Github上源码有误，我已经修正：

代码位于:sparkMLlib模块内的spark.streaming. spark.streaming.DirectKafkaWordCount类

接收两个参数：（brokers列表字符串、topics集合的字符串）

要额外添加spark-streaming-kafka的jar包依赖。

（flume+spark也可以直接集成，也需要加spark-streaming-flume的jar包依赖）

流程：

1.Flume-ng启动(netcat source ——memory channel——kafka sink)：

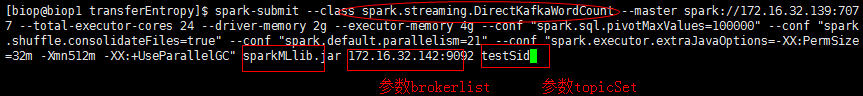
../bin/flume-ng agent --conf conf --conf-file ./myConf/netcatToKafka.properties --name producer -Dflume.root.logger



2.telnet启动，启动参数，一定要和./myConf/netcatToKafka.properties里的绑定好的ip:port一致



3.提交spark-submit



效果图如下：

稍微有点问题：

是wordCount的key有点问题，总是最后一个显示不完整，需要改进

