1.Spark调优的几种方法？

2.Jvm调优了解哪些？

3.JVM内存中存储内存和执行内存保存哪些数据，如果有堆外内存，存储内存还有嘛？·

3.开发中遇到过哪些问题，如何解决的?

4.SparkStreaming如何获取数据的，两种方式有哪些区别？

5.SparkStreaming拉取数据具体调用了哪个api方法，多长时间拉取一次数据。

5.Driver如何与Executer进行联系的。

6.页面单跳转化率是如何实现的，如何进行优化？

7.处理超大型数据如何提高系统效率，如何调优？

8.Spark读取数据生成RDD分区默认多少？

9.快速排序和冒泡排序的时间复杂度是多少，哪个更快？

10.网络编程方面了解哪些，服务器线程方面了解哪些？

11.NIO是否了解?

12.hashMap如何保证线程安全？

13.hive是否了解，SQL语句写的如何？

14.生产者消费者的代码写一下（java方面的生产者消费者案例）

15.spark源码看过哪些？（答：任务提交流程）其它的源码呢？