如何查询主键冲突？

答：这个是异常捕获的。

如何确保抽取数据的准确性？

Update不常用该用什么好？

把目标数据查询出来放入一张表中，再进行更新，把原表进行相关处理，这张表重命名为原表。

大量数据插入该怎样？

使用循环插入。

你做的是离线数据还是实时数据？

离线数据。

Spark 是处理什么的？

你负责负责哪一快业务？

看上级领导安排的。

Hive 的工作原理？

把sql装换成MAOREDUCE程序执行。

大数据你用了哪些东西？你是如何操作的？

使用了hadoop、hive、xftp、Linux、azbaban、sqoop：

用hive对Hadoop的数据进行查询计算，用xftp对数据的调取，用azbaban对数据的定时调度，用sqoop对数据的抽取。

文件形式的数据你做过没有？

没有。

你在工作中的上下关系？

上级是组长对我们进行管理，然后我们这一块就是一个团队。

你对实时数据的理解？