本文基于mysql5.7

目录

·背景

·filesort的两种模式

·外部排序

·mysql的外部排序

·mysql使用filesort时的优化手段

**背景**

在日常开发中，会遇到order by的字段无法使用索引的情况，可能是数据量小不用刻意使用索引优化order by，也可能是业务逻辑复杂导致无法使用索引去优化order by。此时mysql会使用filesort对查询出的结果进行排序。

我们通过explain查看MySQL执行计划时，经常会看到在Extra列中显示Using filesort。

除order by外，Using filesort还经常出现在group by、distinct、join等情况下。

**filesort的两种模式**

1.回表排序模式

a.根据过滤条件获取到查询的数据，获取排序字段值及rowID(行指针)作为键值对。

b.排序。

c.排序完成后进行回表操作，根据rowID读取用户需要返回的数据(字段)。

Tips：由于排序后的rowID不是顺序的，会导致随机IO，Mysql引入了read\_rnd\_buffer来进行优化。Mysql在回表时，会读取一批rowID到read\_rnd\_buffer中并进行排序，再通过rowID读取数据，此时为顺序IO。可以通过read\_rnd\_buffer\_size来配置read\_rnd\_buffer的大小。

2.不回表排序模式

a. 根据过滤条件获取到查询的数据，获取用户需要返回的数据及排序字段作为键值对。

b.排序

c.排序后直接返回，而不用回表查询。

不回表排序模式是典型的空间换时间的方法，但是如果用户查询的数据非常大的话，会花费很多时间在磁盘外部排序上，导致更多IO操作，效率可能还不如回表排序模式。

因此mysql提供了max\_length\_for\_sort\_data参数，当“排序的键值对大小” > max\_length\_for\_sort\_data时，MySQL认为磁盘外部排序的IO效率不如回表的效率，会选择回表排序模式，否则选择不回表排序模式。

mysql使用快速排序算法对数据进行排序，但当需要排序的数据非常多，超过了sort buffer的大小时，则需要使用外部排序。

**外部排序**

1.普通外部排序:

https://blog.csdn.net/joy0921/article/details/80125563