Real-Time- Stock Market Data Analysis using Kafka

Commands Used:

Login to EC2 Machine

Navigate to the directory where the key is present and then,

- cd documents/data_engineering-projects-kafka_stock_market_project (*my key is present in this directory)

- ssh -i "kafka-stock-market-project.pem" ec2-user@ec2-**-***-***-***.compute-1.amazonaws.com

----

- wget https://downloads.apache.org/kafka/3.6.0/kafka_2.12-3.6.0.tgz
  (install kafka)

- tar -xvf kafka_2.12-3.6.0.tgz
  (x – extract the zip file,
  v- be verbose about it,
  f-the next argument is going to be the file name that is to be extracted)

Install Java

- java –version

- wget https://download.oracle.com/java/20/archive/jdk-20_linux-x64_bin.rpm

- sudo rpm -ivh jdk-20_linux-x64_bin.rpm

- java –version

- cd kafka_2.12-3.6.0

Start the Zookeeper Server

- bin/zookeeper-server-start.sh config/zookeeper.properties

Open another terminal window to start kafka
But first ssh to ec2 machine as done above.

Start Kafka Server

- export KAFKA_HEAP_OPTS="-Xmx256M -Xms128M" ****TO INCREASE THE MEMORY****

- cd kafka_2.12-3.6.0

- bin/kafka-server-start.sh config/server.properties

It points to private server, change server.properties so that it can run in public IP

To do this,
- sudo nano config/server.properties

   In this, change ADVERTISED_LISTENERS to public ip of the EC2 instance

---

Create the Topic

Open another terminal window to create topic and start producer.
But first ssh to ec2 machine as done above.

- cd kafka_2.12-3.6.0

- bin/kafka-topics.sh --create --topic demo_test --bootstrap-server **-***-***-***:9092 --replication-factor 1 --partitions 1

bin/kafka-topics.sh --create --topic demo_testing2 --bootstrap-server {**Put the Public IP of your EC2 Instance**:9092} --replication-factor 1 --partitions 1

(**Error**:: Timed out waiting for a node assignment
Resolution: In EC2->Security->Security Groups-> edit inbound rules -> instead of myIP, select all IP4 ---------→ This worked for me and the topic got created successfully.)

Start Producer

- bin/kafka-console-producer.sh --topic demo_test --bootstrap-server **-***-***-***:9092

Start Consumer

Open another terminal window to create start consumer.
But first ssh to ec2 machine as done above.

- cd kafka_2.12-3.6.0

- bin/kafka-console-consumer.sh --topic demo_test --bootstrap-server **-***-***-***:9092

So now in
  Terminal 1 – ZooKeeper will be running.
  Terminal 2 – Kafka will be running.
  Terminal 3 – Producer will be running.
  Terminal 4 – Consumer will be running.