

# 数据库的最简单实现

---

作者： 阮一峰

日期： 2014年7月 4日

所有应用软件之中，数据库可能是最复杂的。

MySQL的手册有3000多页，PostgreSQL的手册有2000多页，Oracle的手册更是比它们相加还要厚。



但是，自己写一个最简单的数据库，做起来并不难。Reddit上面有一个[帖子](#)，只用了几百个字，就把原理讲清楚了。下面是我根据这个帖子整理的内容。

## 一、数据以文本形式保存

---

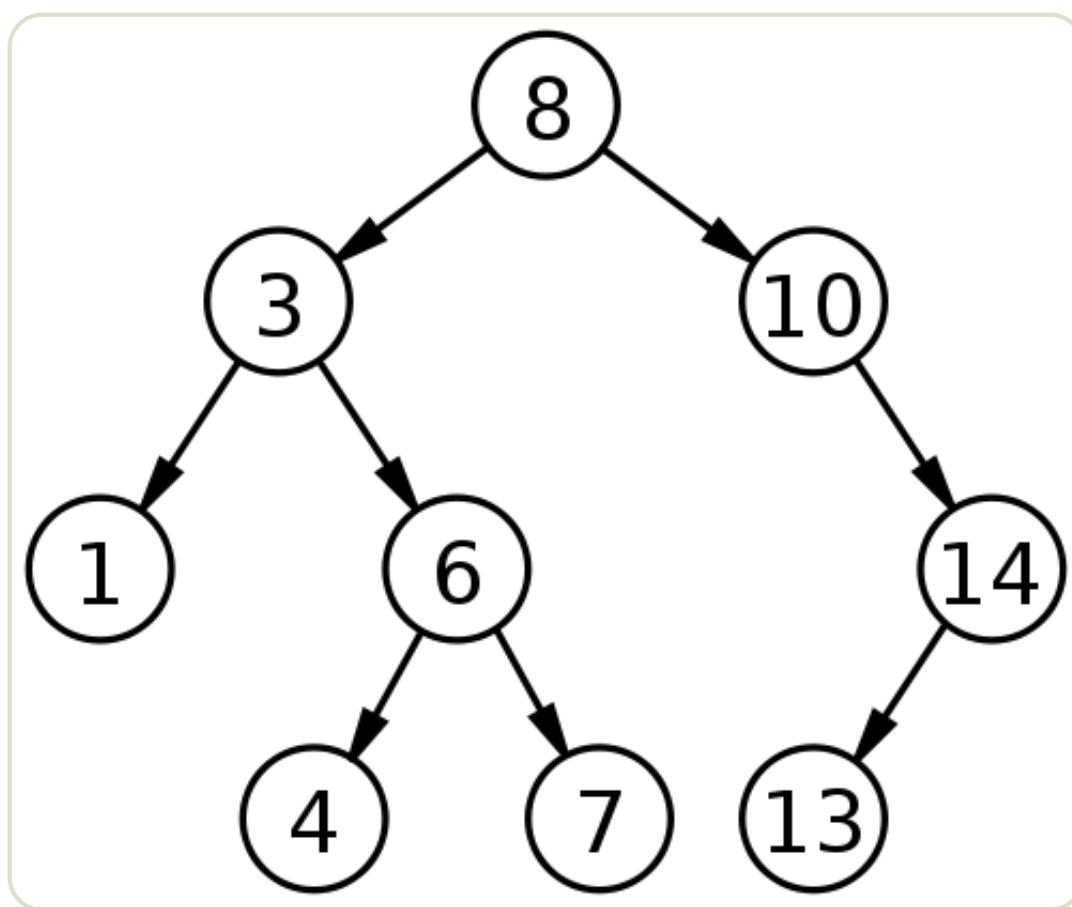
第一步，就是将所要保存的数据，写入文本文件。这个文本文件就是你的数据库。

为了方便读取，数据必须分成记录，每一条记录的长度规定为等长。比如，假定每条记录的长度是800字节，那么第5条记录的开始位置就在3200字节。

大多数时候，我们不知道某一条记录在第几个位置，只知道主键（primary key）的值。这时为了读取数据，可以一条条比对记录。但是这样做效率太低，实际应用中，数据库往往采用B树（B-tree）格式储存数据。

## 二、什么是B树？

要理解B树，必须从二叉查找树（Binary search tree）讲起。



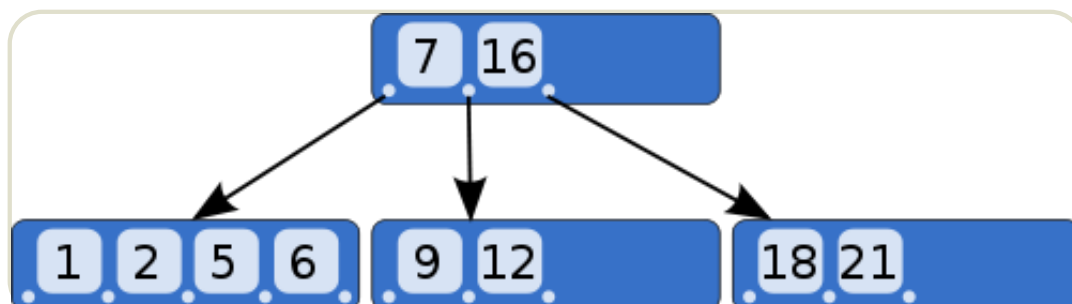
二叉查找树是一种查找效率非常高的数据结构，它有三个特点。

- （1）每个节点最多只有两个子树。
- （2）左子树都为小于父节点的值，右子树都为大于父节点的值。
- （3）在n个节点中找到目标值，一般只需要 $\log(n)$ 次比较。

二叉查找树的结构不适合数据库，因为它的查找效率与层数相关。越处在下层的数据，

就需要越多次比较。极端情况下， $n$ 个数据需要 $n$ 次比较才能找到目标值。对于数据库来说，每进入一层，就要从硬盘读取一次数据，这非常致命，因为硬盘的读取时间远远大于数据处理时间，数据库读取硬盘的次数越少越好。

**B树**是对二叉查找树的改进。它的设计思想是，将相关数据尽量集中在一起，以便一次读取多个数据，减少硬盘操作次数。



B树的特点也有三个。

- (1) 一个节点可以容纳多个值。比如上图中，最多的一个节点容纳了4个值。
- (2) 除非数据已经填满，否则不会增加新的层。也就是说，B树追求"层"越少越好。
- (3) 子节点中的值，与父节点中的值，有严格的大小对应关系。一般来说，如果父节点有 $a$ 个值，那么就有 $a+1$ 个子节点。比如上图中，父节点有两个值（7和16），就对应三个子节点，第一个子节点都是小于7的值，最后一个子节点都是大于16的值，中间子节点就是7和16之间的值。

这种数据结构，非常有利于减少读取硬盘的次数。假定一个节点可以容纳100个值，那么3层的B树可以容纳100万个数据，如果换成二叉查找树，则需要20层！假定操作系统一次读取一个节点，并且根节点保留在内存中，那么B树在100万个数据中查找目标值，只需要读取两次硬盘。

## 三、索引

数据库以B树格式储存，只解决了按照"主键"查找数据的问题。如果想查找其他字段，就需要建立索引（index）。

所谓索引，就是以某个字段为关键字的B树文件。假定有一张"雇员表"，包含了员工号

（主键）和姓名两个字段。可以对姓名建立索引文件，该文件以B树格式对姓名进行储存，每个姓名后面是其在数据库中的位置（即第几条记录）。查找姓名的时候，先从索引中找到对应第几条记录，然后再从表格中读取。

这种索引查找方法，叫做"索引顺序存取方法"（Indexed Sequential Access Method），缩写为ISAM。它已经有多种实现（比如C-ISAM库和D-ISAM库），只要使用这些代码库，就能自己写一个最简单的数据库。

## 四、高级功能

部署了最基本的数据存取（包括索引）以后，还可以实现一些高级功能。

（1）**SQL**语言是数据库通用操作语言，所以需要有一个SQL解析器，将SQL命令解析为对应的ISAM操作。

（2）数据库连接（**join**）是指数据库的两张表通过"外键"，建立连接关系。你需要对这种操作进行优化。





（3）数据库事务（**transaction**）是指批量进行一系列数据库操作，只要有一步不成功，整个操作都不成功。所以需要有一个"操作日志"，以便失败时对操作进行回滚。

（4）备份机制：保存数据库的副本。

（5）远程操作：使得用户可以在不同的机器上，通过TCP/IP协议操作数据库。

（完）

### 文档信息

- 版权声明：自由转载-非商用-非衍生-保持署名（[创意共享3.0许可证](#)）
- 发表日期：2014年7月 4日
- 更多内容： [档案](#) » [理解计算机](#)
- 购买文集：  《如何变得有思想》
- 社交媒体：  twitter,  weibo
- Feed订阅： 



## 相关文章

- **2016.08.05:** [布尔代数入门](#)

布尔代数是计算机的基础。没有它，就不会有计算机。

- **2016.06.16:** [DNS 原理入门](#)

DNS 是互联网核心协议之一。不管是上网浏览，还是编程开发，都需要了解一点它的知识。

- **2015.09.28:** [为什么主引导记录的内存地址是0x7C00?](#)

《计算机原理》课本说，启动时，主引导记录会存入内存地址0x7C00。

- **2015.07.16:** [图解 Monad](#)

函数式编程有一个重要概念，叫做Monad。

联系方式 | [ruanyifeng.com](http://ruanyifeng.com) 2003 - 2016