A REPORT ON

KAFKA Deployment And Creating Publisher/Consumer For ReST Based Servies

International Institute Of Information Technology, Bangalore

Computer Science and Engineering

Guidance by

Prof. Samar Shailendra

Submitted by

Shubham Gupta (IMT2016118)

Suparn S Lele (MT2018122)

INDEX

1. **Abstract**

2. **Introduction**
   - 2.1 Why Dev-Ops
   - 2.2 About Application
3. **SDLC Phase**
   - 3.1 Scope
   - 3.2 Architecture
     - 3.2.1 Jenkins in Docker Configuration
   - 3.3 SCM
   - 3.4 Build
   - 3.5 Test
   - 3.6 Deploy
   - 3.7 Monitor
4. **Results**
   - 4.1 Scalability
5. **Future Work**
6. **Conclusion**
7. **References**

1.0 Kafka Introduction:

Apache Kafka is an open-source stream-processing software platform developed by LinkedIn and donated to the Apache Software Foundation, written in Scala and Java. The project aims to provide a unified, high-throughput, low-latency platform for handling real-time data feeds.

Kafka uses a binary TCP design that is optimized for efficiency and relies on a "message set" abstraction that naturally groups messages together to reduce the overhead of the network roundtrip. This "leads to larger network packets, larger sequential disk operations, contiguous memory blocks [...] which allows Kafka to turn a bursty stream of random message writes into linear writes."

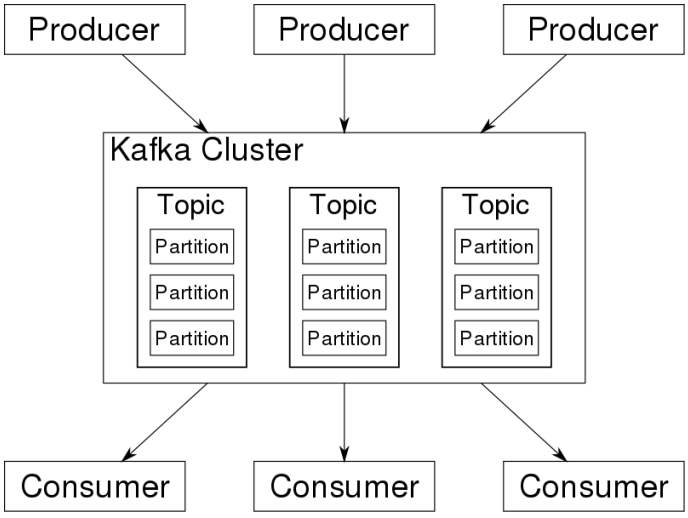


Fig 1. Kafka Architecture

2.0 Components :

2.1 Zookeeper

Apache Zookeeper is a distributed, open-source configuration,synchronization service along with naming registry for distributed applications. ZooKeeper stores a lot of shared information about **Kafka Consumers** and Kafka **Brokers.**

2.2 Broker

A Kafka cluster is made up of multiple Kafka Brokers. Each Kafka Broker has a unique ID (number). Kafka Brokers contain topic log partitions. Connecting to one broker bootstraps a client to the entire Kafka cluster. For failover, you want to start with at least three to five brokers. A Kafka cluster can have, 10, 100, or 1,000 brokers in a cluster if needed.

2.3 Producer

Producers push data to brokers. When the new broker is started, all the producers search it and automatically sends a message to that new broker.

2.4 Consumer

Consumers read data from brokers. Consumers subscribes to one or more topics and consume published messages by pulling data from the brokers.

2.5 Topic

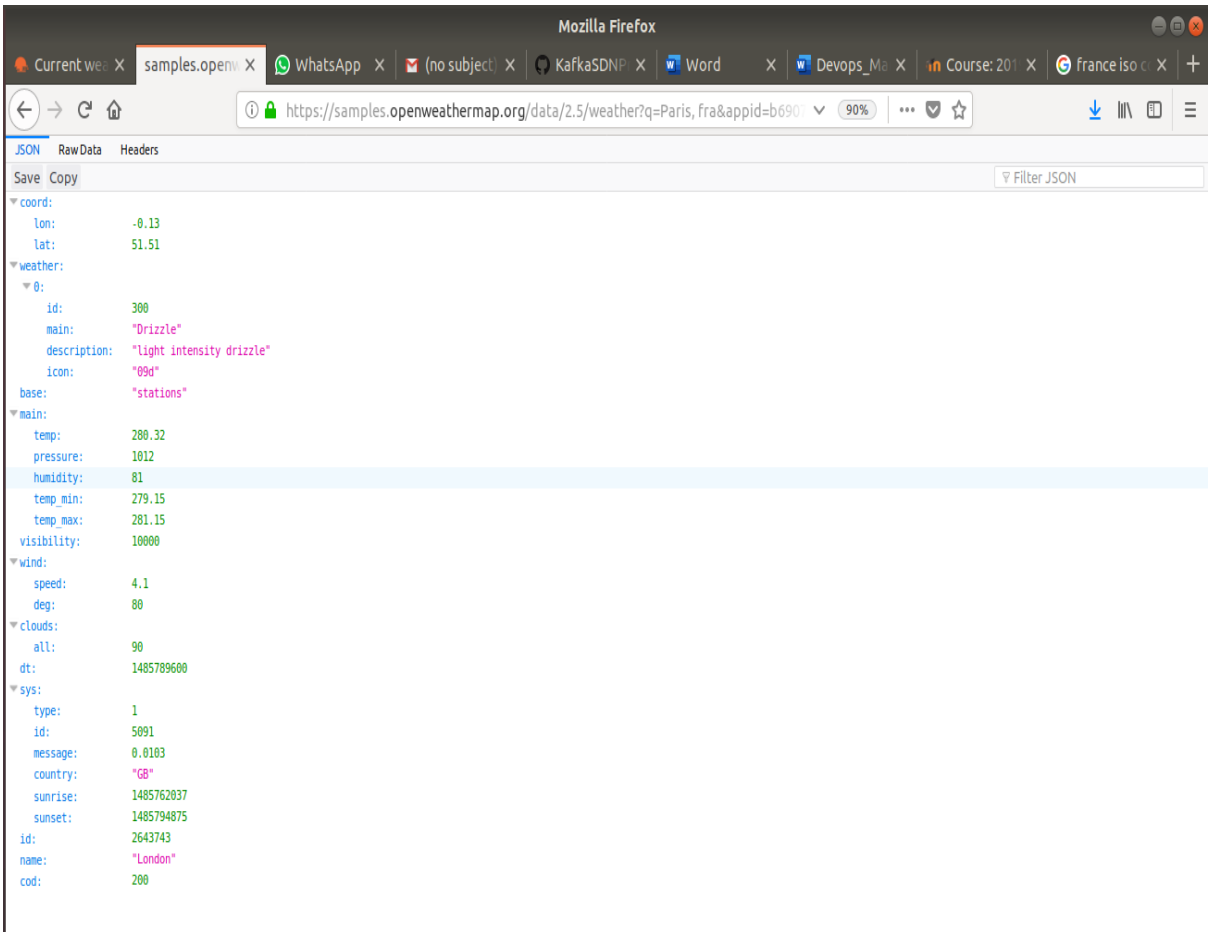
Kafka topic is a named stream of records. Kafka stores topics in logs. A topic log is broken up into partitions. Kafka spreads log's partitions across multiple servers or disks.

3.0 Data :

We are collecting weather data from https://openweathermap.org/current#name. Sample output from this website when we put London as a city is provided below.

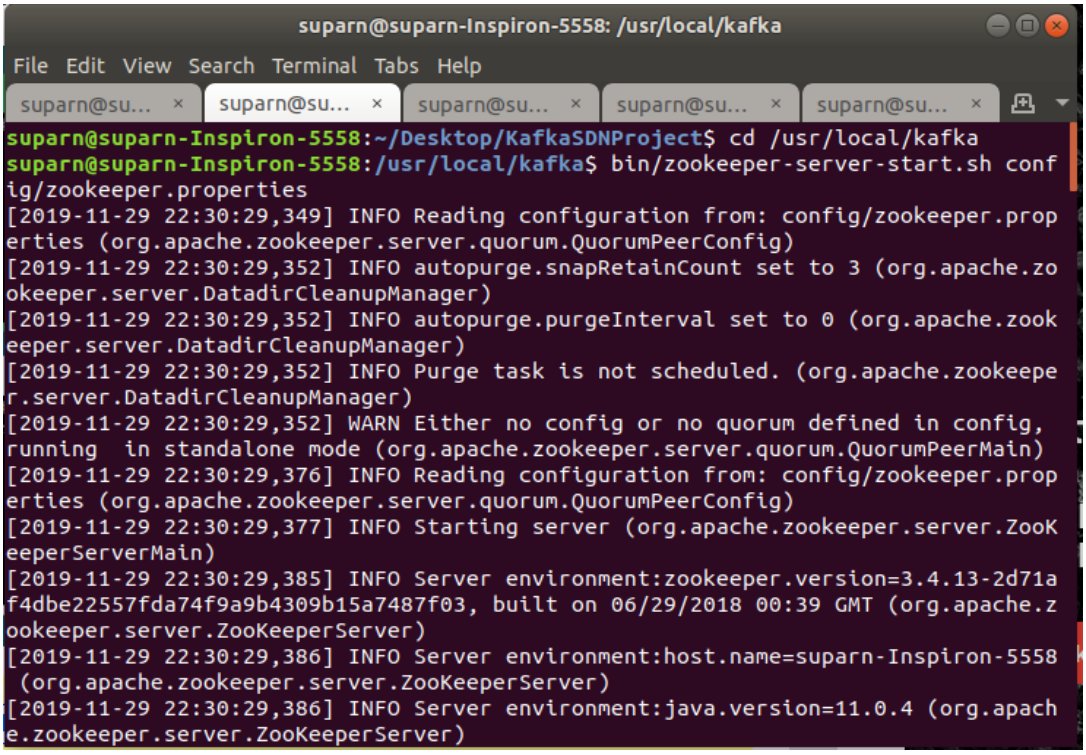
Link :-

https://samples.openweathermap.org/data/2.5/weather?q=London,uk&appid=b6907d289e10d714a6e88b30761fae22

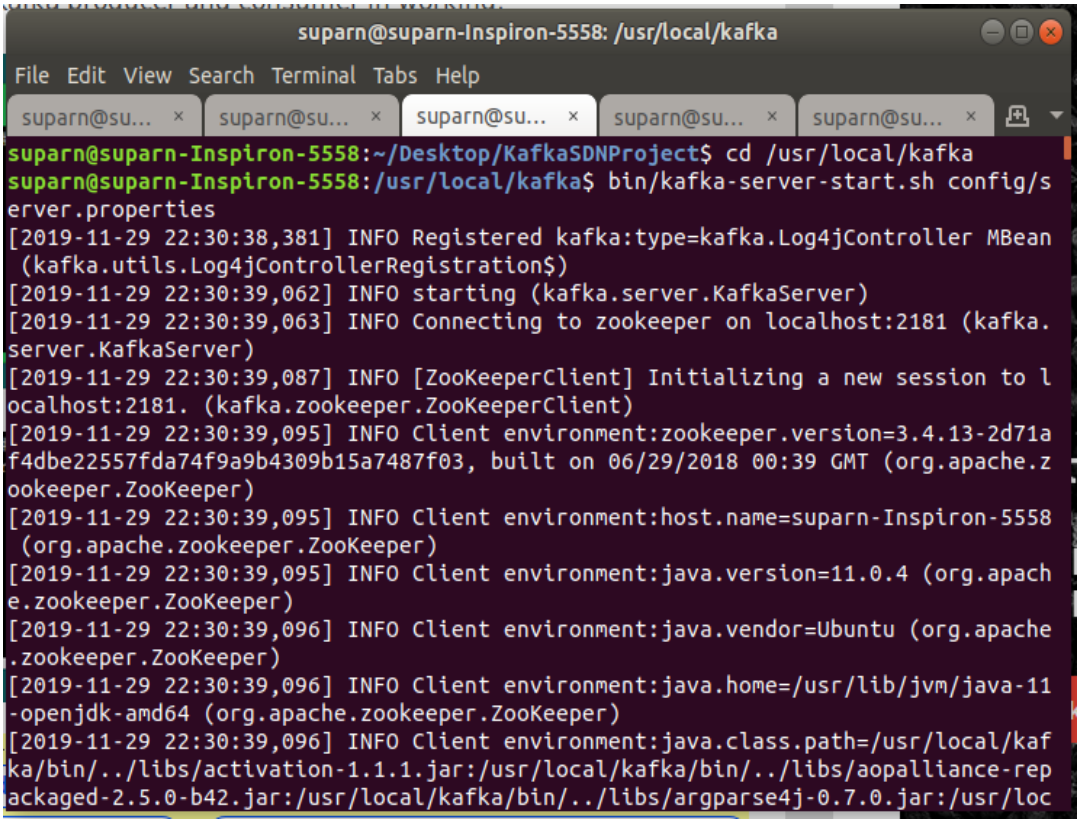


4.0 Screenshots of Steps :

Zoo-Keeper Server



Kafka Server


suparn@suparn-Inspiron-5558: /usr/local/kafka
File Edit View Search Terminal Tabs Help
suparn@su...
suparn@su...
suparn@su...
suparn@su...
suparn@su...
suparn@suparn-Inspiron-5558:~/Desktop/KafkaSDNProject$ cd /usr/local/kafka
suparn@suparn-Inspiron-5558:/usr/local/kafka$ bin/kafka-server-start.sh config/server.properties
[2019-11-29 22:30:38,381] INFO Registered kafka:type=kafka.Log4jController MBean (kafka.utils.Log4jControllerRegistration$)
[2019-11-29 22:30:39,062] INFO starting (kafka.server.KafkaServer)
[2019-11-29 22:30:39,063] INFO Connecting to zookeeper on localhost:2181 (kafka.server.KafkaServer)
[2019-11-29 22:30:39,087] INFO [ZooKeeperClient] Initializing a new session to localhost:2181. (kafka.zookeeper.ZooKeeperClient)
[2019-11-29 22:30:39,095] INFO Client environment:zookeeper.version=3.4.13-2d71af4dbe22557fda74f9a9b4309b15a7487f03, built on 06/29/2018 00:39 GMT (org.apache.zookeeper.ZooKeeper)
[2019-11-29 22:30:39,095] INFO Client environment:host.name=suparn-Inspiron-5558 (org.apache.zookeeper.ZooKeeper)
[2019-11-29 22:30:39,095] INFO Client environment:java.version=11.0.4 (org.apache.zookeeper.ZooKeeper)
[2019-11-29 22:30:39,096] INFO Client environment:java.vendor=Ubuntu (org.apache.zookeeper.ZooKeeper)
[2019-11-29 22:30:39,096] INFO Client environment:java.home=/usr/lib/jvm/java-11-openjdk-amd64 (org.apache.zookeeper.ZooKeeper)
[2019-11-29 22:30:39,096] INFO Client environment:java.class.path=/usr/local/kafka/bin/../libs/activation-1.1.1.jar:/usr/local/kafka/bin/../libs/aopalliance-repackaged-2.5.0-b42.jar:/usr/local/kafka/bin/../libs/argparse4j-0.7.0.jar:/usr/loc