

1^η Εργασία στη Σχεδίαση ΒΔ 2018-2019

Ευρετήρια

Στόχος των εργασιών είναι η εξοικείωση με θεωρητικά και πρακτικά προβλήματα των Βάσεων Δεδομένων, μέσα από χρηστικά παραδείγματα. Στην πρώτη εργασία θα ασχοληθούμε με τις εντολές δημιουργίας ευρετηρίων και με τον υπολογισμό των χαρακτηριστικών ενός ευρετηρίου.

A – Ερωτήματα

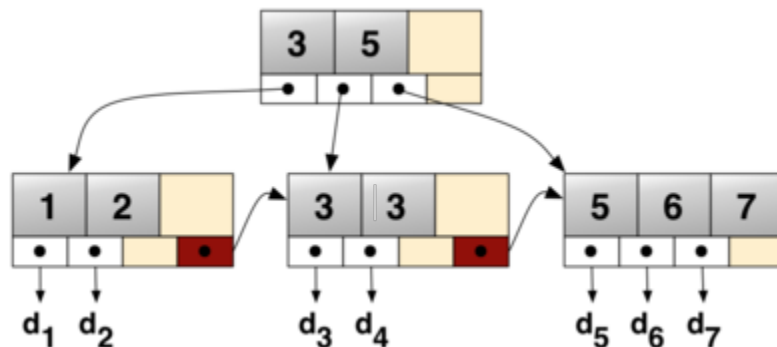
1^ο μέρος – Εντολές δημιουργίας ευρετηρίων

Βρείτε τη σύνταξη της εντολής CREATE INDEX σε ORACLE, MySQL και PostgreSQL. Εξηγείστε τις παραμέτρους κάθε εντολής σε σχέση με τους διαθέσιμους τύπους ευρετηρίου που έχουμε δει στο μάθημα. Παρουσιάστε συγκριτικά (σε ένα πίνακα) τις δυνατότητες των τριών ΣΔΒΔ, ποια ευρετήρια υποστηρίζουν και ποια όχι η κάθε μια.

Αν χρησιμοποιήσατε πηγές από το δίκτυο να τις παραθέσετε (URL διευθύνσεις).

2^ο μέρος – Ευρετήριο με B+δένδρο

Ένα αρχείο έχει 20.000.000 εγγραφές αταξινόμητες σε σωρό. Οι εγγραφές έχουν κατά μέσο όρο μέγεθος 250 bytes και αποθηκεύονται σε αρχεία σε block μεγέθους 2048 bytes με μη εκτατή καταχώρηση.



Πάνω σε ένα αριθμητικό πεδίο (μεγέθους 8 bytes) που μπορεί και να περιέχει διπλότυπα, έχει χτιστεί ευρετήριο με χρήση B+ δένδρου. Κάθε δείκτης προς τις εγγραφές του αρχείου (δείκτες d στην εικόμα) έχει μέγεθος 16 Bytes, κάθε δείκτης προς block του ευρετηρίου μαζί και ο δείκτης προς επόμενο φύλλο έχει μέγεθος 8 Bytes. Το ευρετήριο δημιουργείται μαζικά για όλο το αρχείο και μπορείτε να θεωρήσετε ότι οι κόμβοι στο τελευταίο επίπεδο του δένδρου είναι όσο γίνεται πλήρεις.

Υπολογίστε:

- 1) Το μέγεθος του αρχείου αν είναι αρχείο σωρού.
 - 2) Το μέγεθος του αρχείου αν είναι αρχείο κατακερματισμού.
 - 3) Πόσα επίπεδα έχει το B+ δένδρο συμπεριλαμβανομένου και του τελευταίου επιπέδου
 - 4) Πόσους κόμβους θα περιέχει το κάθε επίπεδό του, ποιο το μέγεθος του ευρετηρίου συνολικά.
 - 5) Αν το δένδρο σας γίνει B* ποια η απάντηση στο 4;
- Εξηγείστε τη διαδικασία επίλυσης σε κάθε βήμα.
- 6) Ποιο θα είναι το κόστος αναζήτησης ισότητας για μια συγκεκριμένη τιμή που γνωρίζετε ότι εμφανίζεται 5 φορές σε όλο το αρχείο;

B – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των 3 ατόμων (το πολύ), αν και επιτρέπονται μικρότερες ομάδες. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους ΑΜ των μελών της ομάδας: π.χ. **AM1-AM2-AM3.zip**

- Το zip θα περιλαμβάνει:
 - ένα αρχείο readme.txt
 - με τα ονοματεπώνυμα και τους ΑΜ των φοιτητών της ομάδας
 - το αρχείο pdf με την τελική εργασία

Γ – Άλλες Οδηγίες

Όσες εργασίες δεν τηρούν τις οδηγίες παράδοσης, θα έχουν επίπτωση στο βαθμό.

Όσες εργασίες κριθούν ότι είναι αντιγραφές θα μηδενίζονται.

Ημερομηνία παράδοσης: **Στο e-class με οριστική τελική ημερομηνία 13-11-2018**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 15-11-2018 θα έχουν μείωση 2 μονάδων στο βαθμό.