Principle of DBMS Fall 2023 COP 5725

Final Exam

[30 points] Problem 1. Disk Storage

(a) Describe the process of reading one sector from the disk.

(b) In the procedure described in (a), which step is the most time-consuming (on average)?

(c) What is the average rotational delay for a 15000 RPM disk? Express your answer in milliseconds. (RPM: revolutions per minute, which means the number of rotations in one minute)

(d) What can we do to reduce the rotational delay?

[20 points] Problem 2. File Organization

(a) Consider the following record: <120, 'Jeffrey'>. Its data type is <int, varchar>. Now we want to store this record as a variable length record in the file. Draw a diagram that shows what content will be stored in which bytes.

(b) Consider the following record: <0, 'student', 'book', 10>. Its data type is <int, varchar, varchar, int>. Now we want to store this record as a variable length record in the file. Draw a diagram that shows what content will be stored in which bytes.

[50 points] Problem 3. B+ tree

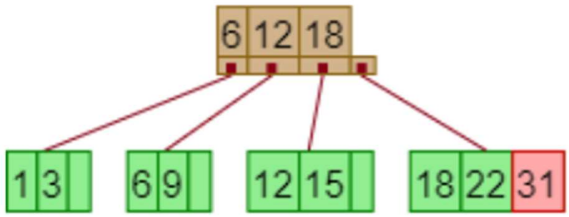

Figure 1

a) (10 points) The B+ tree in figure 1 has order=4 and height=2.
What would the tree's height be after executing the following operations? (Assume each operation is executed independently of each other, i.e., before "insert 5, then insert 7", the tree is exactly as shown in figure 1)

- Insert 17, height of the tree = _____
- Insert 5, then insert 7, height of the tree = _____
- Insert 32, height of the tree = _____
- Insert 32, then delete 32, height of the tree = _____
- Insert 32, then delete 32, then delete 31, height of the tree = _____

b) (10 points) The B+ tree in figure 2 has an order of 4. Assume when storing the B-tree, a disk error happened, causing search-key 5 in the root node to be replaced by 12 (see the right tree in figure 2). The other part of the tree remains unchanged. Obviously, after this error, the tree is in an invalid state.

If we perform searches on this tree, searching which of the following values will return the wrong result (i.e., the search algorithm will say value doesn't exist, but it actually exists)? List of values: 4, 5, 9, 11, 17.

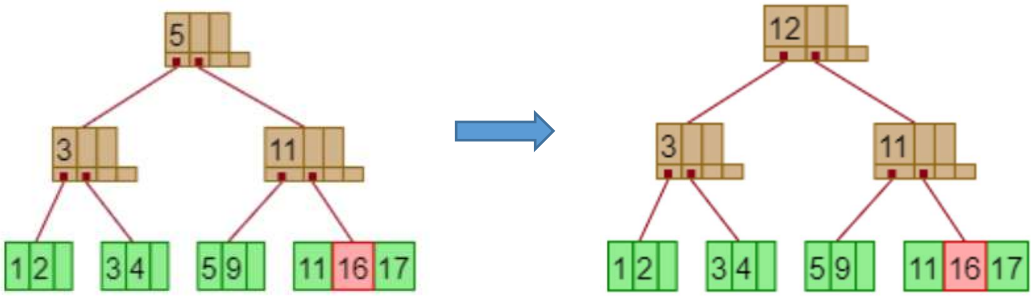

Figure 2

c) In what situation, **deleting** one value is guaranteed to **decrease** the height of the tree by one? Make sure your explanation applies to B+ tree of **any order and any height.** (15 points)

d) The following B+ tree has an order of 4 (see figure 3). Now let's delete random values from the tree until the **height of the tree decreases from 3 to 2**. For example, we delete 3, 5, 7, 9, 17. And only after 17 is deleted, the tree decreases height from 3 to 2. Let X denote the number of values deleted (in this example, X=5). Let S denote the list of deleted values (in this example, S=3,5,7,9,17).

If we start over (restore the tree) and delete different values (meaning S is different), X may still be 5, or a different value.

- Give two more examples in which after the last value is deleted, the tree will decrease height from 3 to 2 (you only need to write down S and X). (5 points)
- What is the minimum possible value of X? Explain your reasoning. (5 points; no points if no reasoning is presented)
- What is the maximum possible value of X? Explain your reasoning. (5 points; no points if no reasoning is presented)

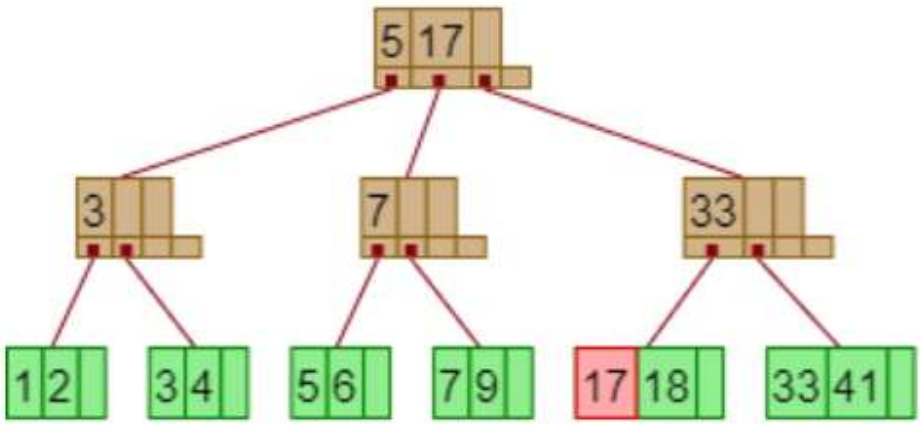

Figure 3. B+ tree of order 4