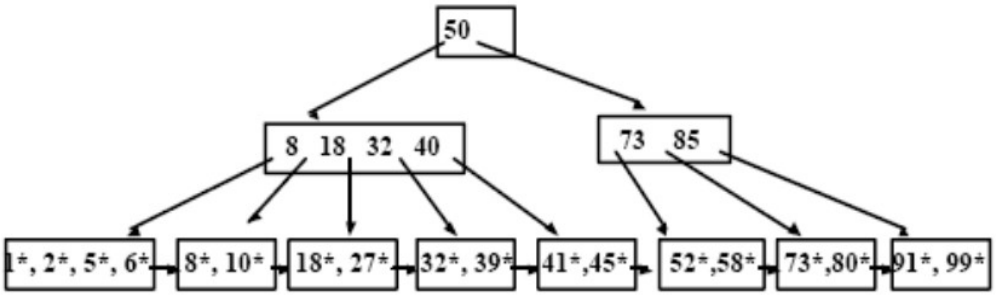
ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΙΙ – Ασκήσεις με B+trees

ΑΣΚΗΣΗ 1. Θεωρείστε το παρακάτω B+tree (κάθε κόμβος ευρετηρίου χωρά 4 καταχωρίσεις ευρετηρίου και κάθε κόμβος φύλλο χωρά 4 καταχωρίσεις δεδομένων):



1.1 Ποιο είναι το μικρότερο πλήθος διαγραφών που πρέπει να γίνουν για να μειωθεί το ύψος του δέντρου κατά 1;

3

1.2 Δώστε μια τέτοια ακολουθία διαγραφών:

Υποθέτουμε ότι κατά τη διαγραφή πρώτα γίνεται προσπάθεια ανακατανομής με το αριστερό ή το δεξιό αδέλφι κατά σειρά, και μετά συγχώνευσης με το αριστερό ή το δεξιό αδέρφι κατά σειρά.

(α) διαγραφή του 18*, συγχώνευση με τον αριστερό κόμβο, νέος κόμβος (8*, 10*, 27*), διαγραφή του 18 από τον κοινό γονιό τους.

(β) διαγραφή του 41*, συγχώνευση με τον αριστερό κόμβο, νέος κόμβος (32*, 39*, 45*), διαγραφή του 40 από τον κοινό γονιό τους. Ο γονιός τώρα είναι ο (8, 32).

(γ) διαγραφή του 80*, συγχώνευση με τον αριστερό κόμβο, νέος κόμβος (52*, 58*, 73*), διαγραφή του 73 από τον κοινό γονιό τους. Ο γονιός γίνεται (85) πέφτει κάτω από το 50% οπότε συγχωνεύεται με το αδέρφι του (8, 32) καθώς ανακατανομή δεν μπορεί να γίνει. Η ρίζα δεν έχει λόγο ύπαρξης καθώς μένει με ένα παιδί, οπότε το 50 κατεβαίνει και δημιουργείται η νέα ρίζα (8, 32, 50, 85) με πέντε παιδιά.

1.3 Ποιο είναι το μικρότερο πλήθος εισαγωγών που πρέπει να γίνουν ώστε ο κόμβος με τις τιμές 73 και 85 να γεμίσει;

5

1.4 Δώστε μια τέτοια ακολουθία εισαγωγών;

Για να γεμίσει ο κόμβος αυτός πρέπει να δεχτεί δυο καταχωρίσεις ευρετηρίου. Αυτό θα συμβεί όταν γίνουν δυο διασπάσεις στα παιδιά του (δηλαδή αποκτήσει άλλα δυο παιδιά). Εισάγουμε τα 75*, 79* και 83* στον (73*, 80*), ο οποίος υπερχειλίζει και διασπάται στους (73*, 75*, 79*) και (80*, 83*) [υποθέτουμε ότι κατά τη διάσπαση περιττού αριθμού καταχωρίσεων η έξτρα καταχώριση πηγαίνει αριστερά]. Ανεβαίνει το 80 στο γονιό και τώρα είναι ο (73, 80, 85). Συνεχίζουμε με εισαγωγή των 76* και 78* στον (73*, 75*, 79*), ο οποίος υπερχειλίζει και διασπάται στους (73*, 75*, 76*) και (78*, 79*). Ανεβαίνει το 78* στο γονιό ο οποίος τώρα γίνεται (73, 78, 80, 85).

ΑΣΚΗΣΗ 2. Ένα B+-tree έχει εσωτερικούς κόμβους και κόμβους-φύλλα που η μέγιστη χωρητικότητα τους είναι 2 κλειδιά αναζήτησης και 3 δείκτες. Σχεδιάστε το τελικό B+-tree που προκύπτει από τη μαζική φόρτωση των τιμών 3, 30, 44, 2, 14, 4, 20 σε ένα άδειο τέτοιο δέντρο.

Ταξινομούμε τις καταχωρίσεις δεδομένων: 2*, 3*, 4*, 14*, 20*, 30*, 44*. Δημιουργούνται 4 κόμβοι φύλλα: (2*, 3*), (4*, 14*), (20*, 30*), (44*). Προσθέτουμε σταδιακά από τα αριστερά προς τα δεξιά τα φύλλα στο κενό δέντρο. Τα τρία πρώτα φύλλα ευρετηριοποιούνται από ένα νέο γονιό που γίνεται και η ρίζα του δέντρου (4, 20). Η προσθήκη του τέταρτου φύλλου, διασπά τη ρίζα στους κόμβους (4) και (44) και δημιουργεί τη νέα ρίζα (20).