Kafka @ PhonePe

Task 1:

- setup a 2 broker kafka cluster, use the already created ZK cluster for this as well

Prerequisites   :

1. **Java should be installed.**
2. **Set up a 3 node zookeeper cluster .**

Task:

1. Create Two More Vm's and set up kafka on top of it using the following commands. I am using 2.13_3.1.0 version of kafka.
   sudo wget https://dlcdn.apache.org/kafka/3.1.0/kafka_2.13-3.1.0.tgz

Kafka-Broker-node:

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo wget https://dlcdn.apache.org/kafka/3.
1.0/kafka_2.13-3.1.0.tgz
--2022-03-10 01:04:07--  https://dlcdn.apache.org/kafka/3.1.0/kafka_2.13-3.1.0.
tgz
Resolving dlcdn.apache.org (dlcdn.apache.org)... 151.101.2.132, 2a04:4e42::644
Connecting to dlcdn.apache.org (dlcdn.apache.org)|151.101.2.132|:443... connect
ed.
HTTP request sent, awaiting response... 200 OK
Length: 88130011 (84M) [application/x-gzip]
Saving to: 'kafka_2.13-3.1.0.tgz'

kafka_2.13-3.1.0.tg 100%[===================>]  84.05M  7.05MB/s    in 12s

2022-03-10 01:04:19 (6.81 MB/s) - 'kafka_2.13-3.1.0.tgz' saved [88130011/881300
11]
```

2. Extract the tar file using the tar command.

Kafka Broker Node:

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo tar -xzvf kafka_2.13-3.1.0.tgz -C /opt
kafka_2.13-3.1.0/
kafka_2.13-3.1.0/LICENSE
kafka_2.13-3.1.0/NOTICE
kafka_2.13-3.1.0/bin/
```

```
zookeeper1@zookeeper1-VirtualBox:~$ cd /opt
zookeeper1@zookeeper1-VirtualBox:/opt$ ls
kafka_2.13-3.1.0
zookeeper1@zookeeper1-VirtualBox:/opt$ sudo mv kafka_2.13-3.1.0 kafka
zookeeper1@zookeeper1-VirtualBox:/opt$ ls
kafka
```

3. Once the installation is done , let's just start configuring it.

We'll need to assign the broker id , Connection with the pre-set zookeeper nodes and the listener which is that kafka node itself.

Kafka-Broker-node:

```
zookeeper1@zookeeper1-VirtualBox:/opt/kafka/config$ sudo nano server.properties
```

```
  GNU nano 4.8                        server.properties
# limitations under the License.

# see kafka.server.KafkaConfig for additional details and defaults

######################### Server Basics #########################

# The id of the broker. This must be set to a unique integer for each broker.
broker.id=1

######################### Socket Server Settings #####################>

# The address the socket server listens on. It will get the value returned fro>
# java.net.InetAddress.getCanonicalHostName() if not configured.
#    FORMAT:
#      listeners = listener_name://host_name:port
#    EXAMPLE:
#      listeners = PLAINTEXT://your.host.name:9092
listeners=PLAINTEXT://10.0.2.14:9092
```

```
  GNU nano 4.8                        server.properties
# to the retention policies
log.retention.check.interval.ms=300000

######################### Zookeeper #########################

# Zookeeper connection string (see zookeeper docs for details).
# This is a comma separated host:port pairs, each corresponding to a zk
# server. e.g. "127.0.0.1:3000,127.0.0.1:3001,127.0.0.1:3002".
# You can also append an optional chroot string to the urls to specify the
# root directory for all kafka znodes.
zookeeper.connect=10.0.2.2:2181,10.0.2.6:2181,10.0.2.10:2181
```

Do the same steps on the other kafka broker node.

4. Before running the kafka brokers , make sure that the zookeeper server on all the three nodes is started.
   - ☐ Sudo /opt/zookeeper/bin/zkServer.sh start

   Run both the kafka brokers using the following command.
   - ☐ Sudo /opt/kafka/kafka-server-start.sh /opt/kafka/config/server.properties

Kafka-Broker-node:

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-server-start.sh /
opt/kafka/config/server.properties
[sudo] password for zookeeper1:
[2022-03-10 01:28:30,854] INFO Registered kafka:type=kafka.Log4jController MBea
n (kafka.utils.Log4jControllerRegistration$)
[2022-03-10 01:28:31,353] INFO Setting -D jdk.tls.rejectClientInitiatedRenegoti
a Files true to disable client-initiated TLS renegotiation (org.apache.zookeeper.
Common.X509Util)
[2022-03-10 01:28:31,458] INFO Registered signal handlers for TERM, INT, HUP (o
rg.apache.kafka.common.utils.LoggingSignalHandler)
[2022-03-10 01:28:31,461] INFO starting (kafka.server.KafkaServer)
[2022-03-10 01:28:31,462] INFO Connecting to zookeeper on 10.0.2.2:2181,10.0.2.
6:2181,10.0.2.10:2181 (kafka.server.KafkaServer)
[2022-03-10 01:28:31,482] INFO [ZooKeeperClient Kafka server] Initializing a ne
w session to 10.0.2.2:2181,10.0.2.6:2181,10.0.2.10:2181. (kafka.zookeeper.ZooKe
eperClient)
[2022-03-10 01:28:31,489] INFO Client environment:zookeeper.version=3.6.3--6401
e4ad2087061bc6b9f80dec2d69f2e3c8660a, built on 04/08/2021 16:35 GMT (org.apache
.zookeeper.ZooKeeper)
[2022-03-10 01:28:31,489] INFO Client environment:host.name=zookeeper1-VirtualB
ox (org.apache.zookeeper.ZooKeeper)
[2022-03-10 01:28:31,489] INFO Client environment:java.version=11.0.14 (org.apa
che.zookeeper.ZooKeeper)
[2022-03-10 01:28:31,489] INFO Client environment:java.vendor=Ubuntu (org.apach
e.zookeeper.ZooKeeper)
[2022-03-10 01:28:31,489] INFO Client environment:java.home=/usr/lib/jvm/java-1
1-openjdk-amd64 (org.apache.zookeeper.ZooKeeper)
[2022-03-10 01:28:31,489] INFO Client environment:java.class.path=/opt/kafka/bi
n/../libs/activation-1.1.1.jar:/opt/kafka/bin/../libs/aopalliance-repackaged-2.
```

The connection is set up and this we can tell because we have the brokers listed on the zookeeper node.

Zookeeper node 1:

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/zookeeper/bin/zkServer.sh start
[sudo] password for zookeeper1:
/usr/bin/java
ZooKeeper JMX enabled by default
Using config: /opt/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
zookeeper1@zookeeper1-VirtualBox:~$ echo dump | nc 10.0.2.2 2181 | grep brokers
        /brokers/ids/1
        /brokers/ids/2
```

Task 2:

- ● Create a topic by name pp-intern with 2 replication and 2 partitions

Start the kafka server and run the following on the terminal.

Sudo /opt/kafka/bin/kafka-topics.sh --create --topic pp-intern --bootstrap-server
10.0.2.18:9092,10.0.2.14:9092 --replication-factor 2 --partitions 2

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-topics.sh --creat
e --topic pp-intern --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092 --replicat
ion-factor 2 --partitions 2
Created topic pp-intern.
```

List the topics that are there on the kafka via this command :

Sudo /opt/kafka/bin/kafka-topics.sh --list --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-topics.sh --list
--bootstrap-server 10.0.2.18:9092,10.0.2.14:9092
pp-intern
```

- **Produce to this topic from console**

Producing the topic to console means writing events to the producer .
This can be done via this command:

Sudo /opt/kafka/bin/kafka-console-producer.sh --topic pp-intern --broker-list
10.0.2.14:9092,10.0.2.18:9092

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-console-producer.
sh --topic pp-intern --broker-list 10.0.2.18:9092,10.0.2.14:9092
>hello pp-intern!
>hi all!
>add these messages to kafka broker.
```

- **Consume from this topic from console**

Read the events via the kafka consumer.

Sudo /opt/kafka/bin/kafka-console-consumer.sh --topic pp-intern --from-beginning --bootstrap-server
10.0.2.14:9092,10.0.2.18:9092

```
>^Czookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-console-consum
er.sh --topic pp-intern --from-beginning --bootstrap-server 10.0.2.18:9092,10.0
.2.14:9092
hi all!
hello pp-intern!
add these messages to kafka broker.
```

- Alter the message retention time for this topic to 2 days

Alter the retention period for this topic to 2 days.

Sudo /opt/kafka/bin/kafka-configs.sh --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092 --alter --entity-type topics --entity-name pp-intern --add-config retention.ms=172800000

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-configs.sh --boot
strap-server 10.0.2.18:9092,10.0.2.14:9092 --alter --entity-type topics --entit
y-name pp-intern --add-config retention.ms=172800000
Completed updating config for topic pp-intern.
```

- Try to increase the partitions from 2 to 3

We can increase the partition number .

Sudo /opt/kafka/bin/kafka-topics.sh --alter --bootstrap-server 10.0.2.14:9092,10.0.2.18:9092 --topic pp-intern --partitions 3

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-topics.sh --alter
 --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092 --topic pp-intern --partition
s 3
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-topics.sh --descr
ibe --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092 --topic pp-intern
Topic: pp-intern         TopicId: pdCGx1FuQ_ijKJ3w4IjlHA PartitionCount: 3      R
eplicationFactor: 2      Configs: segment.bytes=1073741824,retention.ms=17280000
0
        Topic: pp-intern        Partition: 0    Leader: 1       Replicas: 1,2 I
sr: 1,2
        Topic: pp-intern        Partition: 1    Leader: 2       Replicas: 2,1 I
sr: 2,1
        Topic: pp-intern        Partition: 2    Leader: 1       Replicas: 1,2 I
sr: 1,2
```

- Try to reduce the partitions from 3 to 2

   Apache Kafka doesn't support decreasing the partition number. The partitions are a way for scaling out improving performance. So all data sent to the topic flows to all partitions and removing one of them means data loss.

```
zookeeper1@zookeeper1-VirtualBox:~$ sudo /opt/kafka/bin/kafka-topics.sh --alter
 --bootstrap-server 10.0.2.18:9092,10.0.2.14:9092 --topic pp-intern --partition
s 1
Error while executing topic command : Topic currently has 3 partitions, which i
s higher than the requested 1.
[2022-03-10 02:08:22,012] ERROR org.apache.kafka.common.errors.InvalidPartition
sException: Topic currently has 3 partitions, which is higher than the requeste
d 1.
 (kafka.admin.TopicCommand$)
```