

Hadoop面试题系列(11/11)

1、Hive内部表和外部表的区别？

1、在导入数据到外部表，数据并没有移动到自己的数据仓库目录下，也就是说外部表中的数据并不是由它自己来管理的！而表则不一样；

2、在删除表的时候，Hive将会把属于表的元数据和数据全部删掉；而删除外部表的时候，Hive仅仅删除外部表的元数据，数据是不会删除的！

那么，应该如何选择使用哪种表呢？在大多数情况没有太多的区别，因此选择只是个人喜好的问题。但是作为一个经验，如果所有处理都需要由Hive完成，那么你应该创建表，否则使用外部表！

[《Hive表与外部表》](#)

2、Hbase的rowkey怎么创建比较好？列族怎么创建比较好？

3、用mapreduce怎么处理数据倾斜问题？

4、hadoop框架中怎么来优化？

5、Hbase内部是什么机制？

6、hdfs的数据压缩算法

7、hive底层与数据库交互原理

8、hbase过滤器实现原则

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（[过往记忆](#)）所有，未经许可不得转载。

本文链接: [【】](#) ()