

# 2024-06-14

## 오전

hadoop 1.3TB

1. hadoop을 연결할수있는 producer 키기 vm
2. 통신되는지 ping해보기 0.201 ~ 205
3. sudo vim /etc/hosts에서 2번의 ip dns지정 broker1 ~ broker5
4. su kafka
5. 오류날시 ssh-keygen -f "/home/kafka/.ssh/known\_hosts" -R "broker3" 발생하는 모든 known\_hosts에 따라서 에러가난다.
6. ssh로 restart할수있따(꿀팁)
7. kafka-topics.sh --bootstrap-server broker1:9092 --create --topic encore\_mort --partitions 3 --replication-factor 3 이러면 encore가 선언이 안되어있으므로 3번에다가 다시 encore1~encore5 를 dns를 추가한다.
8. kafka-topics.sh --bootstrap-server broker1:9092 --create --topic encore\_mort --partitions 3 --replication-factor 3 생성완료!!

```
kafka@producer:~$ kafka-topics.sh --bootstrap-server broker1:9092 --create --topic encore_mort --partitions 3 --replication-factor 3
WARNING: Due to limitations in metric names, topics with a period ('.') or underscore ('_') could collide. To avoid issues it is best to use either, but not both.
Created topic encore_mort.
```

9. 혹시나 문제가 생기면 sudo systemctl status ufw 방화벽문제일수도있다.
11. kafka-topics.sh --bootstrap-server broker1:9092 --describe --topic encore\_mort

```
kafka@producer:~$ kafka-topics.sh --bootstrap-server broker1:9092 --describe --topic encore_mort
Topic: encore_mort    TopicId: dFdx-3HWT-yfS7qZlvcW-Q PartitionCount: 3      ReplicationFactor: 3    Configs: segment
.bytes=1073741824
    Topic: encore_mort Partition: 0    Leader: 1      Replicas: 1,4,2 Isr: 1,4,2
    Topic: encore_mort Partition: 1    Leader: 2      Replicas: 2,1,3 Isr: 2,1,3
    Topic: encore_mort Partition: 2    Leader: 3      Replicas: 3,2,5 Isr: 3,2,5
```

12. kafka@producer:~\$
13. 주피터를 키기위해 su hadoop 왜 안먹히지???
14. su kafka에서 하기 ls -al /opt/ 권한이 root이다 응용프로그램 관련
15. cd ~ 한다음 sudo mv ./kafka /opt/
16. su hadoop

```
(base) hadoop@producer:/opt$ nohup jupyter notebook --ip=0.0.0.0 &
[1] 4516
nohup: ignoring input and appending output to '/home/hadoop/nohup.out'
: command not found
(base) hadoop@producer:/opt$ |
```

- 17.
18. nohup jupyter notebook --ip=0.0.0.0 &

```
(base) hadoop@producer:/opt$ sudo chown hadoop:hadoop -R ./kafka/
[sudo] password for hadoop:
Sorry, try again.
[sudo] password for hadoop:
(base) hadoop@producer:/opt$ ls -al
total 16
drwxr-xr-x  4 root    root    4096 Jun 14 01:17 .
drwxr-xr-x 21 root    root    4096 Jun  5 00:29 ..
drwx--x--x  4 root    root    4096 May 30 07:21 containerd
drwxr-xr-x  8 hadoop  hadoop  4096 Jun  3 06:27 kafka
19. (base) hadoop@producer:/opt$ |
```

20. sudo chown hadoop:hadoop -R ./kafka/

21. vim ~/.bashrc해서

```
##### zookeeper path #####
ZOOKEEPER_HOME=/opt/kafka
PATH=$PATH:$ZOOKEEPER_HOME/bin
```

```
##### kafka-2.13.0 path #####
KAFKA_HOME=/opt/kafka
PATH=$PATH:$KAFKA_HOME/bin
```

22.

23. source ~/.bashrc하고

24. 잘들어갔는지 확인해보려면 echo \$PATH

25. 하둡계정에서 kafka계정도 접속하려고 하는거임

26. 31기-hadoop-data에 있는걸 /home/hadoop에 filezilla로 옮긴다.

27. hdfs dfs -put ./hadoop\_data/ /mort

28. hdfs에서 만든 파일을 kafka로 보내기 - 2024-06-14.ipynb참고

29. kafka@encore4:/data/kafka-logs/encore\_mort-0에 00...log에 내가 보낸게있다.

30. 파이썬에서 api로 받아와서 kafka에 topic저장 producer가 hadoop으로 보내고 hive로 했을 때 sql쿼리가 돌면 나와야한다

31. 2024-06-14-1.ipynb

32. 2024-06-14-2.ipynb 여기서 csv로 저장해서 hdfs dfs put하기

33. hdfs dfs -put

```
/home/hadoop/hadoop_test1/demo/src/main/java/com/example/20240614_data.csv /mort
주피터에서 작업한것을 하둡에 올리기
```

34. xxx.xxx.0.1로 들어가면



The image shows a login interface for a device labeled "ipTIME A9004M-X2". At the top, there is a header bar with a logo consisting of four squares and the text "ipTIME A9004M-X2". Below the header, there are two input fields: "로그인 이름" (Login Name) and "로그인 암호" (Login Password). The "로그인 이름" field is currently empty and has a cursor. Below the input fields is a large grey button labeled "로그인" (Login). To the right of the "로그인" button are two icons: a yellow gear icon and a green icon representing a network or connection.

35.