Ramya Krishnan
A20506653

Part 1 : Result size estimation:

Question 1.1 Estimate Result Size

$$T(q) = T(\text{course}) \times \frac{1}{V(\text{course, instructor})} \times \frac{\text{Course\_Max} + 1}{\text{Max} - \text{Min} + 1}$$

$$= 80 \times \frac{1}{50} \times \frac{3 - 1 + 1}{6 - 1 + 1} = \underline{0.8}$$

Question 1.2 Estimate Result Size

$$T(q) = T(\text{Students}) \times \left[1 - \left[1 - \frac{1}{V(\text{Student, Major})}\right]\right] \cdot \left[1 - \frac{1}{V(\text{Student, Credit})}\right]$$

$$= 30{,}000 \times \left[1 - \left[1 - \frac{1}{20}\right] \times \left[1 - \frac{9}{32}\right]\right]$$

$$= 30000 \times \left[1 - \left(\frac{19}{20} \times \frac{23}{32}\right)\right]$$

$$= 30000 \times 0.3171$$

$$= 9515.625$$

$$= \underline{9515}$$

Question 1.3 : Estimate Result size.

$$T(q) = \frac{(5-4+1) \times T(\text{Course})}{\text{Max}(\text{Course}, \text{Credits}) - \text{Min}(\text{Course}, \text{Credits}) + 1}$$

$$= \frac{2 \times 80}{6-1+1} = \frac{160}{6} = \underline{26.66}$$

Question 1.4 : Estimate Result size

$$q_1 = \sigma_{\text{credit} \geq 20}(\text{Student})$$

$$T(q_1) = \frac{\max(\text{Students}, \text{credits}) - 20 + 1}{\max(\text{Students}, \text{credits}) - \min(\text{Students}, \text{Credits}) + 1} \times T(\text{Student})$$

$$= 30000 \times \frac{32 - 20 + 1}{32 - 1 + 1} = \frac{13}{32} \times 30000$$

$$= \underline{12187.5}$$

$$T(q) = \frac{T(q_1) \cdot T(\text{registered}) \cdot T(\text{course})}{\max(V(q_1, \text{cwID}), V(\text{registered}, \text{student})) \cdot \max(V(\text{registered}, \text{course}) \cdot V(\text{course}, \text{title}))}$$

$$= \frac{12187 \times 10000 \times 80}{\max(12187, 3000) \times \max(30, 80)}$$

$$= \frac{9749600000}{974960} = \underline{10000}$$

Part 2 I/O Cost Estimations

Question 2.1

$B(S) = 4000$ blocks

$S(R) = 1/30$

$S(S) = 1/20$

$B(R) = 300,000$

$M = 101$

$$T(R) = \frac{B(R)}{S(R)} = 300000 \times \frac{1}{1/30}$$

$$= 9,000,000$$

$$T(S) = \frac{B(S)}{S(S)} = \frac{4000}{1} \times \frac{1}{1/20}$$

$$= 80,000$$

(d).

Tuple-based nested-loop join :-

clustered relations

Total I/o cost = $B(R) + T(R) \times B(S)$

$= 300000 + [9000000 \times 4000]$

$= 300000 + 36000000000$

$= 36,000,300,000$

Block-nested-loop join :-

S in the outer loop.

S = 100 1 = R

Read R = 9000000

Read chunk = 100×20

$= 2000$ IO

Number of chunks in S = $\frac{B(S)}{m-1} = \frac{4000}{100} = 40$

Total Io cost = $40 [2000 + 9000000]$

$= 40 [9002000]$

$= 360080000$

Merge-join :-

i/p are not sorted and non-clustered.

R : 3 pass multiway sort as $B(R) \leq m(m-1)$ and R not clustered

3 way Sort cost $= T(R) + 5 * B(R)$

S : 2 pass multiway Sort as $B(S) \leq m(m-1)$

S not clustered

To read & Sort $= T(S) + 3 \times B(S)$

join cost for both R & S $= B(R) + B(S)$

[redacted]

Total I/O cost : Sort Cost + join cost

$= T(R) + 5 \times B(R) + T(S) + 3 \times B(S) + B(R) + B(S)$

$= T(R) + 6\, B(R) + T(S) + 4\, B(S)$

[redacted]

$= 900000 + 6 \times 300000 + 80000 + 4 \times 4000$

$= 10896000$

Non-clustering Index join :-

Relations are Clustered

$$I/O = B(S) + T(S) \times \frac{T(R)}{V(R,C)}$$

$$\frac{T(R)}{V(R,C)} = 120$$

$$I/O = 4000 + 80000(120)$$

$$= 9604000$$