

04讲深入浅出索引（上）



提到数据库索引，我想你并不陌生，在日常工作中会经常接触到。比如某一个SQL查询比较慢，分析完原因之后，你可能就会说“给某个字段加个索引吧”之类的解决方案。但到底什么是索引，索引又是如何工作的呢？今天就让我们一起来聊聊这个话题吧。

数据库索引的内容比较多，我分成了上下两篇文章。索引是数据库系统里面最重要的概念之一，所以我希望你能够耐心看完。在后面的实战文章中，我也会经常引用这两篇文章中提到的知识点，加深你对数据库索引的理解。

一句话简单来说，索引的出现其实就是为了提高数据查询的效率，就像书的目录一样。一本500页的书，如果你想快速找到其中的某一个知识点，在不借助目录的情况下，那我估计你可得找一会儿。同样，对于数据库的表而言，索引其实就是它的“目录”。

索引的常见模型

索引的出现是为了提高查询效率，但是实现索引的方式却有很多种，所以这里也就引入了索引模型的概念。可以用于提高读写效率的数据结构很多，这里我先给你介绍三种常见、也比较简单的数据结构，它们分别是哈希表、有序数组和搜索树。

下面我主要从使用的角度，为你简单分析一下这三种模型的区别。

哈希表是一种以键-值（key-value）存储数据的结构，我们只要输入待查找的值即key，就可以找到其对应的值即Value。哈希的思路很简单，把值放在数组里，用一个哈希函数把key换算成一个确定的位置，然后把value放在数组的这个位置。

不可避免地，多个key值经过哈希函数的换算，会出现同一个值的情况。处理这种情况的一种方法是，拉出一个链表。

假设，你现在维护着一个身份证信息和姓名的表，需要根据身份证号查找对应的名字，这时对应的哈希索引的示意图如下所示：

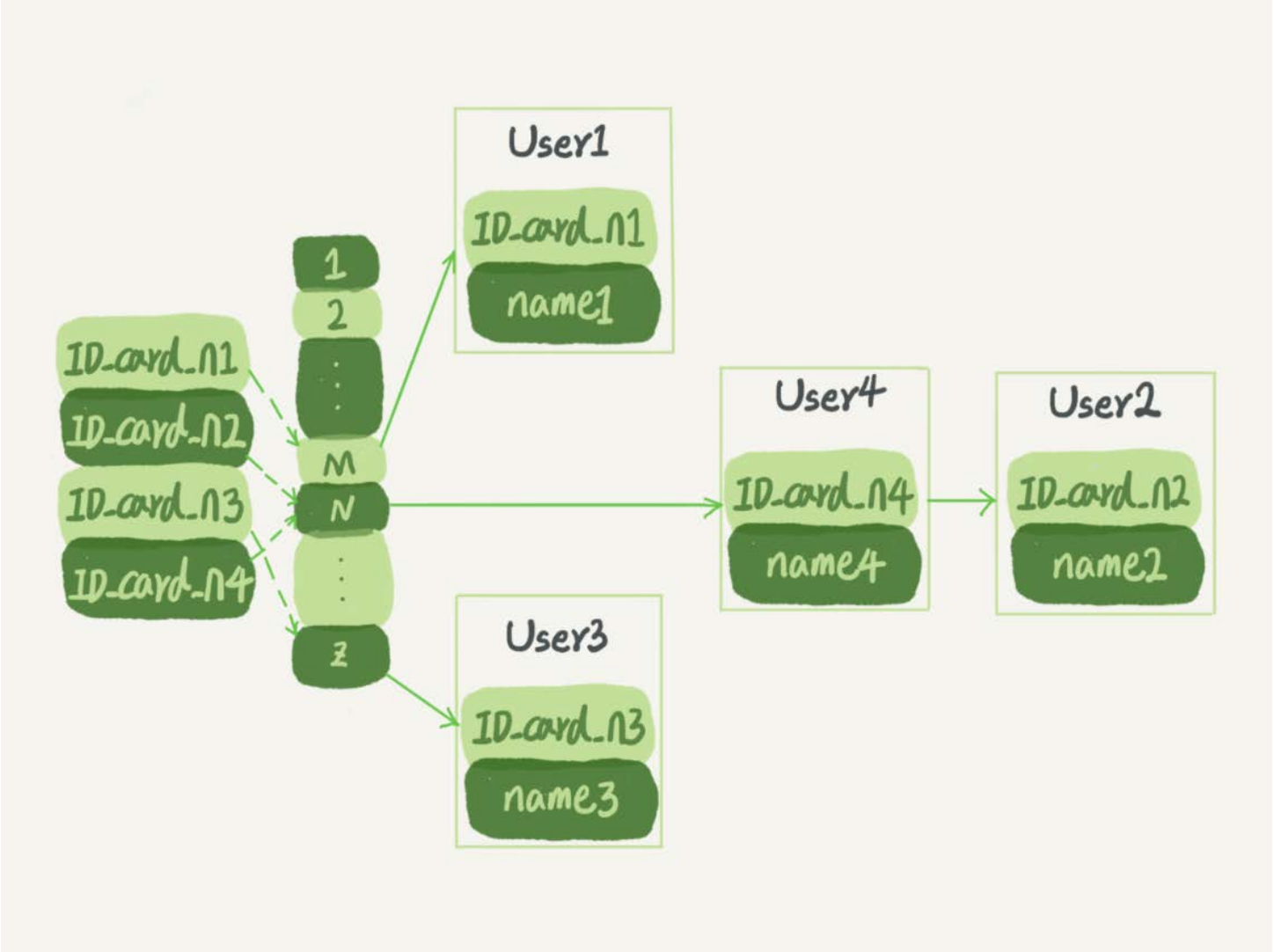


图1 哈希表示意图

图中，User2和User4根据身份证号算出来的值都是N，但没关系，后面还跟了一个链表。假设，这时候你要查ID_card_n2对应的名字是什么，处理步骤就是：首先，将ID_card_n2通过哈希函数算出N；然后，按顺序遍历，找到User2。

需要注意的是，图中四个ID_card_n的值并不是递增的，这样做的好处是增加新的User时速度会很快，只需要往后追加。但缺点是，因为不是有序的，所以哈希索引做区间查询的速度是很慢的。

你可以设想下，如果你现在要找身份证号在[ID_card_X, ID_card_Y]这个区间的所有用户，就必须全部扫描一遍了。

所以，**哈希表这种结构适用于只有等值查询的场景**，比如Memcached及其他一些NoSQL引擎。

而**有序数组在等值查询和范围查询场景中的性能就都非常优秀**。还是上面这个根据身份证号查名字的例子，如果我们使用有序数组来实现的话，示意图如下所示：

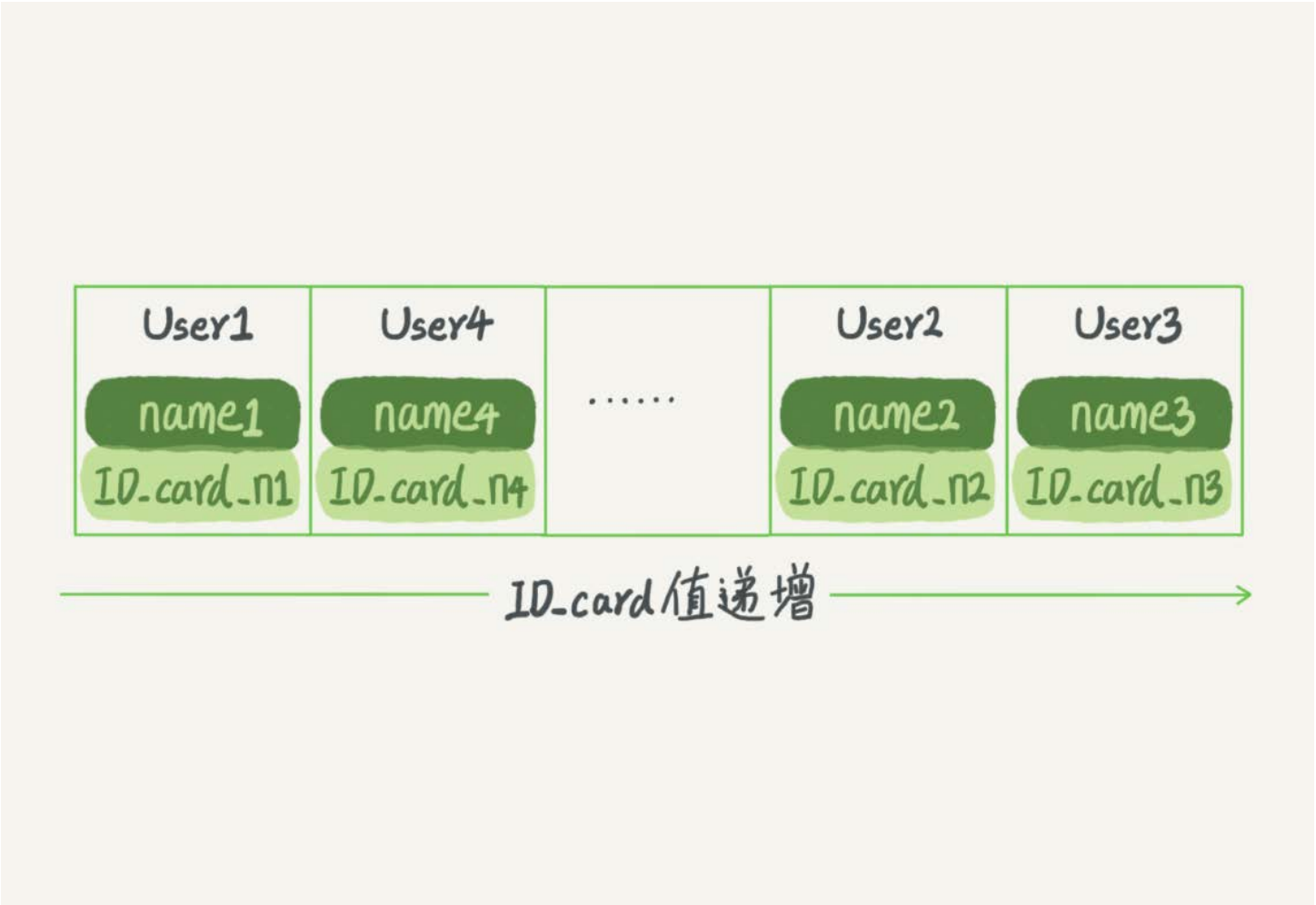


图2 有序数组示意图

这里我们假设身份证号没有重复，这个数组就是按照身份证号递增的顺序保存的。这时候如果你要查ID_card_n2对应的名字，用二分法就可以快速得到，这个时间复杂度是 $O(\log(N))$ 。

同时很显然，这个索引结构支持范围查询。你要查身份证号在[ID_card_X, ID_card_Y]区间的User，可以先用二分法找到ID_card_X（如果不存在ID_card_X，就找到大于ID_card_X的第一个User），然后向右遍历，直到查到第一个大于ID_card_Y的身份证号，退出循环。

如果仅仅看查询效率，有序数组就是最好的数据结构了。但是，在需要更新数据的时候就麻烦了，你往中间插入一个记录就必须得挪动后面所有的记录，成本太高。

所以，**有序数组索引只适用于静态存储引擎**，比如你要保存的是2017年某个城市的所有人口信息，这类不会再修改的数据。

二叉搜索树也是课本里的经典数据结构了。还是上面根据身份证号查名字的例子，如果我们用二叉搜索树来实现的话，示意图如下所示：

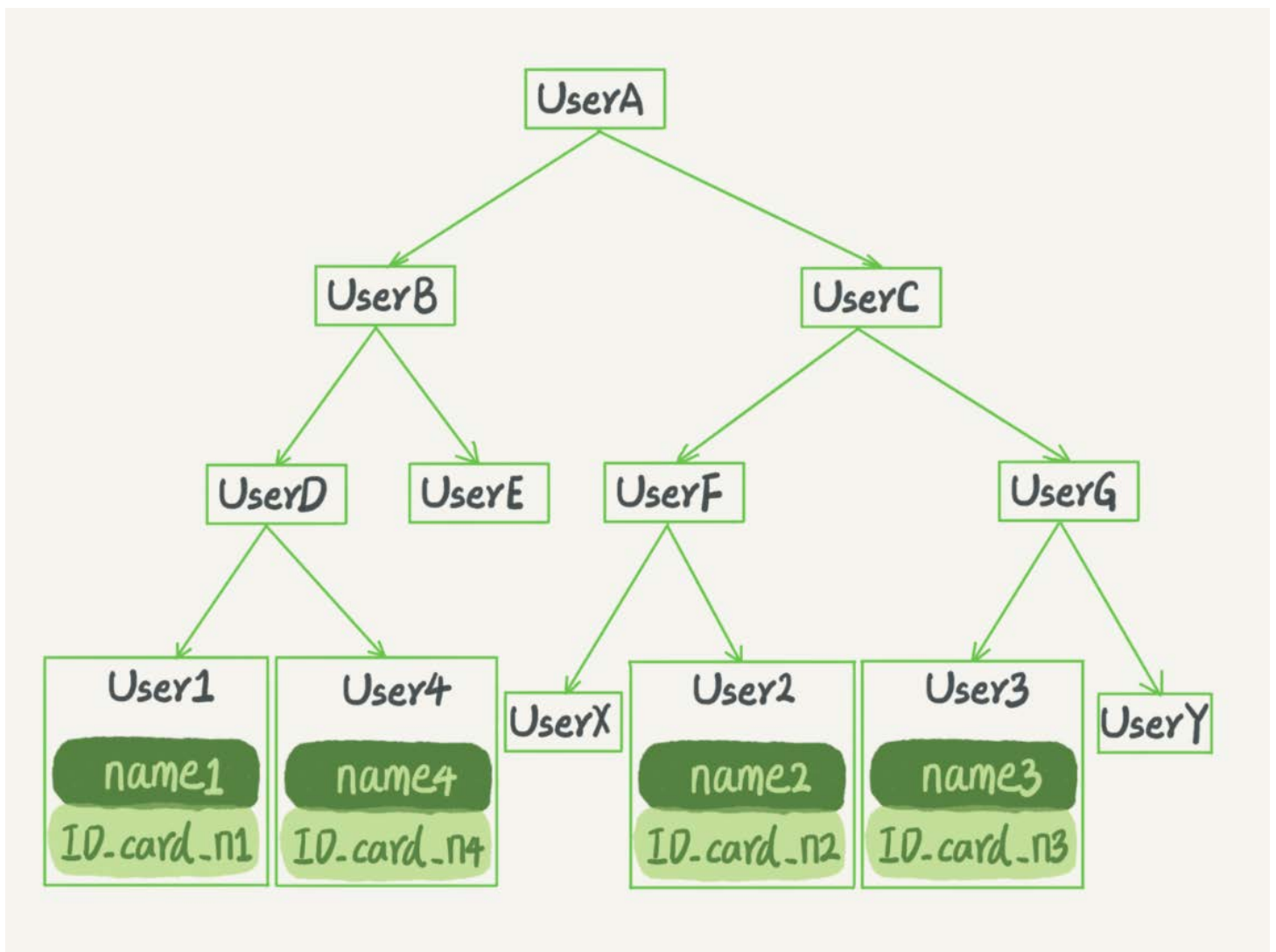


图3 二叉搜索树示意图

二叉搜索树的特点是：每个节点的左儿子小于父节点，父节点又小于右儿子。这样如果你要查ID_card_n2的话，按照图中的搜索顺序就是按照UserA -> UserC -> UserF -> User2这个路径得到。这个时间复杂度是 $O(\log(N))$ 。

当然为了维持 $O(\log(N))$ 的查询复杂度，你就需要保持这棵树是平衡二叉树。为了做这个保证，更新的时间复杂度也是 $O(\log(N))$ 。

树可以有二叉，也可以有多叉。多叉树就是每个节点有多个儿子，儿子之间的大小保证从左到右递增。二叉树是搜索效率最高的，但是实际上大多数的数据库存储却并不使用二叉树。其原因是，索引不止存在内存中，还要写到磁盘上。

你可以想象一下一棵100万节点的平衡二叉树，树高20。一次查询可能需要访问20个数据块。在机械硬盘时代，从磁盘随机读一个数据块需要10 ms左右的寻址时间。也就是说，对于一个100万行的表，如果使用二叉树来存储，单独访问一个行可能需要20个10 ms的时间，这个查询可真够慢的。

为了让一个查询尽量少地读磁盘，就必须让查询过程访问尽量少的数据块。那么，我们就不应该使用二叉树，而是要使用“N叉”树。这里，“N叉”树中的“N”取决于数据块的大小。

以InnoDB的一个整数字段索引为例，这个N差不多是1200。这棵树高是4的时候，就可以存1200的3次方个值，这已经17亿了。考虑到树根的数据块总是在内存中的，一个10亿行的表上一个整数字段的索引，查找一个值最多只需要访问3次磁盘。其实，树的第二层也有很大概率在内存中，那么访问磁盘的平均次数就更少了。

N叉树由于在读写上的性能优点，以及适配磁盘的访问模式，已经被广泛应用在数据库引擎中了。

不管是哈希还是有序数组，或者N叉树，它们都是不断迭代、不断优化的产物或者解决方案。数据库技术发展到现在，跳表、LSM树等数据结构也被用于引擎设计中，这里我就不再一一展开了。

你心里要有个概念，数据库底层存储的核心就是基于这些数据模型的。每碰到一个新数据库，我们需要先关注它的数据模型，这样才能从理论上分析出这个数据库的适用场景。

截止到这里，我用了半篇文章的篇幅和你介绍了不同的数据结构，以及它们的适用场景，你可能会觉得有些枯燥。但是，我建议你还是要多花一些时间来理解这部分内容，毕竟这是数据库处理数据的核心概念之一，在分析问题的时候会经常用到。当你理解了索引的模型后，就会发现在分析问题的时候会有一个更清晰的视角，体会到引擎设计的精妙之处。

现在，我们一起进入相对偏实战的内容吧。

在MySQL中，索引是在存储引擎层实现的，所以并没有统一的索引标准，即不同存储引擎的索引的工作方式并不一样。而即使多个存储引擎支持同一种类型的索引，其底层的实现也可能不同。由于InnoDB存储引擎在MySQL数据库中使用最为广泛，所以下面我就以InnoDB为例，和你分析一下其中的索引模型。

InnoDB 的索引模型

在InnoDB中，表都是根据主键顺序以索引的形式存放的，这种存储方式的表称为索引组织表。又因为前面我们提到的，InnoDB使用了B+树索引模型，所以数据都是存储在B+树中的。

每一个索引在InnoDB里面对应一棵B+树。

假设，我们有一个主键列为ID的表，表中有字段k，并且在k上有索引。

这个表的建表语句是：

```
mysql> create table T(  
  id int primary key,  
  k int not null,  
  name varchar(16),  
  index (k))engine=InnoDB;
```


表中R1~R5的(ID,k)值分别为(100,1)、(200,2)、(300,3)、(500,5)和(600,6)，两棵树的示例示意图如下。

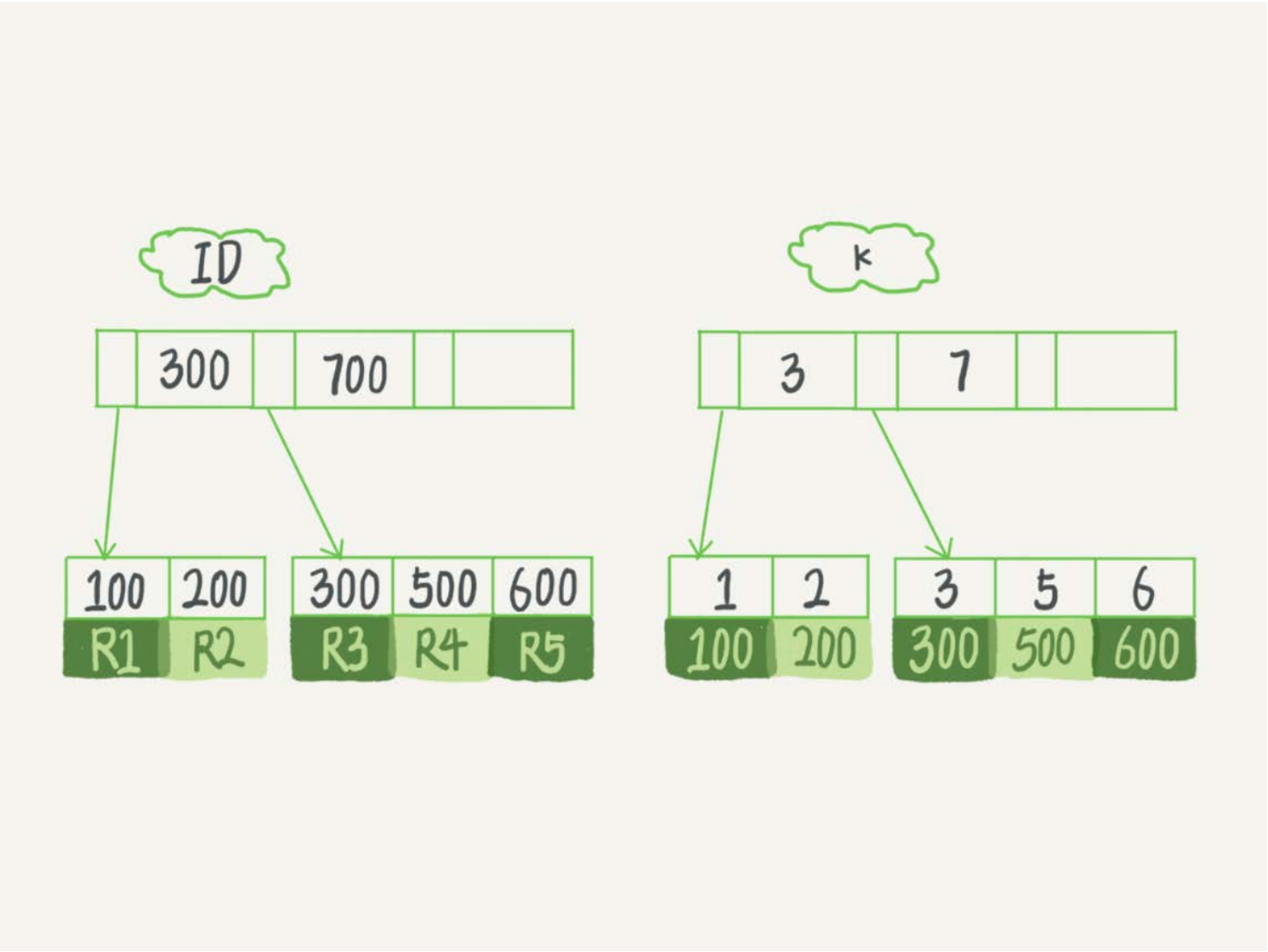


图4 InnoDB的索引组织结构

从图中不难看出，根据叶子节点的内容，索引类型分为主键索引和非主键索引。

主键索引的叶子节点存的是整行数据。在InnoDB里，主键索引也被称为聚簇索引（clustered index）。

非主键索引的叶子节点内容是主键的值。在InnoDB里，非主键索引也被称为二级索引（secondary index）。

根据上面的索引结构说明，我们来讨论一个问题：**基于主键索引和普通索引的查询有什么区别？**

- 如果语句是select * from T where ID=500，即主键查询方式，则只需要搜索ID这棵B+树；
- 如果语句是select * from T where k=5，即普通索引查询方式，则需要先搜索k索引树，得到ID的值为500，再到ID索引树搜索一次。这个过程称为回表。

也就是说，基于非主键索引的查询需要多扫描一棵索引树。因此，我们在应用中应该尽量使用主键查询。

索引|维护

B+树为了维护索引有序性，在插入新值的时候需要做必要的维护。以上面这个图为例，如果插入新的行ID值为700，则只需要在R5的记录后面插入一个新记录。如果新插入的ID值为400，就相对麻烦了，需要逻辑上挪动后面的数据，空出位置。

而更糟的情况是，如果R5所在的数据页已经满了，根据B+树的算法，这时候需要申请一个新的数据页，然后挪动部分数据过去。这个过程称为页分裂。在这种情况下，性能自然会受影响。

除了性能外，页分裂操作还影响数据页的利用率。原本放在一个页的数据，现在分到两个页中，整体空间利用率降低大约50%。

当然有分裂就有合并。当相邻两个页由于删除了数据，利用率很低之后，会将数据页做合并。合并的过程，可以认为是分裂过程的逆过程。

基于上面的索引维护过程说明，我们来讨论一个案例：

你可能在一些建表规范里面见到过类似的描述，要求建表语句里一定要有自增主键。当然事无绝对，我们来分析一下哪些场景下应该使用自增主键，而哪些场景下不应该。

自增主键是指自增列上定义的主键，在建表语句中一般是这么定义的：`NOT NULL PRIMARY KEY AUTO_INCREMENT`。

插入新记录的时候可以不指定ID的值，系统会获取当前ID最大值加1作为下一条记录的ID值。

也就是说，自增主键的插入数据模式，正符合了我们前面提到的递增插入的场景。每次插入一条新记录，都是追加操作，都不涉及到挪动其他记录，也不会触发叶子节点的分裂。

而有业务逻辑的字段做主键，则往往不容易保证有序插入，这样写数据成本相对较高。

除了考虑性能外，我们还可以从存储空间的角度来看。假设你的表中确实有一个唯一字段，比如字符串类型的身份证号，那应该用身份证号做主键，还是用自增字段做主键呢？

由于每个非主键索引的叶子节点上都是主键的值。如果用身份证号做主键，那么每个二级索引的叶子节点占用约20个字节，而如果用整型做主键，则只要4个字节，如果是长整型（bigint）则是8个字节。

显然，主键长度越小，普通索引的叶子节点就越小，普通索引占用的空间也就越小。

所以，从性能和存储空间方面考量，自增主键往往是更合理的选择。

有没有什么场景适合用业务字段直接做主键的呢？还是有的。比如，有些业务的场景需求是这样的：

1. 只有一个索引；

2. 该索引必须是唯一索引。

你一定看出来了，这就是典型的KV场景。

由于没有其他索引，所以也就不用考虑其他索引的叶子节点大小的问题。

这时候我们就要优先考虑上一段提到的“尽量使用主键查询”原则，直接将这个索引设置为主键，可以避免每次查询需要搜索两棵树。

小结

今天，我跟你分析了数据库引擎可用的数据结构，介绍了InnoDB采用的B+树结构，以及为什么InnoDB要这么选择。B+树能够很好地配合磁盘的读写特性，减少单次查询的磁盘访问次数。

由于InnoDB是索引组织表，一般情况下我会建议你创建一个自增主键，这样非主键索引占用的空间最小。但事无绝对，我也跟你讨论了使用业务逻辑字段做主键的应用场景。

最后，我给你留下一个问题吧。对于上面例子中的InnoDB表T，如果你要重建索引k，你的两个SQL语句可以这么写：

```
alter table T drop index k;  
alter table T add index(k);
```

如果你要重建主键索引，也可以这么写：

```
alter table T drop primary key;  
alter table T add primary key(id);
```

我的问题是，对于上面这两个重建索引的作法，说出你的理解。如果有不合适的，为什么，更好的方法是什么？

你可以把你的思考和观点写在留言区里，我会在下一篇文章的末尾给出我的参考答案。感谢你的收听，也欢迎你把这篇文章分享给更多的朋友一起阅读。

上期问题时间

我在上一篇文章末尾给你留下的问题是：如何避免长事务对业务的影响？

这个问题，我们可以从应用开发端和数据库端来看。

首先，从应用开发端来看：

1. 确认是否使用了set autocommit=0。这个确认工作可以在测试环境中开展，把MySQL的general_log开起来，然后随便跑一个业务逻辑，通过general_log的日志来确认。一般框架如果会设置这个值，也就会提供参数来控制行为，你的目标就是把它改成1。
2. 确认是否有不必要的只读事务。有些框架会习惯不管什么语句先用begin/commit框起来。我见过有些是业务并没有这个需要，但是也把好几个select语句放到了事务中。这种只读事务可以去掉。
3. 业务连接数据库的时候，根据业务本身的预估，通过SET MAX_EXECUTION_TIME命令，来控制每个语句执行的最长时间，避免单个语句意外执行太长时间。（为什么会意外？在后续的文章中会提到这类案例）

其次，从数据库端来看：

1. 监控 information_schema.innodb_trx表，设置长事务阈值，超过就报警/或者kill；
2. Percona的pt-kill这个工具不错，推荐使用；
3. 在业务功能测试阶段要求输出所有的general_log，分析日志行为提前发现问题；
4. 如果使用的是MySQL 5.6或者更新版本，把innodb_undo_tablespaces设置成2（或更大的值）。如果真的出现大事务导致回滚段过大，这样设置后清理起来更方便。

感谢 @壹笙 漂泊 @王凯 @易翔 留下的高质量评论。

 极客时间

MySQL 实战 45 讲

从原理到实战，丁奇带你搞懂 MySQL

林晓斌

网名丁奇
前阿里资深技术专家



精选留言



Christain

老师，索引篇结束了么？

有几个问题

- 1:三个字段联合索引时，如果中间的字段使用了范围查询或者模糊查询，最后一个字段还会用到索引么？
- 2:在order by时，索引是如何使用的
- 3:新建一张表，如何界定其索引的数量，有没有选择或者公式
- 4:能不能再详细介绍下字符串前缀索引

2018-11-23 09:09

作者回复

理论篇的先到这里，实践篇的还有六篇索引相关。

1. 我们文中有这样的例子了哦。like “张%”就是对name字段的范围查询/模糊查询。age就是你说的最后一个字段啦
2. 我们两篇的篇幅来讲order by, 第14和16（15是答疑）
3. 根据查询需要
4. 等《如何给字符串字段加索引》发布哈

嗯索引的内容其实很多，基础篇两篇是不可能写完的，而且怕理论类太多大家读着累，把一些知识点放到实践篇了 @all

2018-11-23 09:56



JackPn

老师我可不可以理解为：每一张表其实就是一个B+树，树结点的key值就是某一行的主键，value是该行的其他数据。新建索引就是新增一个B+树，查询不走索引就是遍历主B+树。

2018-11-27 20:06

作者回复

每一个表是好几棵B+树（应该是你理解对了但是手误），其它的完全正确

2018-11-27 20:55



wuxue_123

老师，回表只是普通索引才会有的吗？主键和数据放在同一个树中，根据主键查询的时候，就可以直接获得数据了。

那select * from table where id = xx

和select id from table where id = xx

的效率是一样的吗？（id是主键）

2018-12-20 06:39

作者回复

这两个语句是都不用回表了，在“查找行”这个逻辑上是一样的，

但是select *要读和拷贝更多列到server,还要发送更多列给客户端，所以还是select id更快的。

好问题

2018-12-20 10:00



壹笙 漂泊

总结：

- 1.索引的作用：提高数据查询效率
- 2.常见索引模型：哈希表、有序数组、搜索树
- 3.哈希表：键 - 值(key - value)。
- 4.哈希思路：把值放在数组里，用一个哈希函数把key换算成一个确定的位置，然后把value放在数组的这个位置
- 5.哈希冲突的处理办法：链表
- 6.哈希表适用场景：只有等值查询的场景
- 7.有序数组：按顺序存储。查询用二分法就可以快速查询，时间复杂度是： $O(\log(N))$
- 8.有序数组查询效率高，更新效率低
- 9.有序数组的适用场景：静态存储引擎。
- 10.二叉搜索树：每个节点的左儿子小于父节点，父节点又小于右儿子
- 11.二叉搜索树：查询时间复杂度 $O(\log(N))$ ，更新时间复杂度 $O(\log(N))$
- 12.数据库存储大多不适用二叉树，因为树高过高，会适用N叉树
- 13.InnoDB中的索引模型：B+ Tree
- 14.索引类型：主键索引、非主键索引
主键索引的叶子节点存的是整行的数据(聚簇索引)，非主键索引的叶子节点内容是主键的值(二级索引)
- 15.主键索引和普通索引的区别：主键索引只要搜索ID这个B+ Tree 即可拿到数据。普通索引先搜索索引拿到主键值，再到主键索引树搜索一次(回表)
- 16.一个数据页满了，按照B+ Tree 算法，新增加一个数据页，叫做页分裂，会导致性能下降。空间利用率降低大概50%。当相邻的两个数据页利用率很低的时候会做数据页合并，合并的过程是分裂过程的逆过程。
- 17.从性能和存储空间方面考量，自增主键往往是更合理的选择。

思考题：

如果删除，新建主键索引，会同时去修改普通索引对应的主键索引，性能消耗比较大。
删除重建普通索引貌似影响不大，不过要注意在业务低谷期操作，避免影响业务。

2018-11-21 14:28

作者回复

优秀

2018-11-21 14:50



约书亚

“N叉树”的N值在MySQL中是可以被人工调整的么？曾经面试被问到过这问题，当时就懵逼了...

2018-11-21 08:49

作者回复

面试中题面越简单的问题越暗藏凶险，可见一斑...

可以按照调整key的大小的思路来说；

如果你能指出来5.6以后可以通过page大小来间接控制应该能加分吧

面试回答不能太精减，计算方法、前缀索引什么的一起上

2018-11-21 14:40



张良

结合王争的数据结构与算法看正好

2018-11-21 09:58



路过

顺序应是先删除k列索引，主键索引。然后再创建主键索引和k列索引。

2018-11-21 09:49



张先森森森

drop主键索引会导致其他索引失效，但drop普通索引不会。

2018-11-21 07:24



唐堂@贝壳找房

@约书亚 的问题：“N叉树”的N值在Mysql是否可以被调整？”，老师能否在稍微详细点指导下？谢谢老师了

2018-11-21 20:55



高枕

我来回答这个问题

你可以想象下一棵 100 万节点的平衡二叉树，树高 20。一次查询可能需要访问 20 个数据块。

这里有点不理解，为什么树高20就是20个数据块？

每个叶子结点就是一个块，每个块包含两个数据，块之间通过链式方式链接。树高20的话，就要遍历20个块

2018-11-21 12:33

作者回复

，多谢

多补一句，因为是二叉树结构，每次指针查找很大概率是触发随机磁盘读（比如很难刚好碰上一个节点和他的左右儿子刚好相邻）

2018-11-21 13:45



Richie

二级索引重建应该新建索引再做删除，如果有查询用到这个索引，此时索引已被删除，会导致业务抖动。主键重建不能采用drop这种方式去按操作，因为所有数据都是以主键组织的，删了主键后，InnoDB会自己找一个主键组织数据，再次添加主键又会重新组织数据，重建表的次数已达二次，我们可以直接Optimiz这个表

2018-11-21 03:57



Geek_5b3ccb

请问没有主键的表，有一个普通索引。怎么回表？

2018-11-21 09:25

作者回复

没有主键的表，innodb会给默认创建一个Rowid做主键

2018-11-21 14:11

monkay

你可以想象一下一棵 100 万节点的平衡二叉树，树高 20。一次查询可能需要访问 20 个数据块。

这里有点不理解，为什么树高20就是20个数据块？

2018-11-21 08:53

作者回复

后面有同学答复你了

2018-11-21 14:25



滩涂曳尾

老师我有2个问题，问题1是上一节遗留的没来得及问：

1. 在“读提交”隔离级别下，这个视图是在每个 SQL 语句开始执行的时候创建的。这句话能具体说明吗？
2. 访问磁盘和内存索引涉及磁盘(sata, ssd, nvme)读写性能，以及内存读写性能，可否给一些数值方便直观认识？

2018-11-21 07:42

作者回复

1. 就是语句开始执行前创建一个read-view

2. 我估计你要的是这个：

This group of numbers is from a presentation Jeff Dean gave at a Engineering All-Hands Meeting at Google.

L1 cache reference 0.5 ns

Branch mispredict 5 ns

L2 cache reference 7 ns

Mutex lock/unlock 100 ns

Main memory reference 100 ns

Compress 1K bytes with Zippy 10,000 ns

Send 2K bytes over 1 Gbps network 20,000 ns

Read 1 MB sequentially from memory 250,000 ns

Round trip within same datacenter 500,000 ns

Disk seek 10,000,000 ns

Read 1 MB sequentially from network 10,000,000 ns

Read 1 MB sequentially from disk 30,000,000 ns

Send packet CA->Netherlands->CA 150,000,000 ns

2018-11-22 00:38



大王拍我去巡山

老师你好：之前看过一遍文章，一直有疑惑：一个InnoDB引擎的表，数据量非常大，根据二级索引搜索会比主键搜索快，文章阐述的原因是主键索引和数据行在一起，非常大搜索慢，我的疑惑是：通过

普通索引找到主键ID后，同样要跑一边主键索引，还望老师解惑。。。

2018-11-21 09:52

作者回复

问出这个问题表示你今天的文章看明白了

看完明天的（下），我估计你就有答案了

2018-11-21 14:04



Richie

老师索引只能定位到page，page内部怎么去定位行数据

2018-11-21 16:42

作者回复

内部有个有序数组，二分法

2018-11-21 17:48



老师请问下：

- 1、如果插入的数据是在主键树叶子结点的中间，后面的所有页如果都是满的状态，是不是会造成后面的每一页都会去进行页分裂操作，直到最后一个页申请新页移过去最后一个值
- 2、还有之前看到过说是插入数据如果是在某个数据满了页的首尾，为了减少数据移动和页分裂，会先去前后两个页看看是否满了，如果没满会先将数据放到前后两个页上，不知道是不是有这种情况

2018-11-21 11:13

作者回复

1. 不会不会，只会分裂它要写入的那个页面。每个页面之间是用指针串的，改指针就好了，不需要“后面的全部挪动

2. 对，减为了增加空间利用率

2018-11-21 13:56



大王拍我去巡山

希望老师讲一个联合索引在B+树中的存储方式

2018-11-21 09:53

作者回复

哎呀，下一篇就是

2018-11-21 14:02



jacket

【今日收获】

1. 主键索引的叶子结点存储了整一行的内容（聚簇索引），使用主键可以快速获取到整行的数据。
2. 非主键索引的叶子结点存储的是主键的值，所以主键字段占用空间不宜过大。同时，其查找数据的过程称为“回表”，需要先查找自己得到主键值，再在主键索引上边查找数据内容。
3. 索引的实现由存储引擎来决定，InnoDB使用B+树（N叉树，比如1200叉树），把整颗树的高度维持在很小的范围内，同时在内存里缓存前面若干层的节点，可以极大地降低访问磁盘的次数，提高读的效率。
4. B+树的插入可能会引起数据页的分裂，删除可能会引起数据页的合并，二者都是比较重的IO消耗，所以比较好的方式是顺序插入数据，这也是我们一般使用自增主键的原因之一。
5. 在Key-Value的场景下，只有一个索引且是唯一索引，则适合直接使用业务字段作为主键索引。

2018-11-22 12:56

作者回复

赞

2018-11-22 15:29



到道可道

由于索引中存储的是主键的key值，那么重建主键时，不但会重建主键，所有的普通索引也会重新构建，所以通常不建议重建索引，无普通索引的除外。

2018-11-21 08:30