
Step 1: Install Required Tools

1. Install Java (for Kafka)

Kafka requires Java to run. If you don't have Java installed, run the following:

```
sudo apt update
sudo apt install default-jdk
java -version
```

2. Install Kafka

Download and Extract Kafka: You can get the latest Kafka from [Apache Kafka Downloads](https://kafka.apache.org/downloads).

```
wget https://downloads.apache.org/kafka/3.7.0/kafka_2.13-3.7.0.tgz
tar -xvzf kafka_2.13-3.7.0.tgz
cd kafka_2.13-3.7.0
```

1.

Start Zookeeper: Kafka requires Zookeeper to coordinate cluster management.

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

2.

Start Kafka Broker: Kafka brokers are responsible for storing and managing messages.

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

3.



Step 2: Set Up MySQL

Install MySQL (if not already installed):

```
sudo apt install mysql-server
```

1.

Log in to MySQL:

If you're using the root user with `sudo`:

```
sudo mysql
```

2.

Create a Database and Table:

Once you're inside the MySQL shell:

```
CREATE DATABASE kafka_data;  
USE kafka_data;
```

```
CREATE TABLE messages (  
  id INT AUTO_INCREMENT PRIMARY KEY,  
  message VARCHAR(255),  
  timestamp TIMESTAMP DEFAULT CURRENT_TIMESTAMP  
);
```

3. This will create a `kafka_data` database and a `messages` table to store the incoming data.

Create a New User (Optional but Recommended)

It is recommended to create a new user for connecting to the MySQL database.

```
CREATE USER 'riddhi'@'localhost' IDENTIFIED BY 'mysecurepassword';  
GRANT ALL PRIVILEGES ON kafka_data.* TO 'riddhi'@'localhost';  
FLUSH PRIVILEGES;  
EXIT;
```

4.

Step 3: Install Python Libraries

We need to install the required Python libraries to interact with Kafka and MySQL.

```
pip install kafka-python mysql-connector-python
```

Step 4: Create Kafka Producer (Send Data to Kafka)

The producer will send data to Kafka.

Producer Code (**producer.py**):

```
from kafka import KafkaProducer
import json
import time

# Connect to Kafka
producer = KafkaProducer(
    bootstrap_servers='localhost:9092',
    value_serializer=lambda v: json.dumps(v).encode('utf-8')
)

# Send message every 5 seconds
while True:
    data = {"message": "Hello from Kafka"}
    producer.send('my-topic', value=data)
    print("Sent:", data)
    time.sleep(5) # Send data every 5 seconds
```

Explanation:

- **KafkaProducer**: Connects to the Kafka broker.
 - **value_serializer**: Serializes Python dictionaries to JSON format.
 - **producer.send()**: Sends data to the Kafka topic **my-topic**.
 - **time.sleep(5)**: Sends a message every 5 seconds.
-

Step 5: Create Kafka Consumer (Receive Data and Store in MySQL)

The consumer will listen to the Kafka topic and store the data into MySQL.

Consumer Code (**consumer.py**):

```

from kafka import KafkaConsumer
import mysql.connector
import json

# Kafka Consumer setup
consumer = KafkaConsumer(
    'my-topic', # Topic name
    bootstrap_servers='localhost:9092',
    group_id='my-group',
    value_deserializer=lambda m: json.loads(m.decode('utf-8')) # Deserialize JSON message
)

# MySQL connection
conn = mysql.connector.connect(
    host='localhost',
    user='riddhi',
    password='mysecurepassword', # Your MySQL password
    database='kafka_data'
)
cursor = conn.cursor()

# Listen to Kafka topic and insert messages into MySQL
for msg in consumer:
    data = msg.value # Get the message value
    print("Received:", data)

    # Insert data into MySQL table
    query = "INSERT INTO messages (message) VALUES (%s)"
    cursor.execute(query, (data['message'],))
    conn.commit() # Commit the transaction

# Close the connection when done
cursor.close()
conn.close()

```

Explanation:

- **KafkaConsumer**: Subscribes to **my-topic** and reads messages.
- **value_deserializer**: Decodes the JSON messages from Kafka.
- **mysql.connector**: Connects to MySQL using the **riddhi** user.

- The `INSERT INTO` SQL statement inserts the received message into the MySQL table `messages`.
-

Step 6: Run the Producer and Consumer

➤ Start Kafka and Zookeeper (if not already running):

Open two terminals:

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

➤ Run the Producer:

This will send data to Kafka every 5 seconds.

```
python producer.py
```

➤ Run the Consumer:

This will listen for messages from Kafka and insert them into the MySQL database.

```
python consumer.py
```

Step 7: Verify Data in MySQL

After running both the producer and consumer scripts, the messages will be stored in the MySQL `messages` table.

To check the data, log into MySQL:

```
mysql -u riddhi -p
```

Then select the `kafka_data` database:

```
USE kafka_data;
```

```
SELECT * FROM messages;
```

You should see the messages being stored.

Summary of the Steps

1. **Install Kafka and MySQL:** Install Java, Kafka, and MySQL on your system.
2. **Create Kafka Topic:** Create a topic to send and receive data.
3. **Create MySQL Database and Table:** Set up a MySQL database and table to store the data.
4. **Create Kafka Producer:** Write a Python script to send data to Kafka every 5 seconds.
5. **Create Kafka Consumer:** Write a Python script to consume data from Kafka and insert it into MySQL.
6. **Run Producer and Consumer:** Start the producer and consumer scripts.
7. **Verify the Data:** Check the MySQL table to ensure data is stored.