

什么是基数树

一句话，基数树是一种多叉树。

更官方一点的解释：radix tree是一种多叉搜索树。树的叶子结点是实际的数据条目。每一个结点有一个固定的、 2^n 指针指向子结点（每一个指针称为槽slot，n为划分的基的大小）。看到这里可能还有点懵，继续往下看。

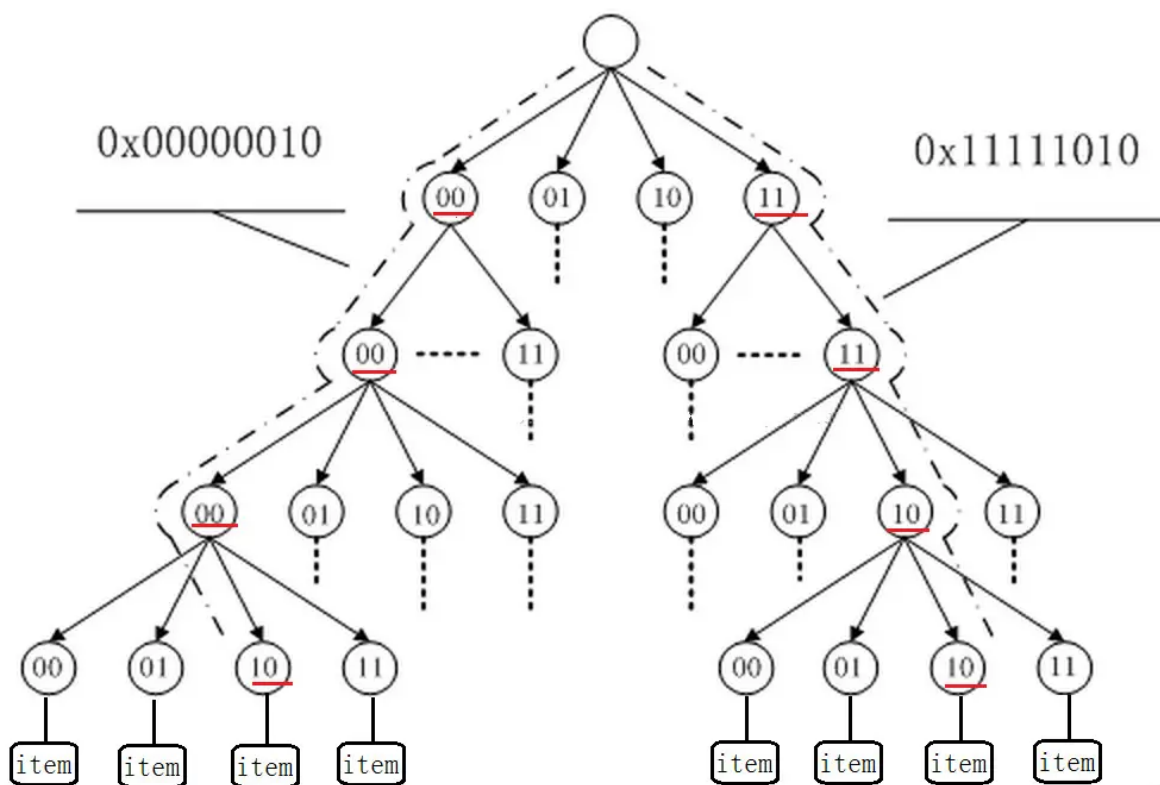
为什么要设计基数树

举个例子，一目了然。对于下面四个kv键值对，我们如何存储？

```
<0101, "abc">
<010, "abcd">
<001, "bcde">
<0110, "cdef">
复制代码
```

有人说用hash表，是可以，但是hash表有两个问题：

- 1.hash冲突。hash函数不好设计，容易产生冲突，需要解决hash冲突
 - 2.hash表大小不好确定。hash表底层还是数组实现的，数组的大小不好确定，涉及到扩容的问题
- 如果用Radix Tree就很容易解决上面两个问题，看下图：



@稀土掘金技术社区

上图是n=4的基数树

怎么实现基数树

1.定义结构体

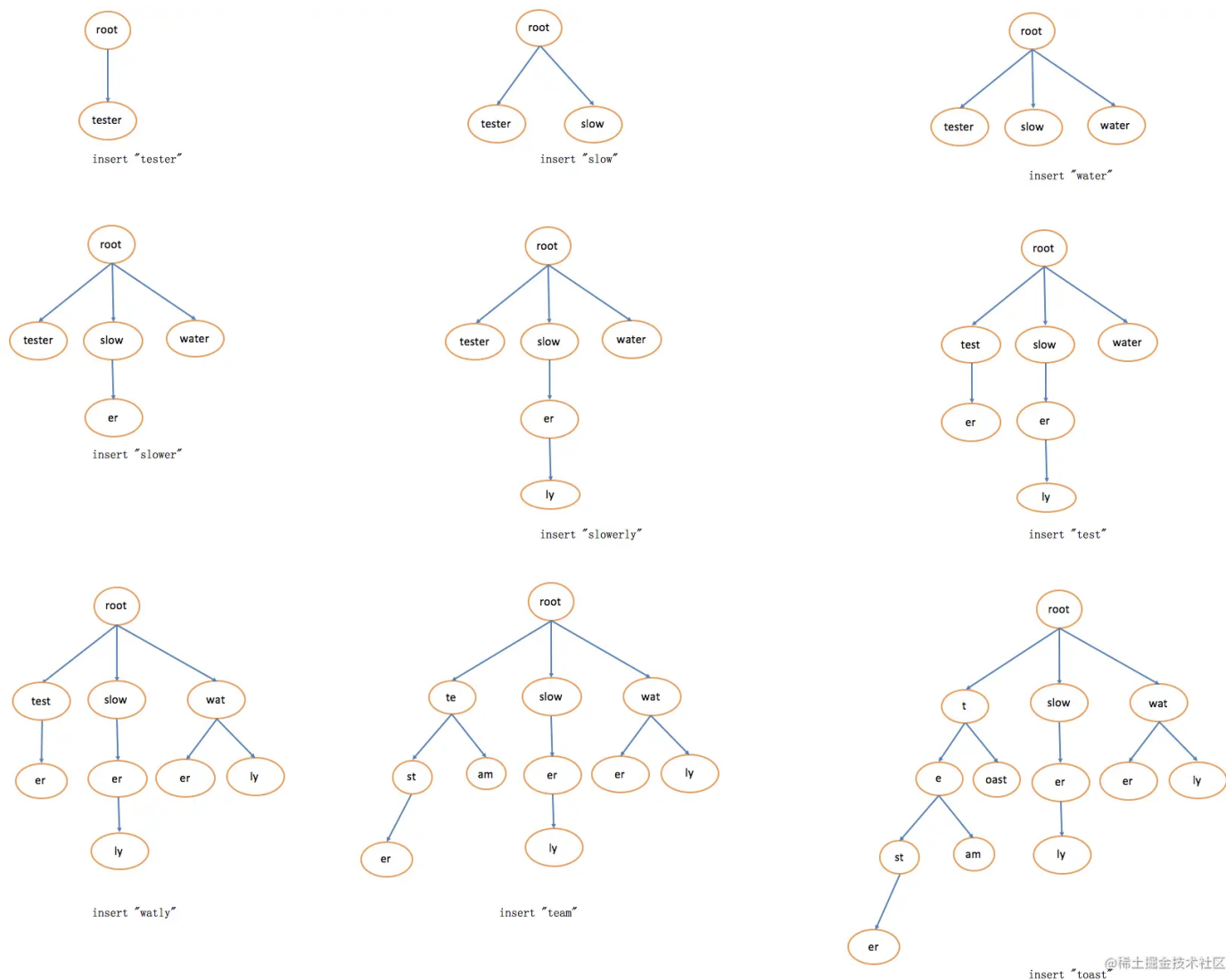
```
// 定义节点
type radix_node struct {
    Val    string    `json:"Val"`
    Childs []*radix_node `json:"Childs"`
}

// 定义RadixTree结构体
type RadixTree struct {
    Root *radix_node `json:"Root"`
}

复制代码
```

2.插入方法

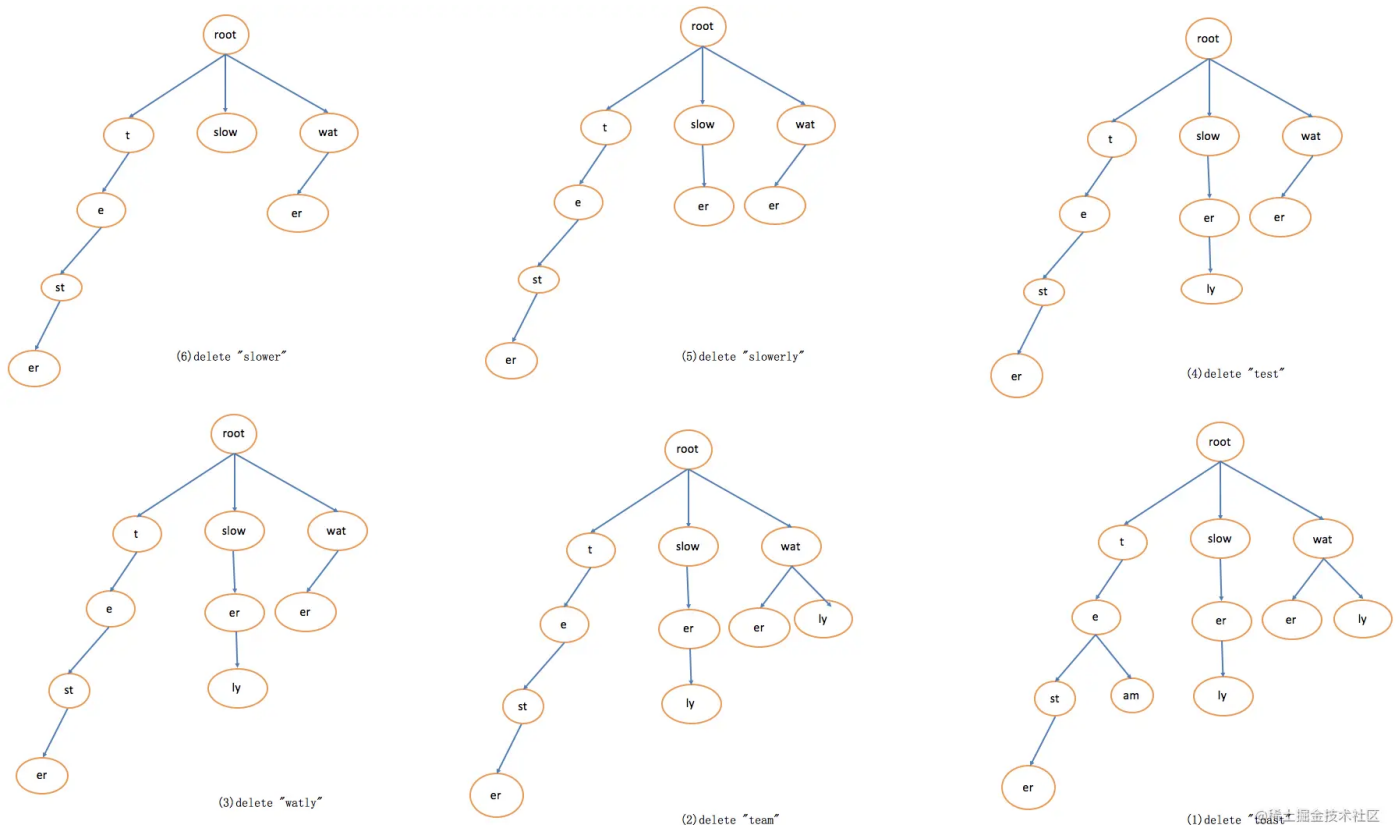
从根节点往下找，如果在某个孩子中找到相同前缀则继续往下找；如果某个孩子部分前缀是字符串前缀，则需要将公共前缀作为新节点，原来孩子节点的孩子作为新节点的孩子。



插入方法有点麻烦，可参考完整代码的insert_test.go文件，有9个测试用例，实现了上图9种case并且打印出了每种case的结果，结合上图和单侧更容易理解代码。

3.删除方法

删除比较简单，沿着根节点往下找，找到某个节点匹配到整个字符串位置，当目标节点是叶子节点（没有孩子）时删除该节点。



删除一个节点之后，与插入该节点之前数据结构不一样，是因为删除的时候只会删除叶子节点，当要删除的节点为非叶子节点的时候并没有删除。正常情况下并没有叶子节点和非叶子节点之分，每个节点都可用保存一个要关联的值。当父节点只有一个孩子节点并且父节点没有关联值的时候，需要将父节点和孩子节点合并，这里为了简单起见，节点没有关联值，因此删除的时候并没有合并父子节点的操作。

4. 查找方法

查找比较简单，类似于Delete方法。只需要一层一层匹配即可，如果在某一层的所有孩子中都没有匹配到则查找失败。

千言万语不如图直接，相信看完上述两张图就很容易理解了Radix Tree的基本原理。下面代码是上面两张图的实现，可对照阅读：github.com/ZBIGBEAR/radix-tree

基数树的应用场景

(1) Radix Tree在Linux中的应用

IDR(ID Radix)机制是将对象的身份鉴别ID与对象指针建立关联，实现从ID与指针之间的转换。说的简单点就是两个长整型变量之间的映射，使用Radix Tree存储节省空间并且查找速度快，这是Radix Tree的最大特点，在linux中的应用也是基于此特点

Radix Tree在Linux中最大的应用是内存管理。使用Radix Tree将页面描述符组织起来，方便查找，详情请查看[Linux Cache管理之radix_tree](#)

(2) Radix Tree存储稀疏数组

一个较小的数据集，每个记录都是比较长的数字或者字符串，且这些记录有较多的公共前缀。

例如：路由表结构，ip数据集不大，且有大量前缀可共享

例如：内存地址映射，连续地址之间有大量前缀可共享

例如：基因搜索。基因有非常长的字符串，且有大量共享片段

(3) Radix Tree在数据库中的应用

参考文档8，SP-GiST索引用到了Radix Tree

(4) Radix Tree在压缩存储方面的应用

当存储的多个记录之前有大量公共前缀时可用Radix Tree做压缩

参考

- 【1】 [查找——图文翔解RadixTree（基数树）](#)
- 【2】 [基数树（Radix Tree）](#)
- 【3】 [利用Radix树作为Key-Value 键值对的数据路由](#)
- 【4】 [Nginx源代码分析-radix tree](#)
- 【5】 [Linux 基数（radix）树](#)
- 【6】 [Linux Cache管理之radix_tree](#)
- 【7】 [基数树wiki](#)
- 【8】 [radix tree在数据库PostgreSQL中的一些应用举例](#)
- 【9】 [数据结构之Radix Tree](#)