

## 作业5: 关系数据库存储与查询执行(2020春)

主讲教师: 邹兆年(znzou@hit.edu.cn)

姓名: \_\_\_\_\_ 学号: \_\_\_\_\_

题目	1	2	3	4	5	6	7	8	9	10	总分
得分											

1. (10分) 在PostgreSQL上首先使用下面的SQL语句创建关系 $t$ :

```
CREATE TABLE t (  
  id INT PRIMARY KEY,  
  val CHAR(1024) NOT NULL DEFAULT 'val');
```

然后使用下面的SQL语句插入元组:

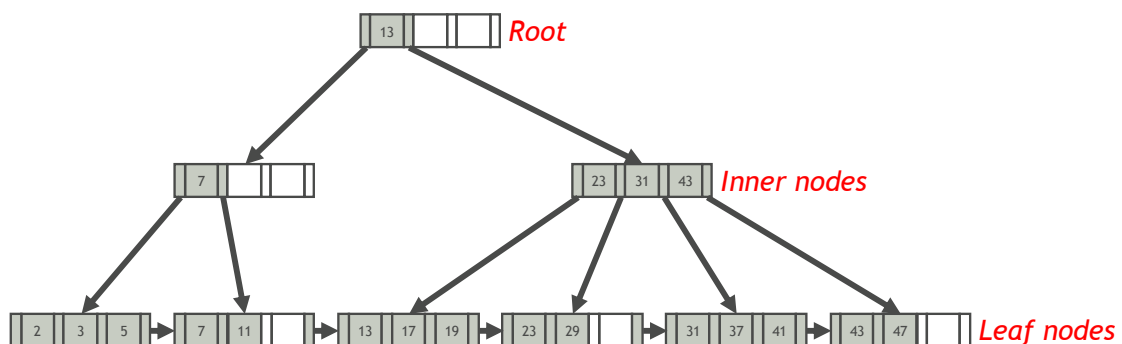
```
INSERT INTO t(id) VALUES (1), (2), (3), (4), (5), (6), (7), (8), (9), (10);
```

执行SQL查询“SELECT ctid, id FROM t;”, 我们得到以下查询结果:

ctid	id
(0,1)	1
(0,2)	2
(0,3)	3
(0,4)	4
(0,5)	5
(0,6)	6
(0,7)	7
(1,1)	8
(1,2)	9
(1,3)	10

其中ctid表示元组的记录号(record ID), 即元组所在页号(page ID)和槽号(slot number)构成的对。请分析该DBMS使用的页的大小是4KB、8KB、16KB中的哪一个? 说明理由。

2. (20分) 已知如下B+树



回答下列问题。

- (a) (10分) 插入键值为35的索引项(index entry)后, 该B+树变成什么样? 请绘制出来。
  - (b) (10分) 删除键值为13的索引项(index entry)后, 该B+树变成什么样? 请绘制出来。
3. (70分) 已知关系 $R(w, x), S(x, y), T(y, z)$ 的块数分别为5000, 10000, 10000。我们准备执行关系代数查询 $(R \bowtie S) \bowtie T$ 。假设缓冲池中有 $M = 101$ 个页可用,  $R, S, T$ 上均无索引且未按连接属性排序。请回答下列问题。
- (a) (10分) 使用什么算法执行 $R \bowtie S$ 最适合? 说明理由。
  - (b) (10分) 使用(a)中选择的算法执行 $R \bowtie S$ 的I/O代价是多少?
  - (c) (10分) 如果 $R \bowtie S$ 的结果不超过49块, 那么在使用(a)中选择的算法执行 $R \bowtie S$ 时,  $R \bowtie S$ 的结果是否需要物化(materialize)到文件中? 说明理由。
  - (d) (10分) 如果 $R \bowtie S$ 的结果不超过49块, 那么使用什么算法将 $R \bowtie S$ 的结果与 $T$ 进行自然连接最合适? 说明理由。
  - (e) (10分) 使用(d)中选择的算法计算连接结果的I/O代价是多少?
  - (f) (10分) 如果 $R \bowtie S$ 的结果大于49块, 那么使用什么算法将 $R \bowtie S$ 的结果与 $T$ 进行自然连接最合适? 说明理由。
  - (g) (10分) 使用(f)中选择的算法计算连接结果的I/O代价是多少?