阿里大数据视频面挂

视频面了50分钟左右

主要是围绕针对做的项目问我的，记不太清

1.spark和hadoop的区别

2.为什么选用spark on yarn这种运行模式。

3.hadoop ha架构

4.hadoop ha当一个namenode挂掉。会有数据丢失吗。如果有，有什么解决方法。

5.spark三种运行模式

6.sql,选出每个班级每个科目最高成绩的那个学生

7.一个有序数组，最优方式找出想要的数

还有好多记不太清了，看到有个同学求大数据面经就随便写了点，8.3号视频面试的