*#agent1 name***agent1.channels** = **c1  
agent1.sources** = **r1 r2 r11 r21 r3  
agent1.sinks** = **k1 k2 k3 k4***#set channel***agent1.channels.c1.type** = **file***#存储在 Channel 当中的最大 events 数***agent1.channels.c1.capacity** = **1000000***#同时刻从Source 获取，或发送到 Sink 的最大 events 数***agent1.channels.c1.transactionCapacity** = **100***#byteCapacity最大内存所有事物允许总字节数***agent1.channels.c1.byteCapacity** = **800000***#添加或删除一个 event 超时的秒数***agent1.channels.c1.keep-alive** = **60**

**1、动态收集指定文件的日志（类似linux收集tail –f 文件名.log命令动态收集）***#agent1.sources.r2.type = exec  
##agent1.sources.r2.command = tail -f F:\\data\\test\\text0315.txt  
#agent1.sources.r2.command = tail -f /data/test/test/text0315-3.txt  
#agent1.sources.r2.channels = c1  
#agent1.sources.r2.inputCharset=utf-8  
  
#agent1.sources.r21.type = exec  
#agent1.sources.r21.command = tail -f F:\\data\\test11\\f21.log  
#agent1.sources.r21.channels = c1  
#agent1.sources.r21.inputCharset=utf-8  
#  
2、收集 静态目录***agent1.sources.r1.type** = **spooldir  
agent1.sources.r1.spoolDir** = **F:\\data\\logdfs***#agent1.sources.r1.spoolDir = /data/test/logdfs/***agent1.sources.r1.fileHeader** = **true  
agent1.sources.r1.channels** = **c1***##  
#agent1.sources.r11.type = spooldir  
#agent1.sources.r11.spoolDir = F:\\data\\logdfs11  
##agent1.sources.r11.spoolDir = /data/test/logdfs11/  
#agent1.sources.r11.fileHeader = true  
#agent1.sources.r11.channels = c1*

*#正则匹配过滤掉收集.log结尾的日志文件  
#agent1.sources.r11.ignorePattern = ^(.)\*\\.log$  
  
3、收集 动态目录  
#agent1.sources.r3.type = TAILDIR  
#agent1.sources.r3.channels = c1  
#agent1.sources.r3.filegroups = f1  
#agent1.sources.r3.filegroups.f1 = F:\\data\\aaaaa\\.\*log.\*  
#agent1.sources.r3.filegroups.f1 = /data/test/test/.\*log.\*  
  
4、保存到本地目录  
#agent1.sinks.k1.channel = c1  
#agent1.sinks.k1.type = file\_roll  
#agent1.sinks.k1.sink.directory = F:\\data\\myfile-test  
#agent1.sinks.k1.sink.rollInterval = 0*

*5、添加拦截器（以便存放到hdfs日志的目录是以 产生日志的主机名开头的*

*比如存放到目录格式为：/* *AM0V4ZRITMRPZCF /20190419/… ）*

*AM0V4ZRITMRPZCF即为产生日志机器名字*

*#host 主机拦截器插入服务器的ip地址或者主机名，  
#agent将这些内容插入到事件的报头中。事件报头中的key使用hostHeader配置，默认是host  
#agent1.sources.r3.interceptors = i2  
#agent1.sources.r3.interceptors.i2.preserveExisting= true  
#agent1.sources.r3.interceptors.i2.type = host  
#agent1.sources.r3.interceptors.i2.hostHeader =hostname  
#agent1.sources.r3.interceptors.i2.useIP = false  
   
6、收集日志到kafka集群***agent1.sinks.k4.channel** = **c1  
agent1.sinks.k4.type** = **org.apache.flume.sink.kafka.KafkaSink  
agent1.sinks.k4.brokerList** = **hadoop000:9092,hadoop001:9092,hadoop002:9092***#agent1.sinks.k4.brokerList = hadoop000:9092***agent1.sinks.k4.topic** = **topicnewtest1  
agent1.sinks.k4.batchSize** = **5  
agent1.sinks.k4.requiredAcks** = **1**

*7、日志通过轮询策略上传到flume（hostname即为要上传的flume机器名）**## set sink1  
#agent1.sinks.k1.channel = c1  
#agent1.sinks.k1.type = avro  
#agent1.sinks.k1.hostname = s3  
#agent1.sinks.k1.port = 52020  
##set sink2  
##aent1.sinks.k2.channel = c1  
##  
#agent1.sinks.k2.channel = c1  
#agent1.sinks.k2.type = avro  
#agent1.sinks.k2.hostname = s4  
#agent1.sinks.k2.port = 52020  
#  
##set sink2  
#agent1.sinks.k3.channel = c1  
#agent1.sinks.k3.type = avro  
#agent1.sinks.k3.hostname = s5  
#agent1.sinks.k3.port = 52020  
#  
###set gruop  
#agent1.sinkgroups = g1  
#####set sink group  
#agent1.sinkgroups.g1.sinks = k1 k2 k3  
#####set failover  
####agent1.sinkgroups.g1.processor.type = failover  
####agent1.sinkgroups.g1.processor.priority.k1 = 10  
####agent1.sinkgroups.g1.processor.priority.k2 = 5  
####agent1.sinkgroups.g1.processor.maxpenalty = 10000  
###  
####设置负载均衡  
#agent1.sinkgroups.g1.processor.type=load\_balance  
#agent1.sinkgroups.g1.processor.backoff=true  
#agent1.sinkgroups.g1.processor.selector=round\_robin*