

Άσκηση 3η
(αυστηρή) Ημερομηνία Παράδοσης: Πέμπτη, 30/05/2024, 16.59

ΘΕΜΑ 1^ο

Έχετε την ερώτηση $\sigma_{R,C>40} (R \bowtie_{R.A=S.B} S)$ και τις παρακάτω πληροφορίες.

- Η σχέση R έχει 150.000 πλειάδες και 30 πλειάδες ανά σελίδα.
- Η σχέση S έχει 100 πλειάδες και 20 πλειάδες ανά σελίδα.
- Το B είναι πρωτεύον κλειδί και το A ξένο κλειδί σ' αυτό.
- Δεν υπάρχουν ευρετήρια ή ταξινόμηση των σχέσεων στο δίσκο
- Διαθέτουμε 52 buffers στην κύρια μνήμη
- Η σχέση R έχει το εξής ιστόγραμμα για την στήλη C: (α) για τιμές στο εύρος 0-50, το 40% των εγγραφών της σχέσης, (β) για τιμές από 51-65 το 45% των εγγραφών και (γ) για τιμές από 66-80 τις υπόλοιπες εγγραφές.

(α) Εκτιμήστε το μέγεθος του αποτελέσματος και τον απαιτούμενο χώρο στο δίσκο για να γραφτεί το αποτέλεσμα αυτό. Προσοχή: δεν σας δίνονται άλλες πληροφορίες (π.χ., εξωτερικά κλειδιά) και άρα πρέπει να σκεφθείτε μια λογική εκτίμηση για τη χειρότερη περίπτωση ☺

Αιτιολογήστε το ακριβές, ελάχιστο και μέγιστο μέγεθος του αποτελέσματος της ερώτησης αν

- το B είναι πρωτεύον κλειδί και το A ξένο κλειδί στο B,
- το B είναι πρωτεύον κλειδί
- τίποτε από τα προηγούμενα δεν ισχύει

(β) Δώστε δύο αλγεβρικά πλάνα για την εκτέλεση της πράξης και διαλέξτε ένα εξ' αυτών που θα χρησιμοποιήσετε για τα επόμενα ερωτήματα. Αιτιολογήστε αν μπορείτε να κάνετε pipelining αποτελεσμάτων ή αν πρέπει να αποθηκεύσετε ενδιάμεσα αποτελέσματα.

(γ) Υπολογίστε το κόστος σε I/O της εν λόγω ερώτησης, με όποιο αλγόριθμο κρίνετε σκόπιμο για την πράξη της επιλογής, αλλά για κάθε έναν από τους παρακάτω τρόπους εκτέλεσης της πράξης σύνδεσης:

- Page-based nested loops join
- Block nested loops join
- Sort-Merge join
- Hash join

(δ) Υπολογίστε το ελάχιστο δυνατό κόστος για την εκτέλεση της ερώτησης και υπό ποιες συνθήκες αυτό μπορεί να επιτευχθεί.