**DSCI 551 HW4**

**Tianze Sun 9612572343**

1.

(a)

read the root node, find the 1st internal node.

read the 1st internal node, find the 5th leaf node.

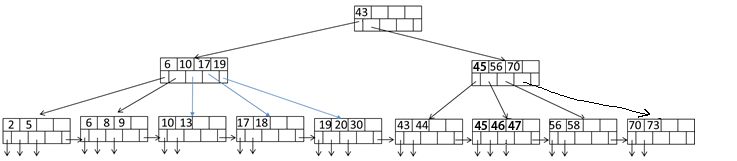
read the 5th leaf node, find the start point 20.

sequential traversals of leaves until 50, so we read 6th and 7th leaf node.

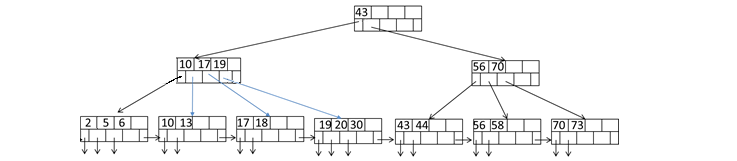
The end point 50 is less than the first data in 7th leaf node, so we stop.

Totally, we did 5 reads and 0 writes, which means 5 block I/O’s.

(b)



(c)



2.

(a)

R ⋈ S

for each 100 blocks br of R do

for each block bs of S do

for each tuple r in br do

for each tuple s in bs do

if r and s join then output (r, s)

Total Cost: B(R) + B(R)B(S)/100 = 505,000 Block I/O’s.

(b)

R ⋈ S

for each 100 blocks bs of S do

for each block br of R do

for each tuple s in bs do

for each tuple r in br do

if r and s join then output (r, s)

Total Cost: B(S) + B(S)B(R)/100 = 510,000 Block I/O’s.

(c)

For B(R)+B(S) > M2

Sort R: in two passes 4B(R) = 20,000

Sort S: in two passes 4B(S) = 40,000

merge: B(R)+ B(S)=15,000

So total cost: 5B(R) + 5B(S) = 75,000 Block I/O’s

(d)

For Min (B(R), B(S)) < M2

So total cost: 3B(R) + 3B(S) = 45,000 Block I/O’s.