注意问题：

在Receiver的方式中，Spark中的partition和kafka中的partition并不是相关的，所以如果我们加大每个topic的partition数量，仅仅是增加线程来处理由单一Receiver消费的主题。但是这并没有增加Spark在处理数据上的并行度。

对于不同的Group和topic我们可以使用多个Receiver创建不同的Dstream来并行接收数据，之后可以利用union来统一成一个Dstream。

如果我们启用了Write Ahead Logs复制到文件系统如HDFS，那么storage level需要设置成 StorageLevel.MEMORY\_AND\_DISK\_SER，也就是KafkaUtils.createStream(..., StorageLevel.MEMORY\_AND\_DISK\_SER)