

Name: Xiang Gu
EID: xg2847
HW5b

Part A:

15.3.2

$10,000/999 * (999 + 10,000) \approx 10 * 11,000 = 110,000$ disk I/O's

15.3.3

a). $10,000/(M-1) * (M-1 + 10,000) \leq 100,000 \Leftrightarrow M \geq 1112$

16.2.6

a). $\pi_{\{b+c \rightarrow x, c+d \rightarrow y\}}(R(a,b,c) \text{ join } S(b,c,d,e)) =$
 $\pi_{\{b+c \rightarrow x, c+d \rightarrow y\}}(\pi_{\{b,c\}}(R(a,b,c)) \text{ join } \pi_{\{b,c,d\}}(S(b,c,d,e)))$
because b,c,d are the input attributes and join attributes.

b). $\pi_{\{a, b, a+d \rightarrow z\}}(R(a,b,c) \text{ join } S(b,c,d,e)) =$
 $\pi_{\{a, b, a+d \rightarrow z\}}(\pi_{\{a,b,c\}}(R(a,b,c)) \text{ join } \pi_{\{b,c,d\}}(S(b,c,d,e))) =$
 $\pi_{\{a, b, a+d \rightarrow z\}}(R(a,b,c) \text{ join } \pi_{\{b,c,d\}}(S(b,c,d,e)))$
Because a,c,c,d are the input attributes and join attributes.

Part B:

In linear cost model, let c = average seek time, so, cost of reading n blocks = $n*c$

In affine cost model, let c = average seek time for the first block and c' = average rotational latency for each additional block, so, cost of reading n (contiguous) blocks = $c + (n-1)*c'$

1). $B(R) * c = 1000c$

2). $c + (B(R) - 1)c' = c + 999c'$

3). This is an ill-posed question because if $W.c$ is the primary key, then $W.c$ should not have duplicates. Anyway, I will still do it. The general approach is the same for subquestion 3) through 7).

a). 10 blocks

b). 1,000 records

4).

a). 1 block

b). 1 record

5).

a). 1 block

b). 100 records

6).

a). 50 records

b). $\max(500, 1000) = 1000$ records

7).

a). $10,000/(3*10) = 333$ records

b). $10,000/3 = 3333$ records

8). 0 because S and U don't have any shared attributes

9). $100*10,000/10 = 100,000$ records

10). $1,000,000*10,000/10,000 = 1,000,000$ records

11). $1,000,000/3 * 10,000 / 200 = 17,000,000$ records

12). I don't know how to solve this question.