NiFi and Kafka to get Twitter tweets

# Pre-requisites

A consumer key/secret and access token/secret for your Twitter application. Developer-user account can be created for free on Twitter. Create a simple app and get the keys

**Steps:**

1. Create AWS Cloud9 environment and Upload template.json, emr\_start\_commands.txt, docker\_nifi.txt, and nifi-files folder.
2. Check the current IP address of your AWS client using <https://checkip.amazonaws.com>
3. Execute commands interactively provided in the file *emr\_start\_commands.txt* until you get to the “aws emr wait”. Copy the nifi-files folder as documented in the *emr\_start\_commands.txt*
4. Execute commands provided in the file *docker\_nifi.txt* to install Docker and Nifi. Check that NiFi is up and running by opening the NiFi console on browser. http://{public\_dns}:9080/NiFi/
5. Install and run Kafka stream as below:

wget https://archive.apache.org/dist/kafka/2.7.0/kafka\_2.13-2.7.0.tgz

tar -xzf kafka\_2.13-2.7.0.tgz

cd kafka\_2.13-2.7.0

nohup bin/zookeeper-server-start.sh \

config/zookeeper.properties \

> /tmp/zookeeper.out 2>&1 &

ss -tnlpu | grep :2181

nohup bin/kafka-server-start.sh \

config/server.properties \

> /tmp/kafka.out 2>&1 &

ss -tnlpu | grep :9092

bin/kafka-topics.sh --create \

--bootstrap-server localhost:9092 \

--replication-factor 1 \

--partitions 1 \

--topic tweets

hdfs dfs -ls /user/hadoop/tweets-en

hdfs dfs -cat /user/hadoop/tweets-en/tweet-sample\*.txt

bin/kafka-console-consumer.sh \

--bootstrap-server localhost:9092 \

--topic tweets