

图文 27 磁盘文件中，40个bit位的数据头以及真实数据是如何存储的？

手机观看

371 人次阅读 2020-02-21 07:00:00

详情 评论

磁盘文件中，40个bit位的数据头以及真实数据是如何存储的？

如何提问：每篇文章都有评论区，大家可以尽情留言提问，我会逐一答疑

如何加群：购买狸猫技术窝专栏的小伙伴都可以加入狸猫技术交流群，一个非常纯粹的技术交流的地方

具体加群方式，请参见目录菜单下的文档：《MySQL专栏付费用户如何加群》（购买后可见）

之前我们已经给大家讲到了在磁盘上存储数据的时候，每一行数据都会有变长字段长度列表，逆序存放这行数据里的变长字段的长度，然后会有NULL值列表，对于允许为NULL的字段都会有一个bit位标识那个字段是否为NULL，也是逆序排列的。

今天我们接着给大家讲每一行数据存储的时候，还得有40个bit位的数据头，这个数据头是用来描述这行数据的。

这40个bit位里，第一个bit位和第二个bit位，都是预留位，是没任何含义的。

然后接下来有一个bit位是**delete_mask**，他标识的是这行数据是否被删除了，其实看到这个bit位，很多人可能已经反映过来了，这么说在MySQL里删除一行数据的时候，未必是立马把他从磁盘上清理掉，而是给他在数据头里搞1个bit

标记他已经被删了？

没错，其实大家现在看这些数据头，只要先留有一个印象就可以了，知道每一行数据都有一些数据头，不同的数据头都是用来描述这行数据的一些状态和附加信息的。

然后下一个bit位是**min_rec_mask**，这个bit位大家现在先不用去关注，他的含义以后我们讲到对应的内容的时候再说，他其实就是说在B+树里每一层的非叶子节点里的最小值都有这个标记。

接下来有4个bit位是**n_owned**，这个暂时我们也先不用去管他，他其实就是记录了一个记录数，这个记录数的作用，后续我们讲到对应的概念时会告诉大家的。

接着有13个bit位是**heap_no**，他代表的是当前这行数据在记录堆里的位置，现在大家可能也很难去理解他，这些概念都要结合后续的一些内容才能理解的，这里只能是初步的给大家介绍下。

然后是3个bit位的**record_type**，这就是说这行数据的类型

0代表的是普通类型，1代表的是B+树非叶子节点，2代表的是最小值数据，3代表的是最大值数据

很多朋友可能也不理解这些什么意思，其实我们也现在不用在乎他，因为很多这些概念都是往后在讲解索引之类的技术的时候才会涉及到的。

最后是16个bit的**next_record**，这个是指向他下一条数据的指针。

今天是周五，我们就讲这些，大家利用周末复习一下我们本周讲的这些内容，务必消化吸收，下周咱们继续。

End

专栏版权归公众号**狸猫技术窝**所有

未经许可不得传播，如有侵权将追究法律责任

狸猫技术窝精品专栏及课程推荐：

[《从零开始带你成为消息中间件实战高手》](#)

[《21天互联网Java进阶面试训练营》（分布式篇）](#)

[《互联网Java工程师面试突击》（第1季）](#)

[《互联网Java工程师面试突击》（第3季）](#)

[《从零开始带你成为JVM实战高手》](#)

Copyright © 2015-2020 深圳小鹅网络技术有限公司 All Rights Reserved. [粤ICP备15020529号](#)



小鹅通提供技术支持