

# Flume\_实时分析\_日志收集

2019年3月10日 18:00

## 1. 修改Flume中心服务器配置

增加kafkaSink，将数据实时导入kafka

hadoop01 hadoop02

```
1  #配置Agent
2  a1.sources = r1
3  a1.sinks = k1 k2
4  a1.channels = c1 c2
5
6  #声明Source
7  a1.sources.r1.type = avro
8  a1.sources.r1.bind = 0.0.0.0
9  a1.sources.r1.port = 44444
10
11 #声明sink
12 a1.sinks.k1.type = hdfs
13 a1.sinks.k1.hdfs.path = hdfs://hadoop01:9000/flux/reportTime=%Y-%m-%d
14 a1.sinks.k1.hdfs.rollInterval = 30
15 a1.sinks.k1.hdfs.rollSize = 0
16 a1.sinks.k1.hdfs.rollCount = 0
17 a1.sinks.k1.hdfs.fileType = DataStream
18 a1.sinks.k1.hdfs.timeZone = GMT+8
19
20
21 a1.sinks.k2.type = org.apache.flume.sink.kafka.KafkaSink
22 a1.sinks.k2.topic = fluxTopic
23 a1.sinks.k2.brokerList = hadoop01:9092,hadoop02:9092,hadoop03:9092
24
25 #声明channel
26 a1.channels.c1.type = memory
27 a1.channels.c1.capacity = 1000
28 a1.channels.c1.transactionCapacity = 100
29
30
31 a1.channels.c2.type = memory
32 a1.channels.c2.capacity = 1000
33 a1.channels.c2.transactionCapacity = 100
34
35 #绑定关系
36 a1.sources.r1.channels = c1 c2
    a1.sinks.k1.channel = c1
    a1.sinks.k2.channel = c2
```

## 2. 配置Kafka

创建对应主题：

```
kafka-topics.sh --create --zookeeper localhost:2181 --replication-factor 1 --partitions 1 --topic fluxTopic
```

启动消费者测试：

```
kafka-console-consumer.sh --zookeeper localhost:2181 --from-beginning --topic fluxTopic
```