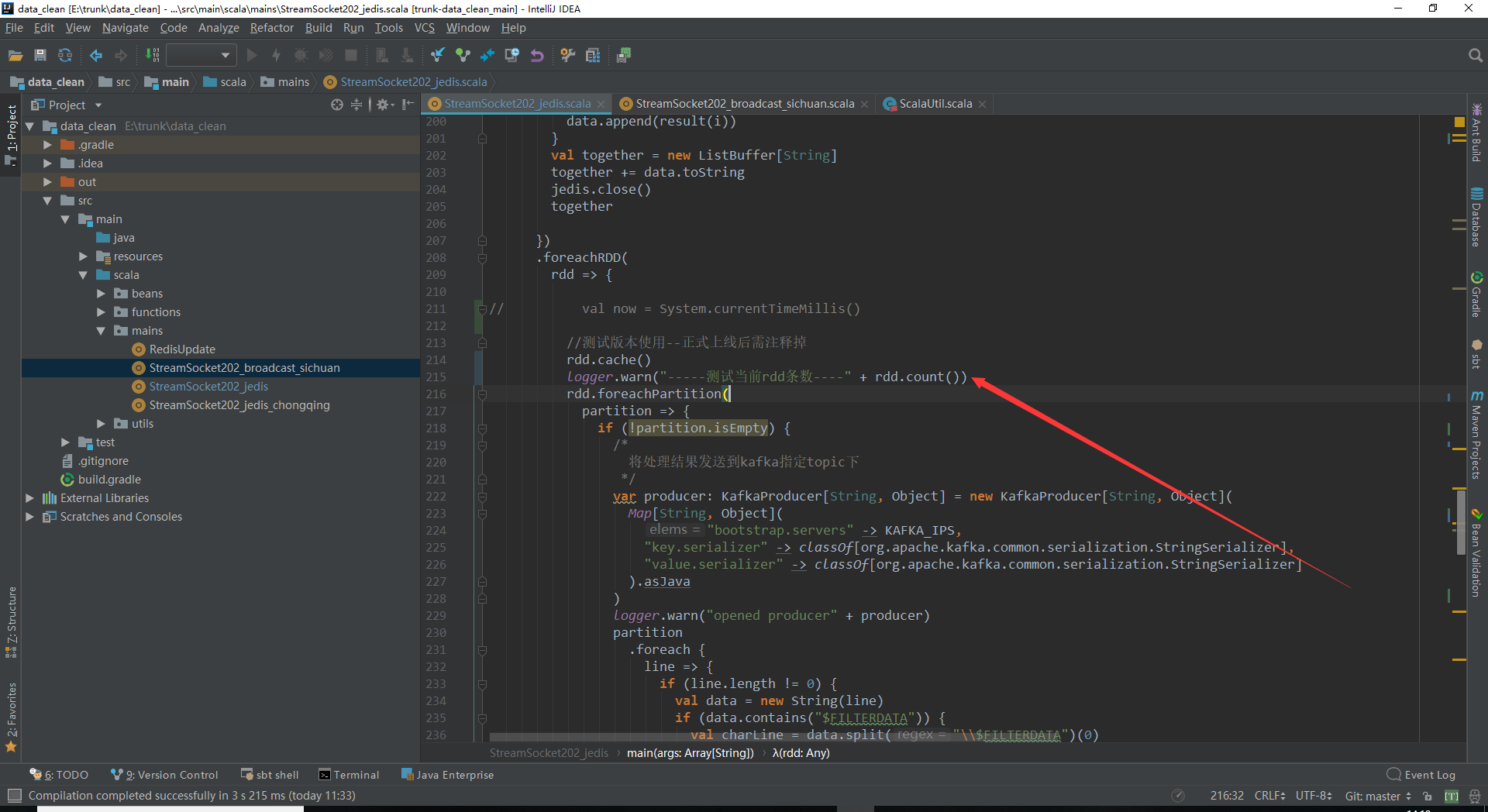
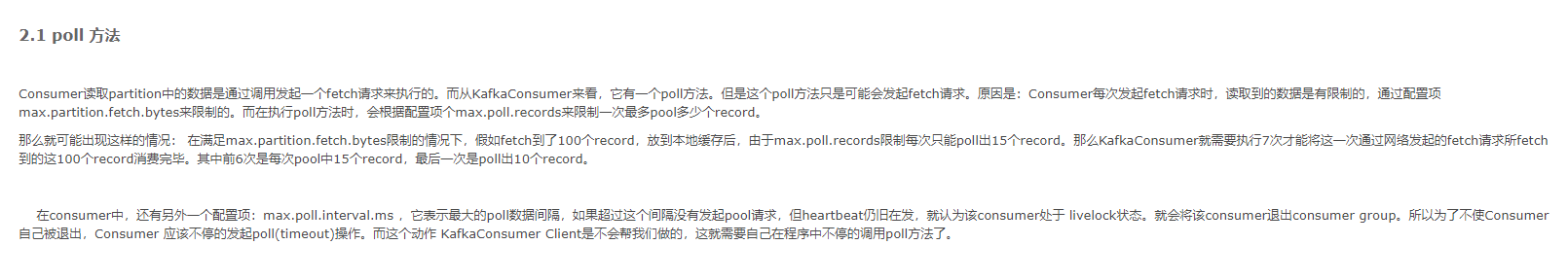
福键延迟问题, 数据一直有延迟

1. 代码中打开注释, 以便在日志中查看spark Streaming中一个批次的数据需要处理多久

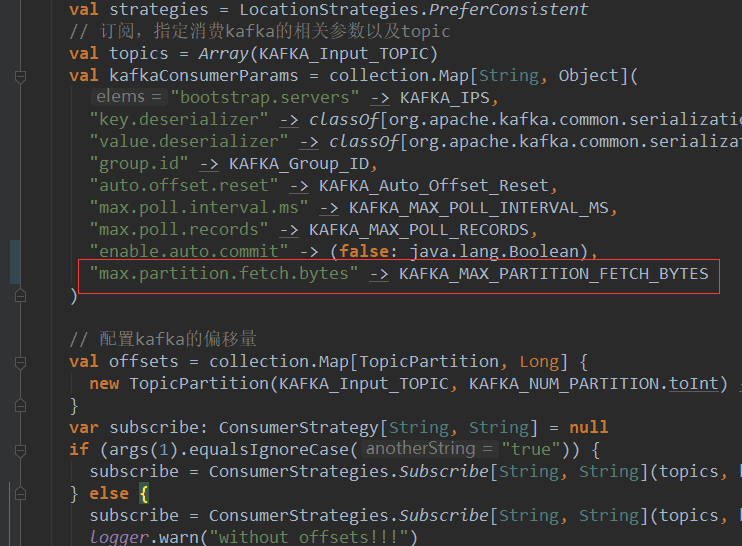


1. 出现问题时可以看到, 程序的窗口大小设置的8s, 但是每个批次处理的时间为11s左右, 在调整poll.max和submit申请的资源后, 无果, 百度相关资料, 可知poll之前还有个fetch操作

参考链接:<https://www.cnblogs.com/rainwang/p/7493742.html>



再结合官网给出的consumer默认配置, 可知, 每个patition的最大fetch值默认1M, poll默认最大500条, poll频率默认512ms, 打开代码注释再修改配置文件, 重启程序, 没有再延迟





Kafka消费者配置官方文档:

<http://kafka.apache.org/documentation/#consumerconfigs>