Dag Truong

Homework 5

10?

12.3

r_1 need 800 blocks $\left(\frac{20000}{25}\right)$

r_2 need 1500 blocks $\left(\frac{45000}{30}\right)$

If M (pages memory) > 80

we can join easily be done in 1500 + 800 = 2300 disk accesses.

If M ≤ 800

a) Nested-loop join.

r_1 is outer relation. $\Leftrightarrow$ 20000 * 1500 + 800

= 30.000.800

r_2 is outer relation = 45000 * 800 + 1500

= 36.001.500

b) Block nested-loop join.

r_1 is outer relation $\Leftrightarrow \left\lceil \frac{800}{M-1} \right\rceil * 1500 + 800$

r_2 ——— $\Leftrightarrow \left\lceil \frac{1500}{M-1} \right\rceil * 800 + 1500$

$$B_S = 1500\ (2\lceil \log_{M-1}(1500/M)\rceil + 2)$$
$$+\ 800\ (2\lceil \log_{M-1}(800/M)\rceil + 2)$$

total $B_S + 1500 + 800$

d) hash join.

if $M > 800$.

cost = $3(1500 + 800) = 6900$

else

cost $= 2(1500 + 800)\lceil \log_{M-1}(800) - 1\rceil + 1500 + 800$.

Additional Exercise:

Assume that we have the following relations:

Puppies(Puppy Number, Puppy Name, Kennel Code)
Kennels(Kennel Code, Kennel Name, Kennel Location)
Tricks(Trick ID, Trick Name)
PuppyTricks(Puppy Number, Trick ID, WhereLearned, Skill Level)

Now say that we have 80,000 puppies, 2,000 kennels, 100 tricks, and puppies know 3 tricks, on the average. There are 5 different "Skill Levels" for each trick. Our pages hold 40 records (same number for all relations). The most common queries to be supported are:

Query1: Given the name of a puppy, find the "Trick Name" and the "Skill Level" for each trick that the puppy has mastered.

Query2: Find the "Puppy Number, Puppy Name", and "Kennel Code, Kennel Name" and "Kennel Location" where the puppy is staying, for each puppy who has mastered a given trick at level 5.

Show which indexes should be used and which indexes should be made as primary indexes to minimize the costs of executing these two queries.

Query 1:

```
Select T.TrickName , PT.Skill Levered
from  Puppies as P
      Trick  as T
      PuppyTricks as PT
Where  P.Puppy name = x
   and T.TrickID  = P.TRickID.
   and P.Puppynumber = PT.Puppynumber.
```

We may index on puppy number or puppyname.
Because Puppy name is not given.
Base on cluster on puppy number

We have:

Unique of tuples = 80000 x 3 = 240000.

skill level tuples = 100 x 5 = 500

if cluster on puppy number.

then we have cost = (240000 / 500)/ 40 = 12 pages.

and the cost with out cluster = (240000 / 500) = 480 page

⇒ We save 468 pages if we use puppy number.

- PuppyName index should be us
- puppyNumber //
- THick ID //

PuppyNumber and trick ID shaub toe primary index.

b)

```
Select P.Number, P.Puppy name
       K.Knellcode, kkennel name. kenel location
from kenels as k , Puppies as P. Puppytrick as Pt
where. Pt.SkillLevel = 5.
       P.Number = Pt number
       P.knellcode = K.kennleddle.
```

- puppyNumber, knell code should be indexed.