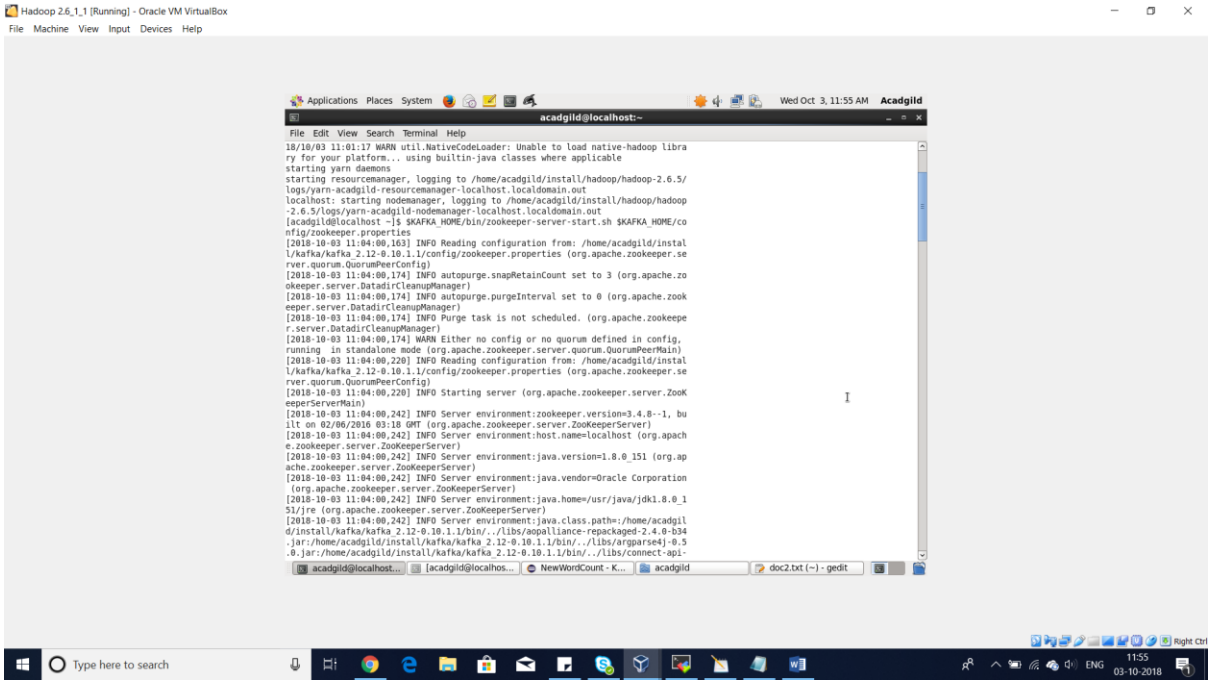
We also instantiate ProducerRecord as producerRecord

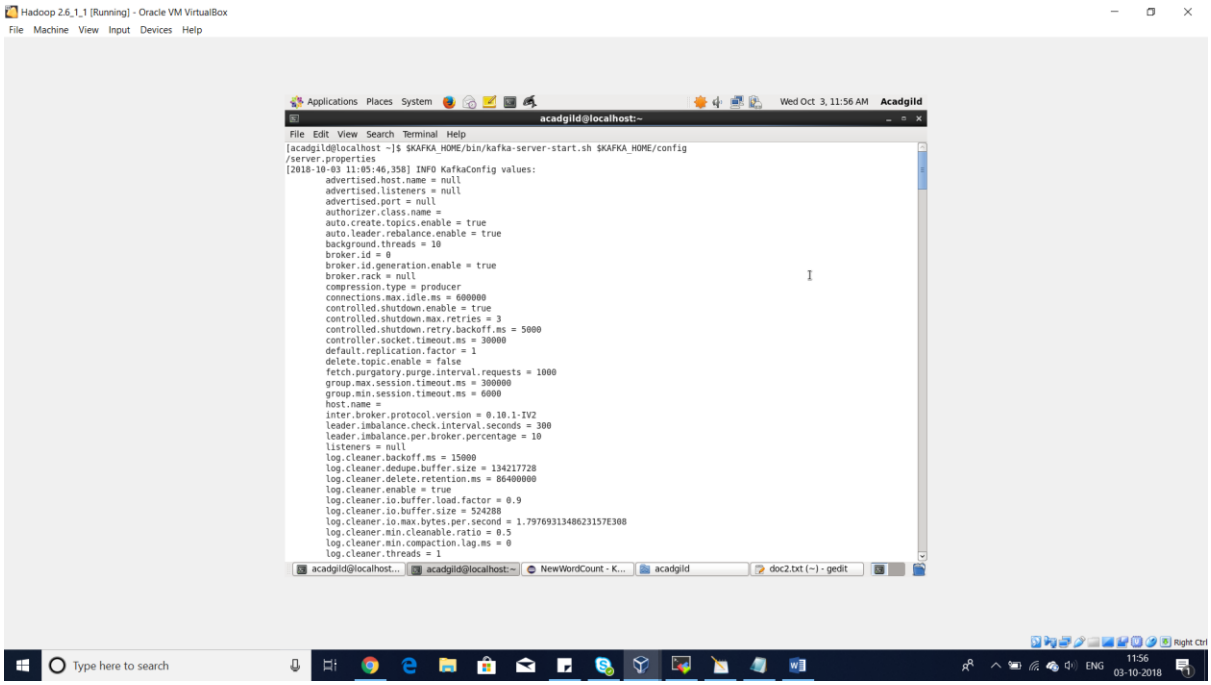
```
KafkaProducer<String, String> producer = new KafkaProducer<>(props);
ProducerRecord<String, String> producerRecord = null;
```

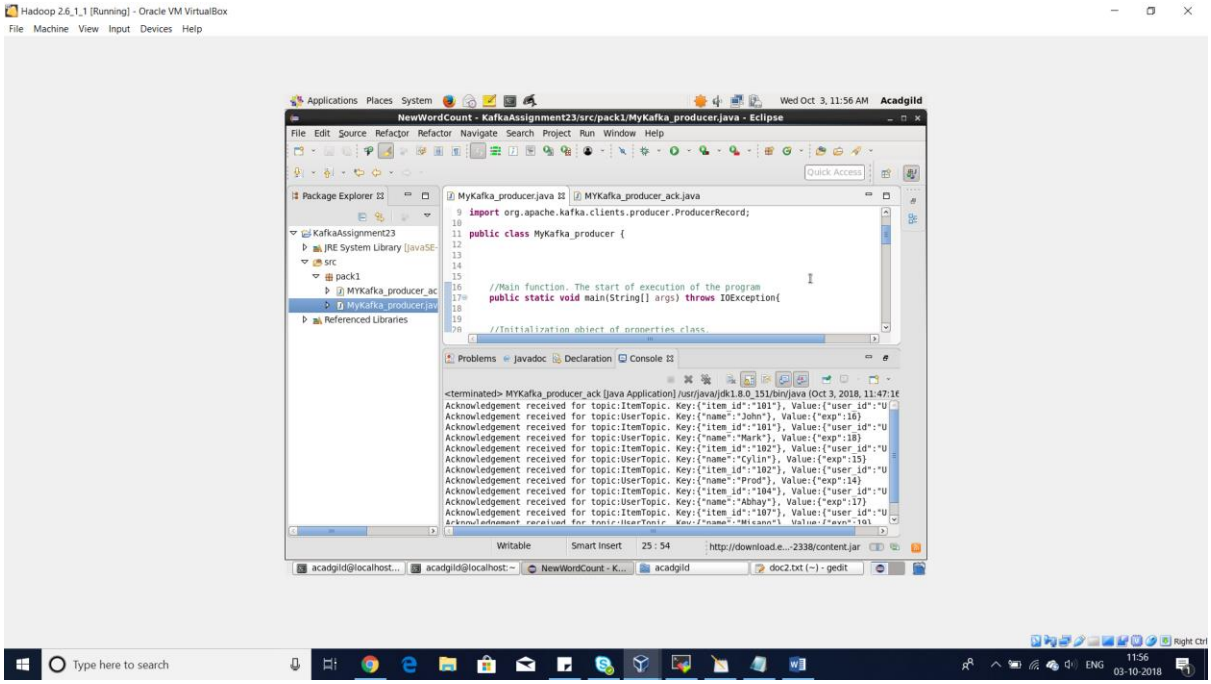
In the terminal we run the following commands:

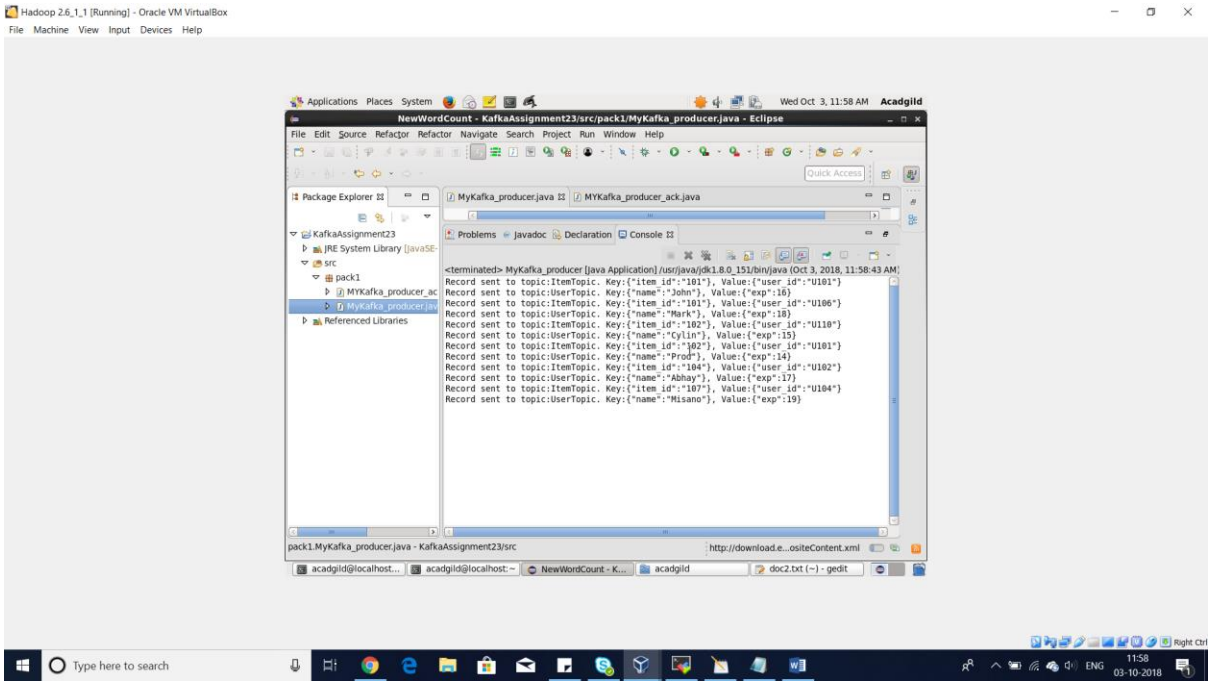
```
##Zookeeper Start Command $KAFKA_HOME/bin/zookeeper-server-start.sh \
$KAFKA_HOME/config/zookeeper.properties
```

```
## Starting broker $KAFKA_HOME/bin/kafka-server-start.sh \
$KAFKA_HOME/config/server.properties
```


Hadoop 2.6_1_1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Wed Oct 3, 11:55 AM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
18/10/03 11:01:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-resourcemanager-localhost.localdomain.out
localhost: starting nodemanager, logging to /home/acadgild/install/hadoop/hadoop-2.6.5/logs/yarn-acadgild-nodemanager-localhost.localdomain.out
[acadgild@localhost ~]$ $KAFKA_HOME/bin/zookeeper-server-start.sh $KAFKA_HOME/config/zookeeper.properties
[2018-10-03 11:04:00,163] INFO Reading configuration from: /home/acadgild/install/kafka/kafka_2.12-0.10.1.1/config/zookeeper.properties (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2018-10-03 11:04:00,174] INFO autopurge.snapRetainCount set to 3 (org.apache.zookeeper.server.DatadirCleanupManager)
[2018-10-03 11:04:00,174] INFO autopurge.purgeInterval set to 0 (org.apache.zookeeper.server.DatadirCleanupManager)
[2018-10-03 11:04:00,174] INFO Purge task is not scheduled. (org.apache.zookeeper.server.DatadirCleanupManager)
[2018-10-03 11:04:00,174] WARN Either no config or no quorum defined in config, running in standalone mode (org.apache.zookeeper.server.quorum.QuorumPeerMain)
[2018-10-03 11:04:00,220] INFO Reading configuration from: /home/acadgild/install/kafka/kafka_2.12-0.10.1.1/config/zookeeper.properties (org.apache.zookeeper.server.quorum.QuorumPeerConfig)
[2018-10-03 11:04:00,220] INFO Starting server (org.apache.zookeeper.server.ZooKeeperServerMain)
[2018-10-03 11:04:00,242] INFO Server environment:zookeeper.version=3.4.8--1, built on 02/06/2016 03:18 GMT (org.apache.zookeeper.server.ZooKeeperServer)
[2018-10-03 11:04:00,242] INFO Server environment:host.name=localhost (org.apache.zookeeper.server.ZooKeeperServer)
[2018-10-03 11:04:00,242] INFO Server environment:java.version=1.8.0_151 (org.apache.zookeeper.server.ZooKeeperServer)
[2018-10-03 11:04:00,242] INFO Server environment:java.vendor=Oracle Corporation (org.apache.zookeeper.server.ZooKeeperServer)
[2018-10-03 11:04:00,242] INFO Server environment:java.home=/usr/java/jdk1.8.0_151/jre (org.apache.zookeeper.server.ZooKeeperServer)
[2018-10-03 11:04:00,242] INFO Server environment:java.class.path=:/home/acadgild/install/kafka/kafka_2.12-0.10.1.1/bin/../libs/aopalliance-repackaged-2.4.0-b34.jar:/home/acadgild/install/kafka/kafka_2.12-0.10.1.1/bin/../libs/argparse4j-0.5.0.jar:/home/acadgild/install/kafka/kafka_2.12-0.10.1.1/bin/../libs/connect-api-
acadgild@localhost... [acadgild@localhos... NewWordCount - K... acadgild doc2.txt (~) - gedit
Type here to search
ENG 11:55 03-10-2018


Hadoop 2.6_1_1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Wed Oct 3, 11:56 AM Acadgild
acadgild@localhost:~
File Edit View Search Terminal Help
[acadgild@localhost ~]$ $KAFKA_HOME/bin/kafka-server-start.sh $KAFKA_HOME/config/server.properties
[2018-10-03 11:05:46,358] INFO KafkaConfig values:
advertised.host.name = null
advertised.listeners = null
advertised.port = null
authorizer.class.name =
auto.create.topics.enable = true
auto.leader.rebalance.enable = true
background.threads = 10
broker.id = 0
broker.id.generation.enable = true
broker.rack = null
compression.type = producer
connections.max.idle.ms = 600000
controlled.shutdown.enable = true
controlled.shutdown.max.retries = 3
controlled.shutdown.retry.backoff.ms = 5000
controller.socket.timeout.ms = 30000
default.replication.factor = 1
delete.topic.enable = false
fetch.purgatory.purge.interval.requests = 1000
group.max.session.timeout.ms = 300000
group.min.session.timeout.ms = 6000
host.name =
inter.broker.protocol.version = 0.10.1-IV2
leader.imbalance.check.interval.seconds = 300
leader.imbalance.per.broker.percentage = 10
listeners = null
log.cleaner.backoff.ms = 15000
log.cleaner.dedupe.buffer.size = 134217728
log.cleaner.delete.retention.ms = 86400000
log.cleaner.enable = true
log.cleaner.io.buffer.load.factor = 0.9
log.cleaner.io.buffer.size = 524288
log.cleaner.io.max.bytes.per.second = 1.7976931348623157E308
log.cleaner.min.cleanable.ratio = 0.5
log.cleaner.min.compaction.lag.ms = 0
log.cleaner.threads = 1
acadgild@localhost... acadgild@localhost:~ NewWordCount - K... acadgild doc2.txt (~) - gedit
Type here to search
ENG 11:56 03-10-2018





we run the console consumer commands on terminal to view the output of the program, using the below command:

- To read contents of ItemTopic:

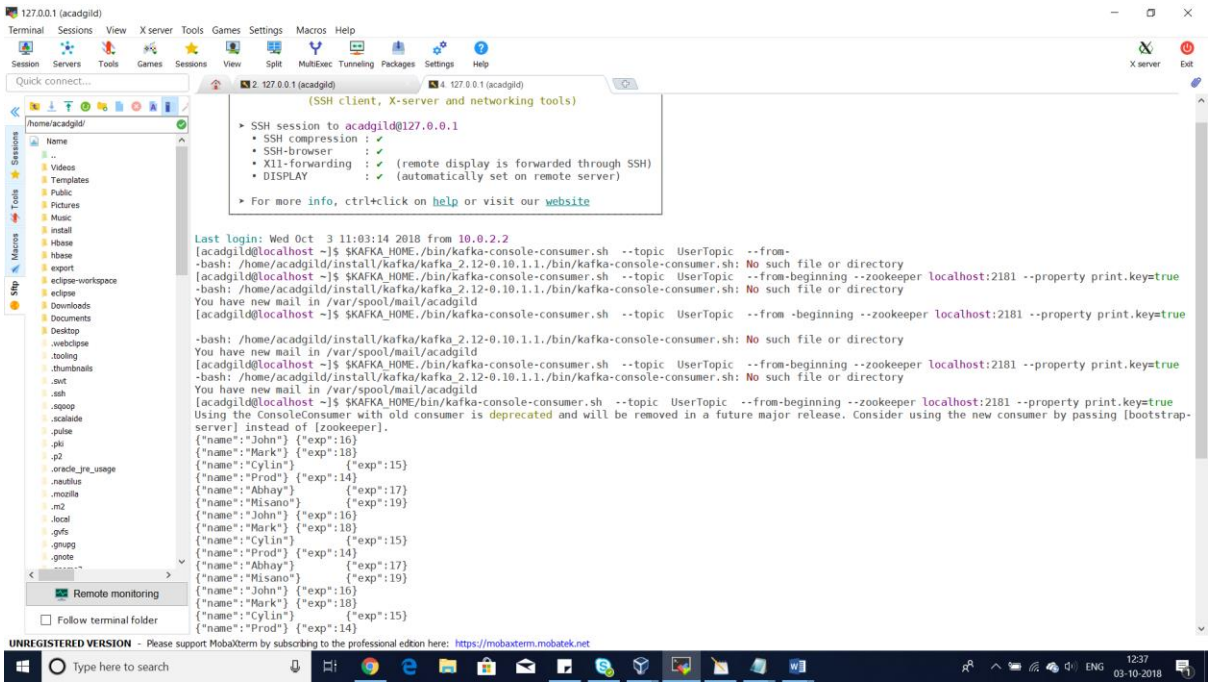
$ $KAFKA_HOME/bin/kafka-console-consumer.sh --topic ItemTopic --from-

beginning --zookeeper localhost:2181 --property print.key=true

- To read contents of UserTopic:

$ $KAFKA_HOME./bin/kafka-console-consumer.sh --topic UserTopic --from-

beginning --zookeeper localhost:2181 --property print.key=true



$ $KAFKA_HOME/bin/kafka-console-consumer.sh --topic ItemTopic --from-

beginning --zookeeper localhost:2181 --property print.key=true

```
[acadgild@localhost ~]$ $KAFKA_HOME/bin/kafka-console-consumer.sh  --topic  ItemTopic  --from-beginning --zookeeper localhost:2181 --property print.key=tru
Using the ConsoleConsumer with old consumer is deprecated and will be removed in a future major release. Consider using the new consumer by passing [bootst
server] instead of [zookeeper].
{"item_id":"101"}       {"user_id":"U101"}
{"item_id":"101"}       {"user_id":"U106"}
{"item_id":"102"}       {"user_id":"U110"}
{"item_id":"102"}       {"user_id":"U101"}
{"item_id":"104"}       {"user_id":"U102"}
{"item_id":"107"}       {"user_id":"U104"}
{"item_id":"101"}       {"user_id":"U101"}
{"item_id":"101"}       {"user_id":"U106"}
```

Task 2:

Modify the previous program MyKafkaProducer.java and create a new Java program

KafkaProducerWithAck.java

This should perform the same task as of KafkaProducer.java with some modification.

When passing any data to a topic, it should wait for acknowledgement.

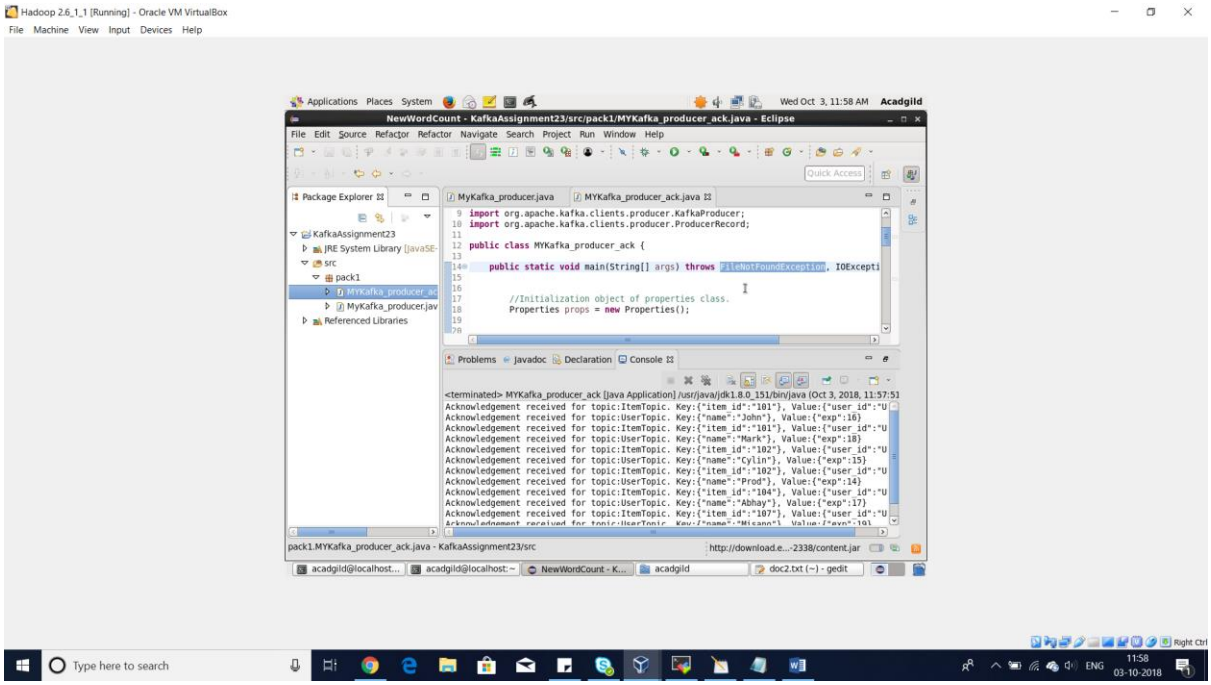
After acknowledgement is received from the broker, it should print the key and value which has been

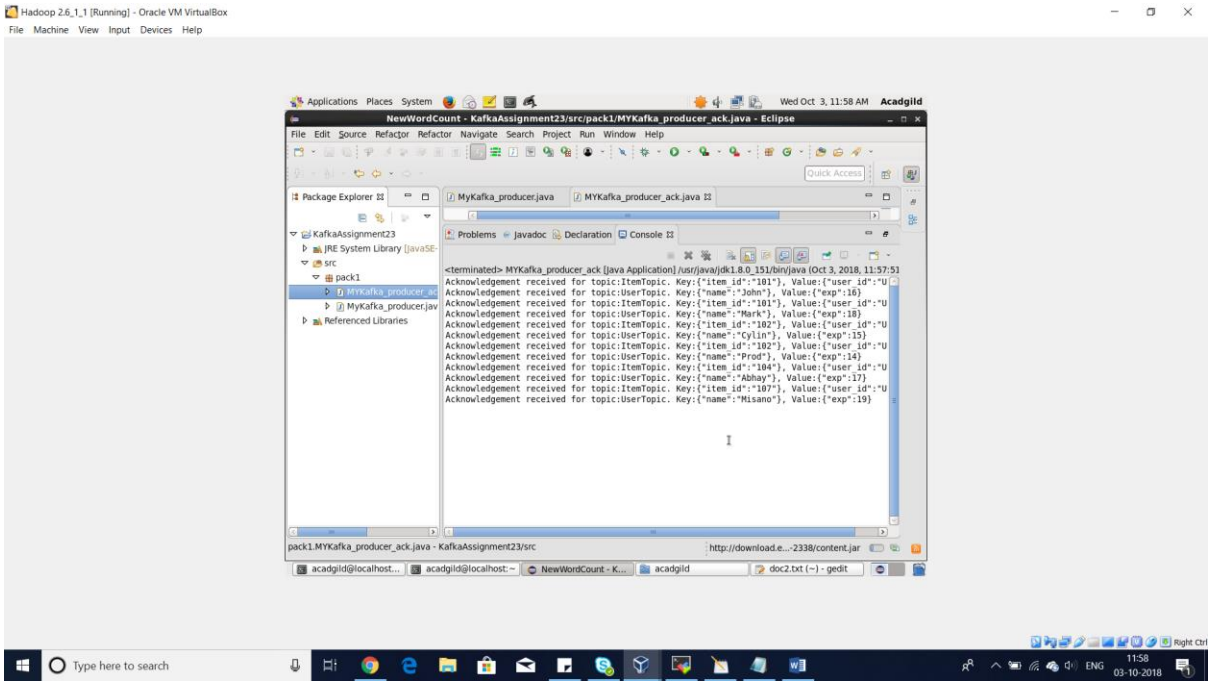
written to a specified topic.

The application should attempt for 3 retries before giving any exception.

Pass dataset_producer.txt as the input file and -as delimiter.

Solution:


Hadoop 2.6_1_1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Wed Oct 3, 11:58 AM Acadgild
NewWordCount - KafkaAssignment23/src/pack1/MYKafka_producer_ack.java - Eclipse
File Edit Source Refactor Refactor Navigate Search Project Run Window Help
Package Explorer
KafkaAssignment23
JRE System Library [javaSE-
src
pack1
MYKafka_producer_ac
MyKafka_producer.jav
Referenced Libraries
MyKafka_producer.java
MYKafka_producer_ack.java
import org.apache.kafka.clients.producer.KafkaProducer;
import org.apache.kafka.clients.producer.ProducerRecord;
public class MYKafka_producer_ack {
public static void main(String[] args) throws FileNotFoundException, IOExcepti
//Initialization object of properties class.
Properties props = new Properties();
Problems Javadoc Declaration Console
<terminated> MYKafka_producer_ack [Java Application] /usr/java/jdk1.8.0_151/bin/java (Oct 3, 2018, 11:57:51
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"John"}, Value:{"exp":16}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Mark"}, Value:{"exp":18}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Cylin"}, Value:{"exp":15}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Prod"}, Value:{"exp":14}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"104"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Abhay"}, Value:{"exp":17}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"107"}, Value:{"user_id":"U
pack1.MYKafka_producer_ack.java - KafkaAssignment23/src
http://download.e...-2338/content.jar
acadgild@localhost...
acadgild@localhost: ~
NewWordCount - K...
acadgild
doc2.txt (~) - gedit
Type here to search
ENG
11:58
03-10-2018


Hadoop 2.6_1_1 [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
Wed Oct 3, 11:58 AM Acadgild
NewWordCount - KafkaAssignment23/src/pack1/MYKafka_producer_ack.java - Eclipse
File Edit Source Refactor Refactor Navigate Search Project Run Window Help
Package Explorer
KafkaAssignment23
JRE System Library [javaSE-
src
pack1
MYKafka_producer_ac
MyKafka_producer.jav
Referenced Libraries
MyKafka_producer.java
MYKafka_producer_ack.java
Problems Javadoc Declaration Console
<terminated> MYKafka_producer_ack [Java Application] /usr/java/jdk1.8.0_151/bin/java (Oct 3, 2018, 11:57:51
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"John"}, Value:{"exp":16}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"101"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Mark"}, Value:{"exp":18}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Cylin"}, Value:{"exp":15}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"102"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Prod"}, Value:{"exp":14}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"104"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Abhay"}, Value:{"exp":17}
Acknowledgement received for topic:ItemTopic. Key:{"item_id":"107"}, Value:{"user_id":"U
Acknowledgement received for topic:UserTopic. Key:{"name":"Misano"}, Value:{"exp":19}
pack1.MYKafka_producer_ack.java - KafkaAssignment23/src
http://download.e...-2338/content.jar
acadgild@localhost...
acadgild@localhost: ~
NewWordCount - K...
acadgild
doc2.txt (~) - gedit
Type here to search
ENG
11:58
03-10-2018