

racheol

MyISAM和InnoDB的索引实现

在 MySQL 中，主要有四种类型的索引，分别为：**B-Tree** 索引，**Hash** 索引，**Fulltext** 索引和 **R-Tree** 索引。我们主要分析B-Tree 索引。

B-Tree 索引是 MySQL 数据库中使用最为频繁的索引类型，除了 Archive 存储引擎之外的其他所有的存储引擎都支持 B-Tree 索引。Archive 引擎直到 MySQL 5.1 才支持索引，而且只支持索引单个 AUTO_INCREMENT 列。

不仅仅在 MySQL 中是如此，实际上在其他的很多数据库管理系统中B-Tree 索引也同样是作为最主要的索引类型，这主要是因为 B-Tree 索引的存储结构在数据库的数据检索中有非常优异的表现。

一般来说，MySQL 中的 B-Tree 索引的物理文件大多都是以 Balance Tree 的结构来存储的，也就是所有实际需要的数据都存放于 Tree 的 Leaf Node(叶子节点)，而且到任何一个 **Leaf Node** 的最短路径的长度都是完全相同的，所以我们大家都称之为 B-Tree 索引。当然，可能各种数据库（或 MySQL 的各种存储引擎）在存放自己的 B-Tree 索引的时候会对存储结构稍作改造。如 **Innodb** 存储引擎的 **B-Tree** 索引实际使用的存储结构实际上是 **B+Tree**，也就是在 B-Tree 数据结构的基础上做了很小的改造，在每一个Leaf Node 上面出了存放索引键的相关信息之外，还存储了指向与该 **Leaf Node** 相邻的后一个 **LeafNode** 的指针信息（增加了顺序访问指针），这主要是为了加快检索多个相邻 **Leaf Node** 的效率考虑。

下面主要讨论**MyISAM**和**InnoDB**两个存储引擎的索引实现方式：

1. MyISAM索引实现：MyISAM索引文件和数据文件是分离的，索引文件仅保存数据记录的地址。

1）主键索引：

MyISAM引擎使用B+Tree作为索引结构，叶节点的**data**域存放的是数据记录的地址。下图是MyISAM主键索引的原理图：

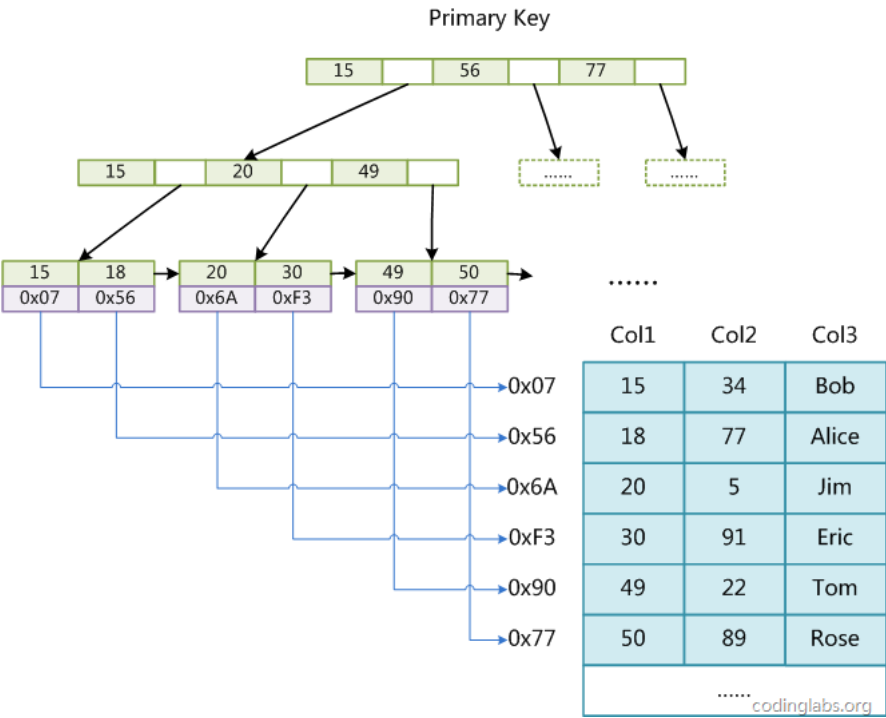
公告

昵称：立超的专栏
园龄：3年2个月
粉丝：1
关注：0
[+加关注](#)

搜索

随笔分类

- C/C++(33)
- Debug Hacks(4)
- Git(1)
- Linux(19)
- Linux使用手册(8)
- mac使用心得(4)
- Mysql(4)
- nginx
- Python
- UNIX网络编程(9)
- 工具(2)
- 海量数据处理(6)
- 环境搭建(5)

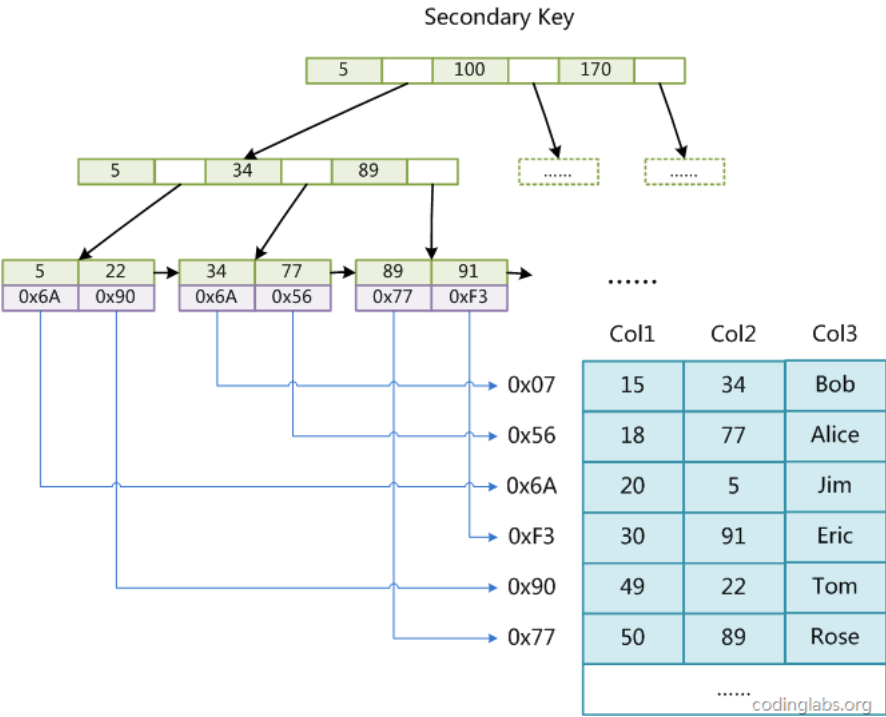


(图myisam1)

这里设表一共有三列，假设我们以Col1为主键，图myisam1是一个MyISAM表的主索引（Primary key）示意。可以看出MyISAM的索引文件仅仅保存数据记录的地址。

2）辅助索引（Secondary key）

在MyISAM中，主索引和辅助索引（Secondary key）在结构上没有任何区别，只是主索引要求key是唯一的，而辅助索引的key可以重复。如果我们在Col2上建立一个辅助索引，则此索引的结构如下图所示：



同样也是一颗B+Tree，data域保存数据记录的地址。因此，MyISAM中索引检索的算法为首先按照B+Tree搜索算法搜索索引，如果指定的Key存在，则取出其data域的值，然后以data域的值作为地址，读取相应数据记录。

MyISAM的索引方式也叫做“非聚集”的，之所以这么称呼是为了与InnoDB的聚集索引区分。

2. InnoDB索引实现

- 机器学习(1)
- 快捷键小抄(3)
- 深入理解计算机系统(14)
- 算法(5)
- 信号处理(2)

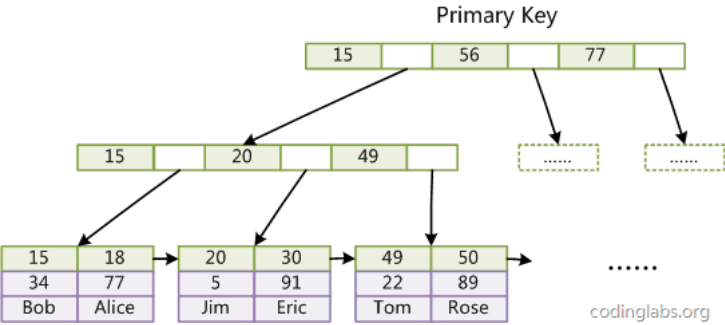
阅读排行榜

- 1. Mysql中MyISAM索引的比较(2459)
- 2. linux下Clang和gcc
- 3. NS3系列——eclipse(1585)
- 4. 内存为程序分配空间565)
- 5. MyISAM和InnoDB 6)

然InnoDB也使用B+Tree作为索引结构，但具体实现方式却与MyISAM截然不同。

1) 主键索引：

MyISAM索引文件和数据文件是分离的，索引文件仅保存数据记录的地址。而在InnoDB中，表数据文件本身就是按B+Tree组织的一个索引结构，这棵树的叶节点data域保存了完整的数据记录。这个索引的key是数据表的主键，因此InnoDB表数据文件本身就是主索引。

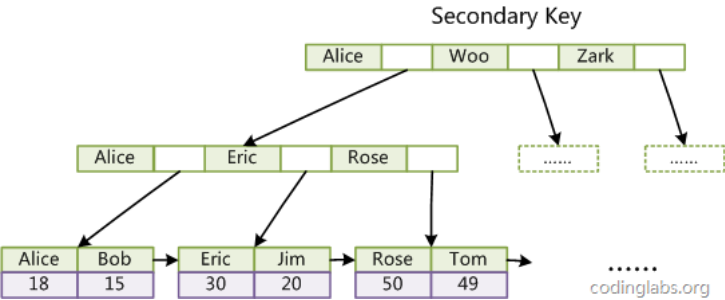


(图innodb主键索引)

(图innodb主键索引)是InnoDB主索引（同时也是数据文件）的示意图，可以看到叶节点包含了完整的数据记录。这种索引叫做**聚集索引**。因为InnoDB的数据文件本身要按主键聚集，所以InnoDB要求表必须有主键（MyISAM可以没有），如果没有显式指定，则MySQL系统会自动选择一个可以唯一标识数据记录的列作为主键，如果不存在这种列，则MySQL自动为InnoDB表生成一个隐含字段作为主键，这个字段长度为6个字节，类型为长整形。

2) . InnoDB的辅助索引

InnoDB的所有辅助索引都引用主键作为data域。例如，下图为定义在Col3上的一个辅助索引：



InnoDB 表是基于**聚集索引**建立的。因此InnoDB 的索引能提供一种非常快速的主键查找性能。不过，它的辅助索引（Secondary Index，也就是非主键索引）也会包含主键列，所以，如果主键定义的比较大会，其他索引也将很大。如果在表上定义、很多索引，则**争取尽量把主键定义得小一些**。InnoDB 不会压缩索引。

文字符的ASCII码作为比较准则。聚集索引这种实现方式使得按主键的搜索十分高效，但是**辅助索引搜索需要检索两遍索引：首先检索辅助索引获得主键，然后用主键到主索引中检索获得记录。**

不同存储引擎的索引实现方式对于正确使用和优化索引都非常有帮助，例如知道了InnoDB的索引实现后，就很容易明白**1、为什么不建议使用过长的字段作为主键，因为所有辅助索引都引用主索引，过长的主索引会令辅助索引变得过大。**再例如，**2、用非单调的字段作为主键在InnoDB中不是个好主意，因为InnoDB数据文件本身是一颗B+Tree，非单调的主键会造成在插入新记录时数据文件为了维持B+Tree的特性而频繁的分裂调整，十分低效，而使用自增字段作为主键则是一个很好的选择。**

InnoDB索引和MyISAM索引的区别：


- 一是主索引的区别，InnoDB的数据文件本身就是索引文件。而MyISAM的索引和数据是分开的。
- 二是辅助索引的区别：InnoDB的辅助索引data域存储相应记录主键的值而不是地址。而MyISAM的辅助索引和主索引没有多大区别。

分类： Mysql

好文要顶

关注我

收藏该文



立超的专栏
关注 - 0
粉丝 - 1
[+加关注](#)

0

0

« 上一篇: B树、B-树、B+树、B*树

» 下一篇: Linux下gcc和g++的区别

posted @ 2016-08-10 16:19 立超的专栏 阅读(1136) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问网站首页](#)。

- 【推荐】50万行VC++源码: 大型组态工控、电力仿真CAD与GIS源码库
- 【推荐】中铁、中石油等大型企业的复杂报表解决方案
- 【福利】阿里云免费套餐升级，更多产品，更久时长

最新IT新闻:

- 版权再亮红灯 乐视体育还想继续讲故事
- 90后喜欢透支未来? 99%的90后更重视信用积累
- 用4425颗卫星建太空互联网，SpaceX后年开始发射
- Bing新功能将替用户“解密”主页背景图背后的故事
- 特斯拉将开设首家自营汽车修理门店 下季度新增100辆机动维修车

» 更多新闻...

最新知识库文章:

- 唱吧DevOps的落地，微服务CI/CD的范本技术解读
- 程序员，如何从平庸走向理想?
- 我为什么鼓励工程师写blog
- 怎么轻松学习JavaScript
- 如何打好前端游击战

» 更多知识库文章...