Apache Kafka Documentation

Introduction:-

Apache Kafka is a distributed publish-subscribe messaging system and a robust queue that can handle a high volume of data and enables you to pass messages from one end-point to another. Kafka is suitable for both offline and online message consumption. Kafka messages are persisted on the disk and replicated within the cluster to prevent data loss. Kafka is built on top of the ZooKeeper synchronization service. It integrates very well with Apache Storm and Spark for real-time streaming data analysis.

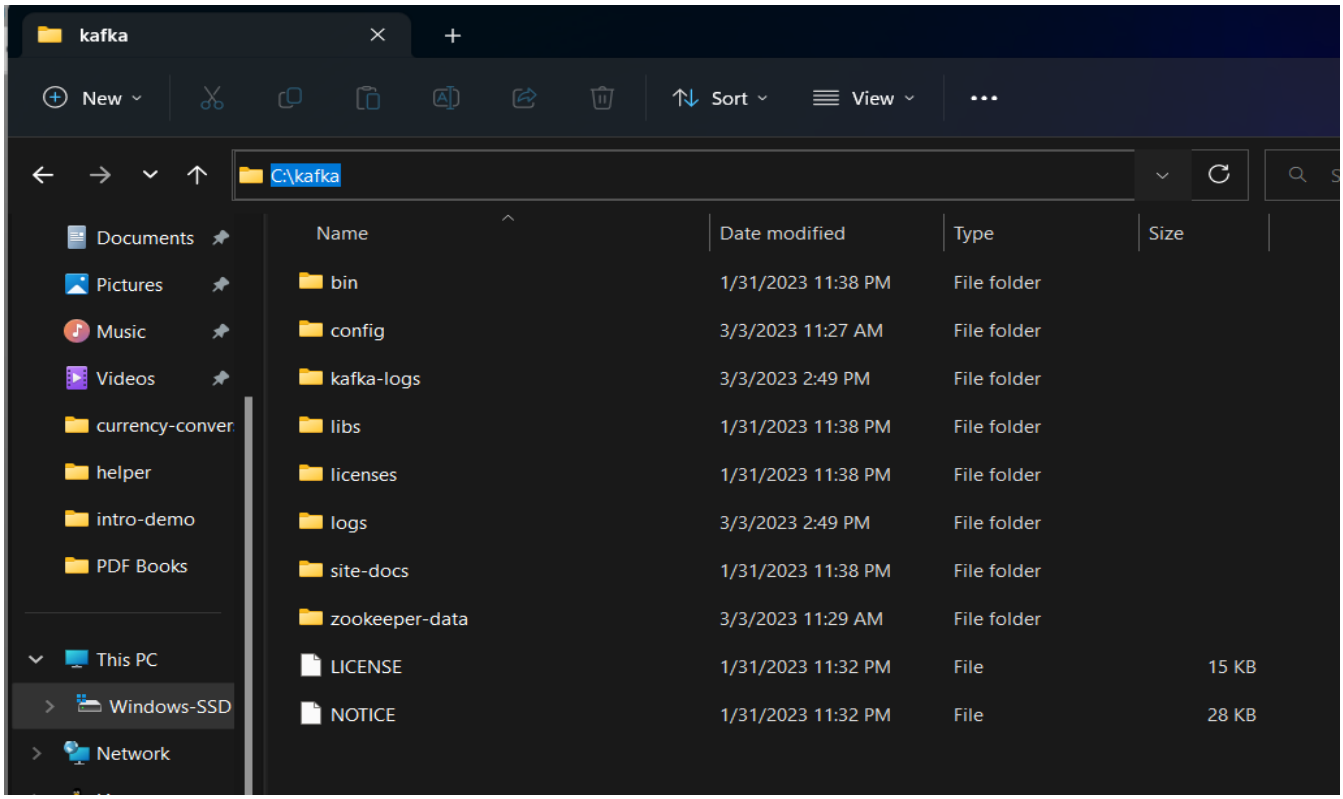
Benefits

Following are a few benefits of Kafka –

1. **Reliability** – Kafka is distributed, partitioned, replicated and fault tolerance.
2. **Scalability** – Kafka messaging system scales easily without down time..
3. **Durability** – Kafka uses Distributed commit log which means messages persists on disk as fast as possible, hence it is durable..
4. **Performance** – Kafka has high throughput for both publishing and subscribing messages. It maintains stable performance even many TB of messages are stored.

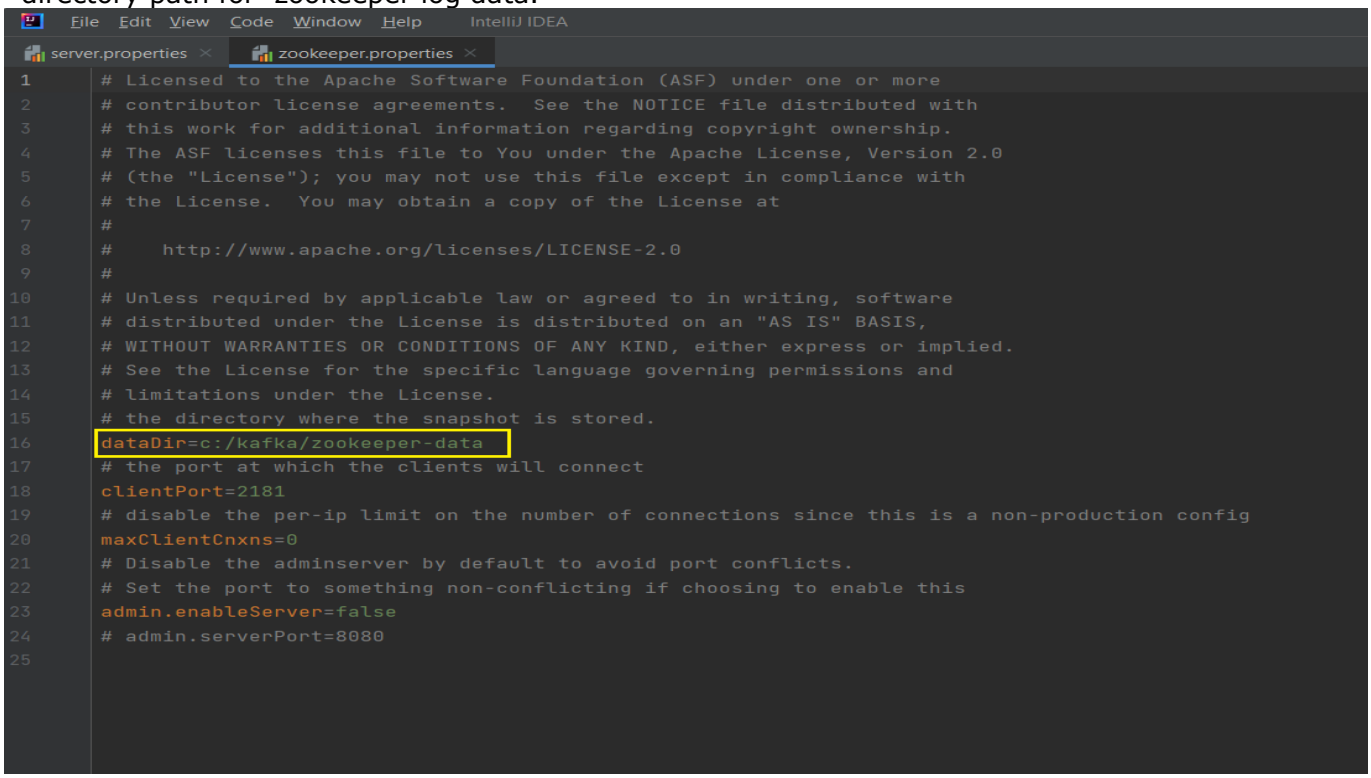
Installation of Windows:-

1. Download Apache Kafka latest release **kafka_x_xx_x.x.x.tgz** file from official site https://kafka.apache.org/downloads.
2. After download extract Tar file using 7zip and save extracted folder using name **kafka** and move to **C Drive.**

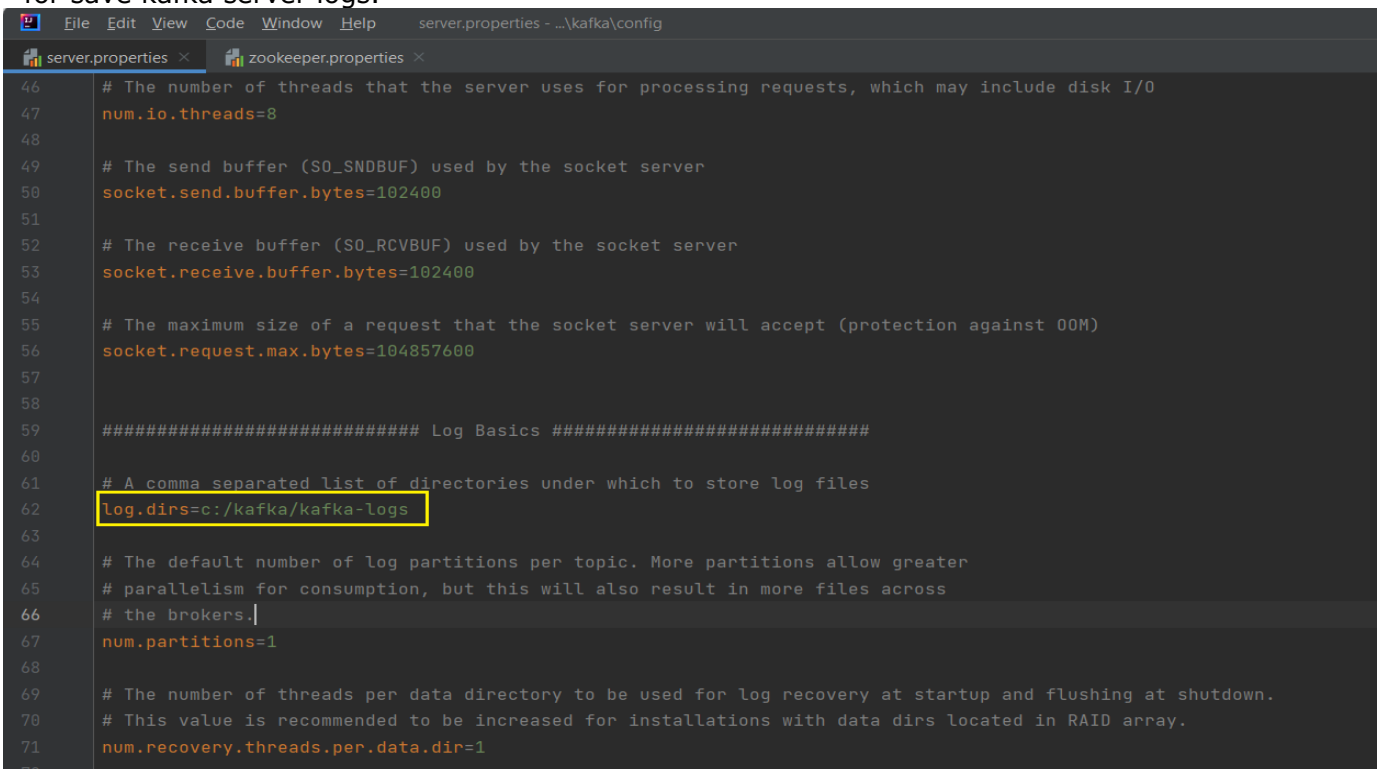


Configuration Kafka Server and ZooKeeper:-

1. In kafka folder go to **config** folder **zookeeper.properties** and edit it, go to **dataDir** section for set directory path for zookeeper log data.



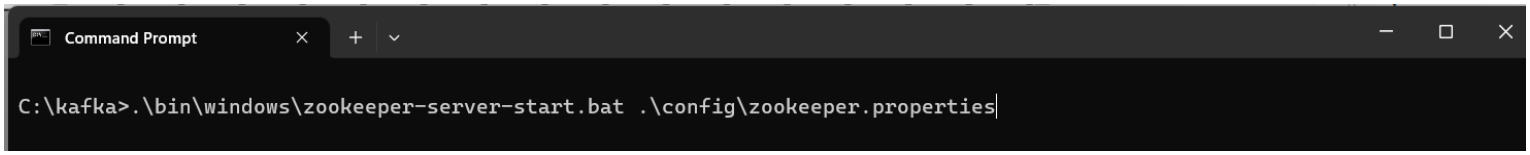
2. Same in config folder go to **server.properties** and edit it, go to **log.dirs** section and set directory path for save kafka server logs.



Run Kafka Server and ZooKeeper:-

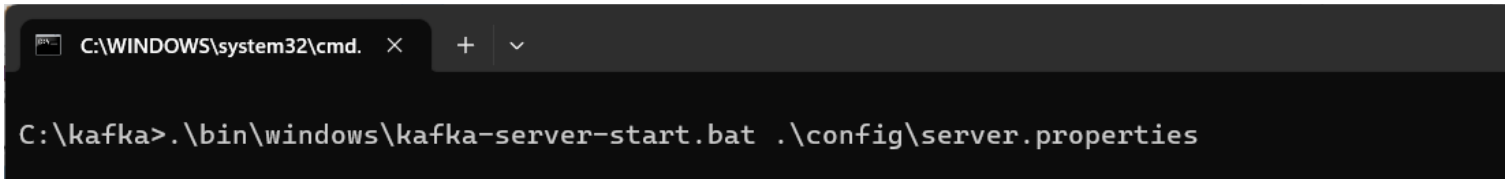
1. Open CMD and go to directory path of kafka stored, **select path cd C:\kafka**
2. For running ZooKeeper run following command,

.\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties



Zookeeper will be start running on default port **2181**.

3. Open another CMD for run kafka server go same above path **cd C:\kafka**
4. For running Kafka Server run following command,



Kafka server will be start on default port **9092.**

Create Topics:-

For creating topic in CMD run following command,

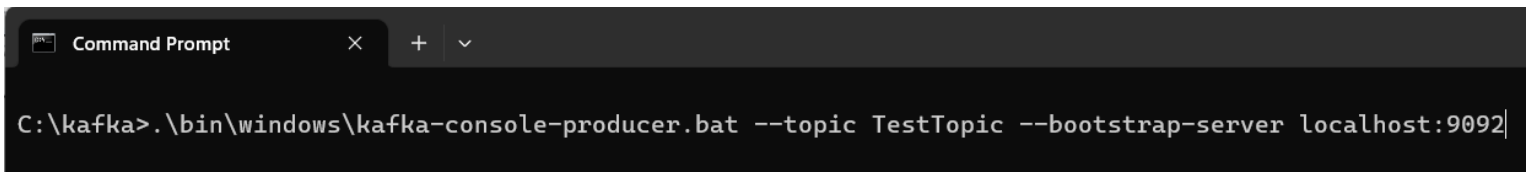
.\bin\windows\kafka-topics.bat --create --topic <topic_name> --bootstrap-server localhost:9092

Create Kafka Events using Kafka Producer:-

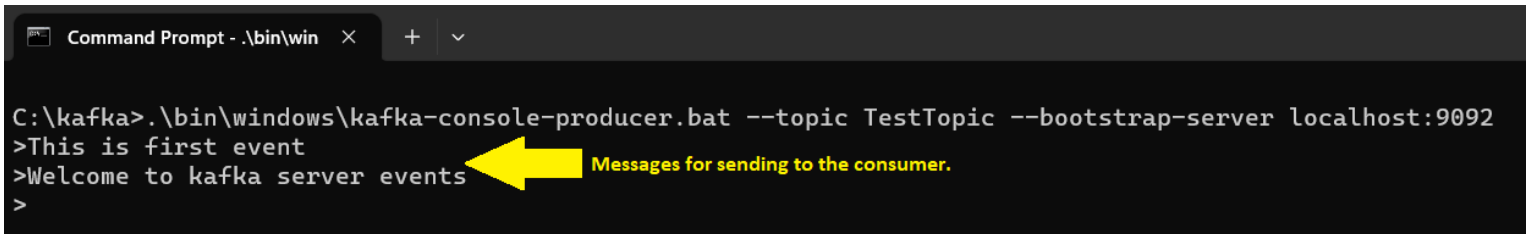
For creating events use following command,

.\bin\windows\kafka-console-producer.bat --topic <topic_name> --bootstrap-server localhost:9092

NOTE: Topic name should be same as created earlier.



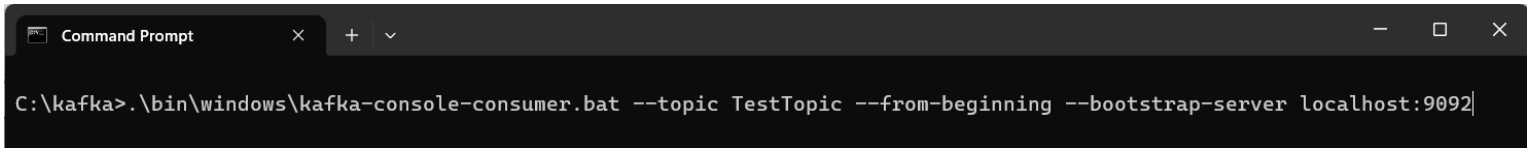
And hit enter and start typing some messages for sent to the consumer.



Listen Kafka Events using Kafka Consumer:-

For listen message coming from producer use following command,

.\bin\windows\kafka-console-consumer.bat --topic <topic_name> --from-beginning --bootstrap-server localhost:9092



Messages will be showing which is entered in producer screen,