

Quiz 5, Fall 2019
CSCI 4380 DB Sys
Time: 20 minutes

Question 1 (10 points). You are given the following information for a relation R. Note that a B-tree index stores key values and pointers to other nodes.

TUPLES(R)=10,000,000

PAGES(R)=5,000

A disk page has 8,000 bytes of data available for storing index information

Sizes of attributes for R:

A: 8 bytes, B: 14 bytes, C: 20 bytes, pointer: 8 bytes

What is the size of a B-tree index on R(B,C,A)? How many nodes in each level of the index? You can make some simplifying assumptions by disregarding +/-1 additional pointer in each node and assume each node stores maximum amount of key value and pointer pairs in each node. Show your work.

Answer. One entry takes: $8+14+20+8=50$ bytes.

One disk page can store a max of $8000/50=160$ entries.

Leaf level: $10,000,000/160 = 62,500$ leaf nodes

Next level: $62500/160 = 391$ nodes

Next level $391/160=3$ nodes

Next level 1 node (root)

Question 2 (15 points). Suppose you are given an index on R(B,C,A) such as the one you calculated in the previous question. The actual size of the index is irrelevant to this question once you write your assumption. For each of the queries below, fill in the table of how the index is searched for this query: (a) how many nodes from internal nodes and (b) which condition is used to scan the leaf nodes.

Write your assumption first: How many levels in the B-tree?

Fill the following:

Query	# internal nodes scanned	Condition used to scan leaf nodes
SELECT * FROM R WHERE B='SEA' AND C='breeze';	1 node from each level	B='SEA' AND C='breeze'
SELECT * FROM R WHERE B='SEA' AND A=10;	1 node from each level	B='SEA'
SELECT * FROM R WHERE B='SEA' AND C>='breeze' AND C<='wind';	1 node from each level	B='SEA' AND C>='breeze' AND C<='wind';
SELECT * FROM R WHERE B>='LAND' AND B<='breeze'; AND C='wind'	1 node from each level	from B>='LAND' AND C='wind' to B<='breeze' AND C='wind'
SELECT * FROM R WHERE A=10;	1 node from each level	TRUE (or all nodes)