

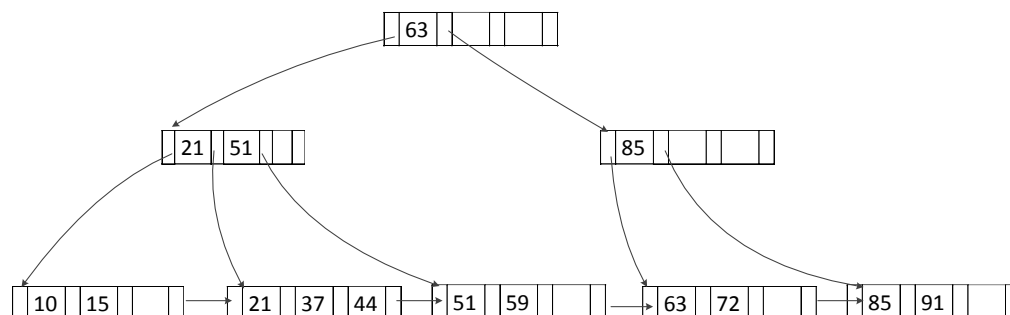
一、假设我们利用可扩展 hash 方法（散列函数 $h(k)$ 是一个 b (足够大) 位二进制序列，后 d 位表示桶的数目），对以下记录进行 hash 存储：

3,8,11,14,15,16,17,19,20,33,43,48

对于 hash 函数 $h(x) = x \bmod 7$ ，要求每个桶至多包含 3 个元素。按照给定的顺序进行添加，回答下列问题：

- (1) 当只用最后一位当索引时，最后能添加的元素是？画出当前的索引结构。
- (2) 在上一问的基础上继续添加元素，当只用最后两位当索引时，最后能添加的元素是？画出当前的索引结构。
- (3) 画出添加完所有元素后，最终的索引结构。

二、已知一棵 B+ 树，如图 1 所示



- (1) 请画出在图 1 中插入 40 后所得的新的 B+ 树（给出图示即可）
- (2) 请画出在图 1 中删除 85 后所得的新的 B+ 树（给出图示即可）

三、用下面的关键码值集合建立一个 B+ 树

(2, 3, 4, 7, 11, 17, 19, 23, 29, 31)

假设树初始为空，值按上升顺序加入，根据一个节点所能容纳指针数的下列情况分别构造 B+ 树。

- (1) 4
- (2) 8

四、考虑关系 $r_1(A, B, C)$ ， $r_2(C, D, E)$ 及 $r_3(E, F)$ 。

(1) 假设 r_1 、 r_2 、 r_3 的主码分别是 A、C、E。 r_1 有 1000 个元组， r_2 有 1500 个元组， r_3 有 750 个元组，估计 $r_1 \bowtie r_2 \bowtie r_3$ 的大小，给出一个有效地计算这个连接的策略。（要求写出具体计算过程）

(2) 假设除了整个模式外没有主码，令 $V(C, r_1)$ 为 900， $V(C, r_2)$ 为 1100， $V(E, r_2)$ 为 50， $V(E, r_3)$ 为 100。假设 r_1 有 1000 个元组， r_2 有 1500 个元组， r_3 有 750 个元组。估计 $r_1 \bowtie r_2 \bowtie r_3$ 的大小，给出一个有效地计算这个连接的策略。（要求写出具体计算过程）

五、图书馆数据库有如下 3 个关系模式：

B(B#, Title, Author, Publisher)

S(S#, Name, Department)

L(S#, B#, Date)

其中 B 为图书信息表：B#为图书编号，Title 为书名，Author 为作者，Publisher 为出版社；

S 为学生信息表：S#为学号，Name 为学生姓名，Department 为学院名；

L 为借阅信息表：S#为借阅人学号，B#为被借阅图书编号，Date 为借阅日期。

(1) 用户有一查询语句：

Select Title

From B, L

Where B.B# = L. B# and L.Date=" 2017 年 5 月 1 日"

检索在"2017 年 5 月 1 日"被借阅的图书名

等价的关系代数表达式有：

$$Q1 = \pi_{Title}(\sigma_{Date="2017 \text{ 年 } 5 \text{ 月 } 1 \text{ 日}"}(B \bowtie L))$$

$$Q2 = \pi_{Title}(B \bowtie \sigma_{Date="2017 \text{ 年 } 5 \text{ 月 } 1 \text{ 日}"}(L))$$

$$Q3 = \pi_{Title}(\pi_{B.B\#, Title}(B) \bowtie \pi_{L.B\#}(\sigma_{Date="2017 \text{ 年 } 5 \text{ 月 } 1 \text{ 日}"}(L)))$$

设 L 表有 20000 条元组，B 表有 3000 条元组，L 表中满足在"2017 年 5 月 1 日"借阅图书的元组数为 100（有多本同名书籍时，图书编号各不相同）。每个缓存块能存放 L 表的 200 条元组或 B 表的 60 条元组。内存中最多可容纳 10 块 B 表元组和 1 块 L 表元组。B 表默认在主键 B#上有序。

求解下列问题，给出具体计算步骤：

- 对于查询计划 Q1, 分别针对 Nest-loop(B 表是外层循环中的关系), Sort-Merge, 计算自然连接操作需要读取的总块数。（注：L 在 Date 上无索引）
- 对于查询计划 Q2, 分别针对 Nest-loop, Sort-Merge, 计算自然连接操作需要读取的总块数。（注：对两种连接操作算法，分别考虑 L 在 Date 上有索引和无索引的情形）
- 对于查询计划 Q3, 分别针对 Nest-loop, Sort-Merge, 计算自然连接操作需要读取的总块数。（注：对两种连接操作算法，分别考虑 L 在 Date 上有索引和无索引的情形）

(2) 用户有一查询语句：

Select Name

From B, S, L

Where L.S#=S.S# and L. B#=B.B# and Title="Jane Eyre"

检索借阅了书名为 "Jane Eyre" 的学生姓名。

求解下列问题：

- 写出以上 SQL 语句所对应的关系代数表达式，并画出对应的逻辑查询计划树。使用启发式查询优化算法，对该逻辑查询计划树进行优化，并画出优化后的逻辑查询计划树。

b. 设 L 表有 10000 条元组，B 表有 2000 条元组，S 表中有 1000 条元组，L 表中满足借阅书名为“Jane Eyre”的元组数为 50，计算优化前与优化后的查询计划中每一步所产生的中间结果大小。（要求写出具体计算过程）