

## Hadoop面试题系列(8/11)

## 一. 问答题

- 1.hive如何调优?
- 2.hive如何权限控制?
- 3.hbase写数据的原理是什么?
- 4.hive能像关系数据库那样,建多个库吗?
- 5.hbase宕机如何处理?
- 6.hive实现统计的查询语句是什么?
- 7.生产环境中为什么建议使用外部表?
- 8.hadoop mapreduce创建类DataWritable的作用是什么?
- 9.为什么创建类DataWritable?

## 二. 思考题

- 1.假设公司要建一个数据中心,你会如何规划?
- 2.用Hadoop分析海量日志文件,每行日志记录了如下数据: TableName(表名), Time(时间), User(用户), TimeSpan(时间开销)。

## 要求:

编写MapReduce程序算出高峰时间段(如上午10 点)哪张表被访问的最频繁,以及这段时间访问这张表最多的用户,以及这个用户的总时间开销。

本博客文章除特别声明,全部都是原创! 原创文章版权归过往记忆大数据(<u>过往记忆</u>)所有,未经许可不得转载。 本文链接:【】()