



0.1. 安装 Kafka

第一节 简介

Kafka 是一种分布式队列工具,可以与 Flume 结合实现日志的收集和分发。需要借助 Zookeeper 协调工作。

第二节 配置文件

Kafka 有很多的配置可以添加。启动简单的 3 个节点的集群需要配置的文件有:

- `server.properties`, 启动服务的基础配置, 设置`broker.id`和 ZK 集群信息。
- `producer.properties`, 配置 Producer 信息。
- `consumer.properties`, 配置 Consumer 信息。
- `log4j.properties`, 配置日志输出信息。

在下列文件夹中加入下面内容:

`server.properties`

```
# 注意这里每台机器要不一样
broker.id=0
port=9092
log.dirs=/home/zhangyu/kafka/data
```

- `broker.id`, 表示区分 Broker 的 ID, 每台机器都要修改, 要不一样。
- `port`, 表示服务的端口。
- `log.dirs`, 存放 log 日志的目录。需要提前创建。

启动流程:

`start-kafka.sh`

```
# 启动Kafka, 需要先启动ZK服务, 假设已经启动了
bin/kafka-server-start.sh config/server.properties
# 创建主题Topic
bin/kafka-topics.sh --create --zookeeper master:2181 \
--replication-factor 1 \
```

```
--partitions 1 \  
--topic test  
# 查看有哪些主题  
bin/kafka-topics.sh --list \  
--zookeeper master:2181  
# 开始生产数据  
bin/kafka-console-producer.sh \  
--broker-list master:9092 \  
--topic test  
# 新建另一个窗口开始消费数据  
bin/kafka-console-consumer.sh \  
--bootstrap-server master:2181 \  
--topic test \  
--from-beginning
```

第三节 测试和运用

注意，将kafka文件夹直接拷贝到不同机器上后，修改broker.id 的值后，每台机器上都需要启动 Kafka 服务。启动之后就可以新建窗口，进行后面的测试了。

```

Terminal 终端 - zhangyu@bd2a17a1c41a: ~
文件(F) 编辑(E) 视图(V) 终端(T) 标签(A) 帮助(H)

zhangyu@bd2a17a1c41a:~$ cd kafka/
zhangyu@bd2a17a1c41a:~/kafka$ ls
bin  data  LICENSE  NOTICE
config  libs  logs  site-docs
zhangyu@bd2a17a1c41a:~/kafka$ kafka-topics.sh --zoo
keeper master:2181 --create --topic t_test \
> --partitions 5 --replication-factor 2
WARNING: Due to limitations in metric names, topic
s with a period('.') or underscore('_') could co
llide. To avoid issues it is best to use either, b
ut not both.
Created topic t_test.
zhangyu@bd2a17a1c41a:~/kafka$ jps
30018 Nimbus
61876 LogviewerServer
38820 Supervisor
773 QuorumPeerMain
26406 Kafka
43479 UIServer
1676 Jps
zhangyu@bd2a17a1c41a:~/kafka$ kafka-console-produc
er.sh --broker-list master:9092 --topic t_test
>Hell world!
>Nihao
>zhelishi Beijing
>Beijing Huanyingni
>Kafka
>Hadoop
>HCT INFO
>HHH
>spark
>zk
>storm
>

zhangyu@bd2a17a1c41a:~/kafka$ kafka-console-consu
mer.sh --bootstrap-server master:2181 --topic t_
test --from-beginning
^CProcessed a total of 0 messages
zhangyu@bd2a17a1c41a:~/kafka$ kafka-console-consu
mer.sh --bootstrap-server master:9092 --topic t_
test --from-beginning
Nihao
HHH
zhelishi Beijing
Beijing Huanyingni
Kafka
Hell world!
Hadoop
spark
zk
storm

```

可以看到左侧 Producer 产生出的数据直接就被右侧 Consumer 消费了。