

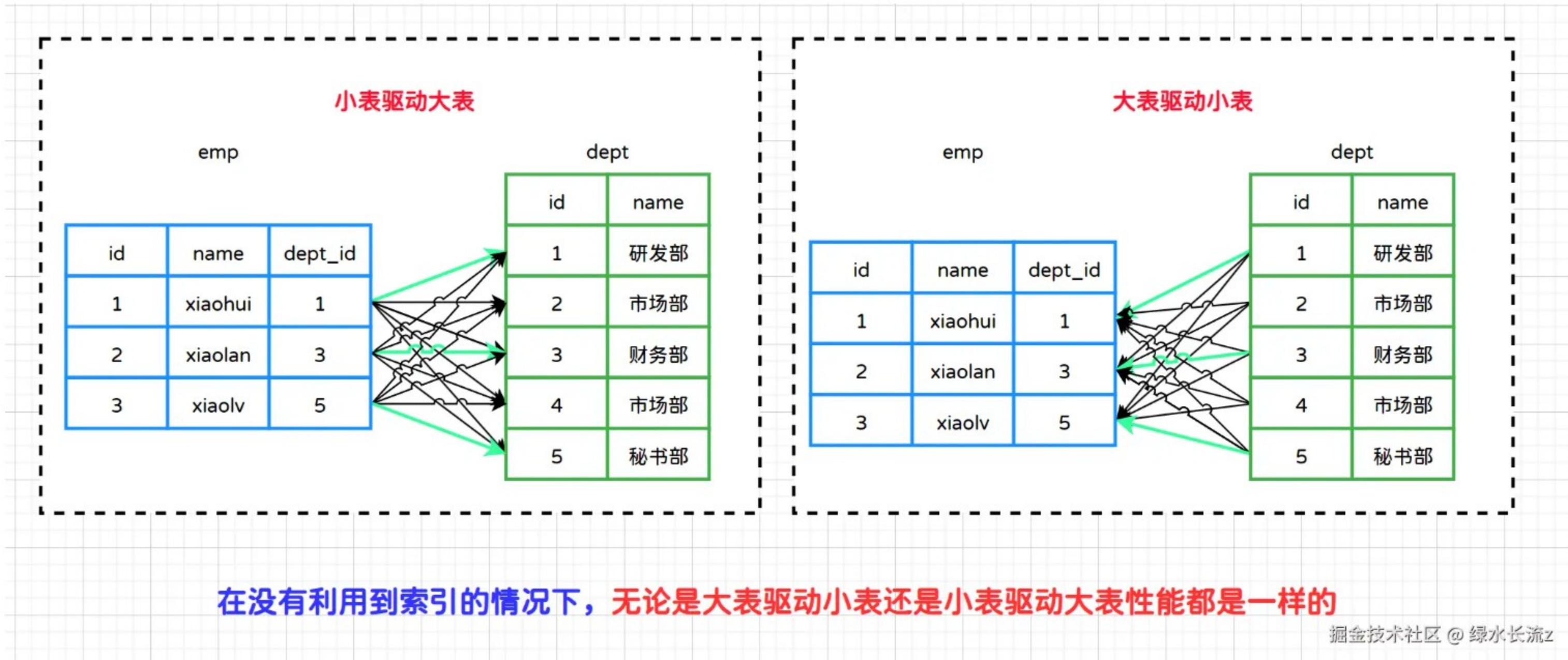
为何要小表驱动大表？

绿水长流z 2025-01-16 2,327 阅读2分钟 专栏：数据库

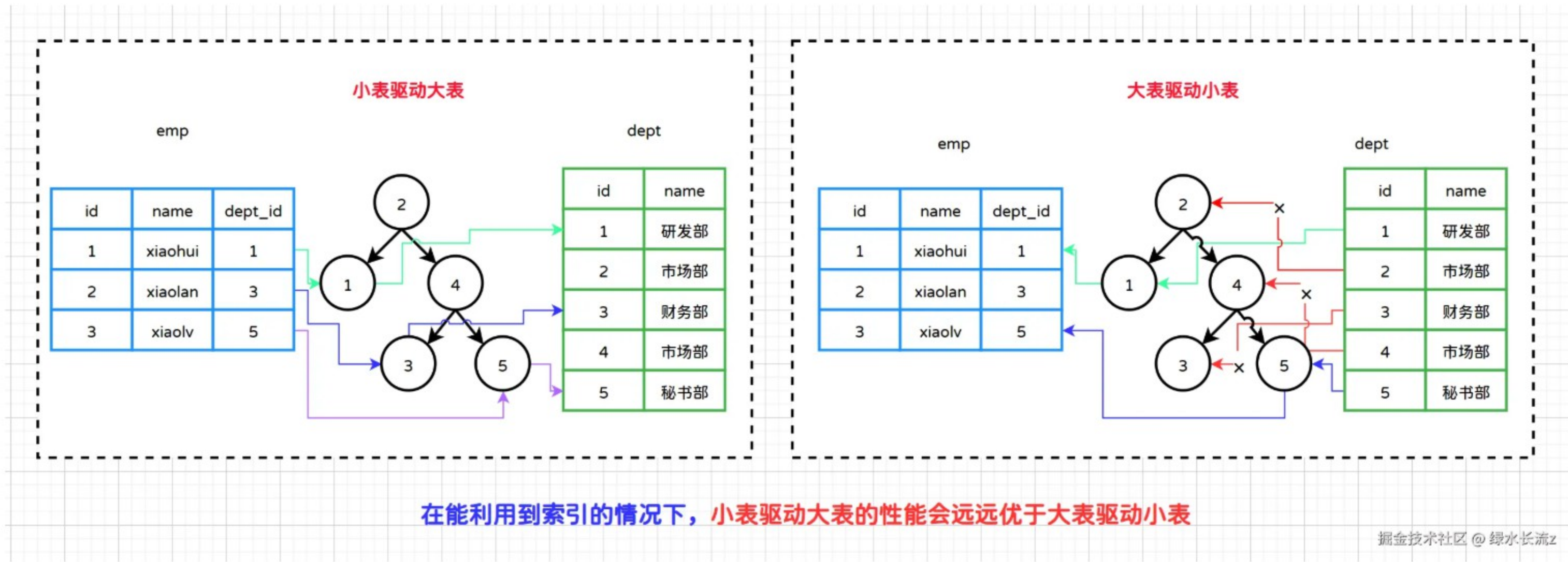
为什么要小表驱动大表？

MySQL在执行Join操作时，优先使用较小的表作为驱动表（也称为外层表）去连接较大的表（也称为被驱动表或内层表）。这样做的原因是，对于连接操作，通常需要对驱动表进行全表扫描或根据索引进行查找，然后根据连接条件与被驱动表进行匹配。

- 没有索引的情况下：



- 有用到索引的情况：



因此，小表驱动大表速度快的前提是：两个表上根据主/外键建立了索引，这样在根据某一条数据查找B+树时，速度就会大大提高，若没有建立索引，则两个表无论谁当作主表，查找数据的次数都是一样的。

如果先对小表进行操作，那么总的扫描行数 and 匹配次数会相对较少，从而提高查询的效率。例如，假设有表 A（小表）和表 B（大表）进行连接，如果以表 A 作为驱动表，那么遍历表 A 的成本相对较低，然后根据连接条件与表 B 进行匹配，效率会更高。

假设我们有两个表 `employees` 和 `departments`，其中 `employees` 表包含大量的员工信息，而 `departments` 表只包含少量的部门信息。我们要找出每个员工所在的部门名称。

- 大表驱动小表伪代码：

```
java
1 FOR each row e IN employees // 假设employees有100万行
2   FOR each row d IN departments // 假设departments有100行
3     IF e.department_id == d.id THEN
4       OUTPUT (e.name, d.name)
```

这里，外层循环会执行 100 万次，每次都要遍历 `departments` 表中的 100 行，总共需要 100 万 * 100 = 1 亿次比较。

- 小表驱动大表伪代码：

```
java
1 FOR each row d IN departments // 只有100行
2   LOOKUP rows FROM employees USING idx_department WHERE department_id = d.id // 使用索引查找
3   FOR each matching row e IN lookup_result
4     OUTPUT (e.name, d.name)
```

这里，外层循环只执行 100 次，每次遍历 `employees` 表中的 100 万行。虽然内层循环的次数看起来很多，但如果 `employees.department_id` 上有索引，每次查找都可以非常快，且外层循环的次数大大减少。

标签：Java 架构 后端 话题：每天一个知识点

 绿水长流z LV.3
Java技术经理

17
文章

44k
阅读

36
粉丝

关注

私信

目录 收起

[为什么要小表驱动大表？](#)

- 相关推荐
- MySQL的驱动表与被驱动表

249阅读 · 4点赞
- MySQL关联查询时，我们为什么建议小...

1.3k阅读 · 0点赞
- 银行家算法：确保数据库事务无死锁经...

292阅读 · 6点赞
- MySQL 语法树解析：深入理解数据库查...

76阅读 · 1点赞
- 若依自定义权限控制

772阅读 · 2点赞

- 精选内容
- 每个程序员都应该知道的计算常识

LeoLi_4 · 81阅读 · 0点赞
- 重新认识一下redis的lua脚本和事务

提前退休的java猿 · 2.1k阅读 · 11点赞
- Laravel11 博客2--集成前端HTML模板

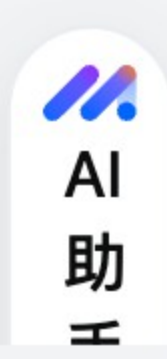
一个人的程序 · 31阅读 · 1点赞
- Java 解决并发问题方案及代码案例：并...

装睡鹿先生 · 71阅读 · 2点赞
- 使用VSCode进行Java开发的指南与优...

世界哪有真情 · 174阅读 · 2点赞

找对属于你的技术圈子

回复「进群」加入官方微信群



本文收录于以下专栏



数据库

专栏目录

记录数据库开发

3 订阅 · 9 篇文章

订阅

上一篇 深入分析MVCC多版本并发控制

下一篇 图解深入剖析数据库的Join算法

评论 6



登录 / 注册 即可发布评论!

最热 | 最新



小码编匠 后端工程师（DotNet技术匠）

不错

3天前 1 点赞 0 评论

...



EnigmaGcl 后端工程师

如果被驱动表列有索引，走 Nest Loop Join，就是题主说的有用到索引的情况，时间复杂度 Mlog2(N)，没有就走 Blocked Nest Loop Join，对应题主说的没有用到索引的情况，时间复杂度 M*N，因此可以看到，有索引，如果小表 1w 行，大表 100w 行，使用 Nest Loop Join，时间复杂度为 1w*log2(100w)，如果没有索引使用 BNL，时间复杂度为 1w*100w，差别非常明显，因此join 的优化方向一般都是在被驱动表上加索引，将 BNL 优化为 NLJ，大幅提升性能，如果不能加索引，那可以考虑将 join buffer 适当调大，确保驱动表能完全放入，或着减少驱动表分批次数，以提升性能...

展开

18天前 1 4 2 评论

...



用户5450406... : 老哥怎么学习这方面的知识，有什么书籍推荐嘛，感谢

10天前 1 点赞 0 回复

...



EnigmaGcl 回复 用户5450406... : 可以参考下我写的文章: juejin.cn 看4.2部分，理解后可以看高性能MySQL这本书

9天前 1 点赞 0 回复

...



大树上的孙大剩

首先要清楚索引在大表上的作用明显，在小表上作用微乎其微

热评

这类问题其实就是把索引作用最大化

小表驱动大表，小表全表扫描导致索引失效，关联的大表索引生效；
大表驱动小表，大表全表扫描索引失效，关联的小表索引生效；

1月前 1 10 1 评论

...



活捉打水怪：这个就是本质,一大群人叭叭叭的讲半天整不清楚,说白了根本没有深入理解索引

9天前 1 点赞 0 回复

...

为你推荐

MySQL表连接算法

程序员小潘 | 2年前 | 1.0k 1 评论

后端

记录一次联表查询优化过程

还有谁_还有谁 | 7月前 | 284 1 评论

后端

MySQL的驱动表与被驱动表

明年就当百万富翁 | 1年前 | 249 4 评论

MySQL

Mysql如何Join一条语句

唐宋xy | 3年前 | 668 1 评论

MySQL

什么，LEFT JOIN 会变成 JOIN?

开坦克的贝吉塔 | 4年前 | 3.5k 5 评论

MySQL

MySQL面试：left join我要怎优化？

不会敲代码了 | 3年前 | 13k 76 19 评论

后端

MySQL

MySQL update语句带子查询的优化探寻

37手游后端团队 | 2月前 | 648 10 1 评论

后端

面试

数据库

MySQL 表连接的原理是什么？如何优化表连接？

掘掘子 | 7月前 | 331 1 评论

后端

MySQL

面试

MySQL JOIN关联查询的原理以及优化手段

刘Java | 2年前 | 1.7k 6 评论

MySQL

Mysql-一些SQL优化的方式

Xiao 斌 | 11月前 | 1.2k | 12 | 8 | 后端 MySQL 面试

mysql优化二记录一下

超级爽朗的郑 | 3年前 | 275 | 4 | 评论 | MySQL

掌握这几个技巧，以后用MySQL查询总比别人快一步！

麒麟改bug | 3年前 | 445 | 4 | 评论 | Java

MySQL Join原理

在下uptown | 1年前 | 1.7k | 7 | 1 | 后端

MySQL之子查询优化

阿布 | 2年前 | 1.4k | 7 | 评论 | 后端

为什么我使用了索引，查询还是慢？

码农开花 | 4年前 | 214 | 1 | 评论 | 分布式