

Άσκηση 3η
(αυστηρή) Ημερομηνία Παράδοσης: 26/05/2026, 16.59

ΘΕΜΑ 1°

Έχετε την ερώτηση $\sigma_{R.C>10}(R \bowtie_{R.A=S.B} S)$ και τις παρακάτω πληροφορίες.

- Η σχέση R έχει 100.000 πλειάδες και 30 πλειάδες ανά σελίδα.
- Η σχέση S έχει 2000 πλειάδες και 10 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.

(α') Αιτιολογήστε το ακριβές, ελάχιστο και μέγιστο μέγεθος του αποτελέσματος της ερώτησης αν, αντί για το ότι το join είναι ένα key-foreign key join, το μόνο που ξέρουμε είναι

(α'1) το B είναι πρωτεύον κλειδί, ή

(α'2) δεν ξέρουμε τίποτα σε σχέση με τα κλειδιά

Στις επόμενες υποερωτήσεις, συνεχίζουμε να θεωρούμε ότι έχουμε ένα key-foreign key join.

(β) Δώστε ένα πλάνο που θα έβγαινε χωρίς βελτιστοποιητή, καθώς και ένα αλγεβρικό πλάνο που βγαίνει από βελτιστοποιητή. Αιτιολογήστε τις διαφορές.

Στο πλάνο που θα έβγαινε από τον βελτιστοποιητή, 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, 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) Κάτω από ποιες συνθήκες μπορεί να έχουμε το ελάχιστο δυνατό κόστος για την εκτέλεση της ερώτησης? Ποιο είναι αυτό το κόστος?