

## HBase的大合并、小合并是什么？

参考答案：

删除一条记录，就会在该记录上打上标记，被打上标记的记录就成了墓碑记录，该记录使用get和scan查询不到，但还是在HFile中。只有进行大合并的时候才会删除HFile中的墓碑记录。

大合并和小合并其目的都是为了优化查询性能和减少存储空间。

**大合并：**指将多个 HFile 合并成一个更大的 HFile，通常会在后台运行，定期或者手动触发。大合并可以清理已经被删除或者过期的数据，并且可以减少 HFile 的数量，提高查询效率。

**小合并：**指将一个 Region 中的一部分 HFile 合并成一个更大的 HFile。小合并可以避免出现太多小的 HFile，减少查询时需要扫描的文件数量，提高查询效率。小合并是自动触发的，当一个 Region 中存储的 HFile 数量达到一定值时，就会触发小合并。

欢迎加入知识星球，获取《大数据面试题 V4.0》以及更多大数据开发内容



知识星球  
长按扫码领取优惠 ▶

