Problem 1.

Consider a hard disk that uses Advanced Format with sector size of 4096 bytes (4KB), 1000 tracks per surface, 50 sectors per track, five (5) double-sided platters, and average seek time of 8 msec. Suppose that the disk platters rotate at 7200 rpm (revolutions per minute).

Suppose also that a block (page) of 4096 bytes is chosen. Now, consider a file with 150,000 records of 100 bytes each that is to be stored on such a disk and that no record is allowed to span two blocks. Also, no block can span two tracks.

1. How many records fit onto a block?
2. How many blocks are required to store the entire file? If the file is arranged "sequentially on the disk, how many cylinders are needed?
3. How many records of 100 bytes each can be stored using this disk?
4. What time is required to read a file containing 100,000 records of 100 bytes each sequentially? You can make an assumption about how long it takes moving the heads from one cylinder to the next here.
5. What is the time required to read a file containing 100,000 records of 100 bytes each in a random order? To read a record, the block containing the record has to be fetched from the disk. Assume that each block request incurs the average seek time and rotational delay. You need also to include the transfer time for each block.

Problem 2.

Consider a B+-tree of order two (d=2). Thus, the maximum number of pointers per node is 5 and the maximum number of entries is 4.

1. Show the results of entering one by one the keys that are three letter strings: (era, ban, bat, kin, day, log, rye, max, won, ace, ado, bug, cop, gas, let, fax) (in that order) to an initially empty B+-tree. Assume that you use lexicographic ordering to compare the strings. Show the state of the tree after every 4 insertions.
2. What is the utilization of the tree? The utilization of the tree is defined as the total number of entries in all the nodes of the tree (both leaf and non-leaf nodes) over the maximum number of entries that the same nodes can store.

SOLUTIONS:

Problem 1

1. 4096 / 100 = 40.96. Which means that 40 records can fit in a block.

2. The file requires 150,000 records / 40 records per block = 3750 blocks to be saved. In our case one block is equal to one sector(4KB)
1 track holds 50 blocks
10 tracks (two for each platter), which is one cylinder holds 500 blocks.

The file is 3750 blocks. Thus 3750/500 = 7,5 cylinders. So 8.

3. 1 block can hold 40 records.

The disk has 50*1000*10 number of blocks (blocks per track * tracks * platters)

Thus, this disk can store 40*50*1000*10 = 20M pages.

4. A file containing 100,000 records of 100 bytes is stored in 100,000/40 = 2500 blocks.

From answer 1, we have that 1 cylinder holds 500 blocks. Our file will be saved in 2500/500 = 5 cylinders.

In order to read one track, we need 60s/7200(rpm) = 0,0083 seconds. In order to read one cylinder, we need 10 * 0,0083 = 0,083 seconds. We need to read 5 cylinders, so the total time will be 5 * 0,083 = **0.415 seconds**. We assume that the time for moving from one cylinder to another is very small. The average seek time is 8msec = 0,008 seconds and the average rotational delay is 0.00415 seconds. Therefore, the total time is 0,415 + 0,00415 + 0,008 = 0,42715 seconds.

5. A file containing 100,000 records of 100 bytes is stored in 100,000/40 = 2500 blocks

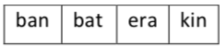
For any block of data, average access time =avg seek time +avg rotational delay + transfer time
Average seek time = 0,008s s and average rotational-delay = 0,00415s .
To find transfer time:

Each track has 50 blocks. To read all 50 blocks it takes 0,0083 seconds. To read just one takes 0,0083/50 = 0,000166 seconds

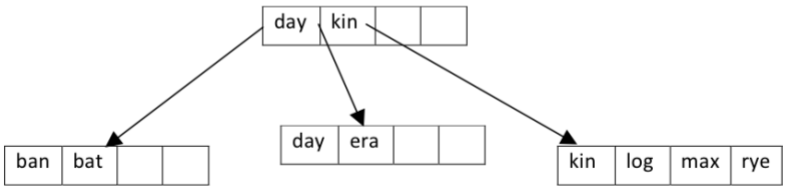
As a result, the total time is (0,008 + 0,00415 + 0,000166) * 2500 = 30.79 seconds.

Problem 2

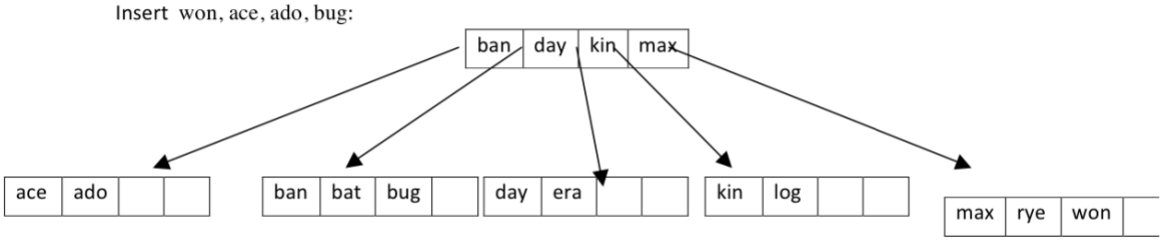
1. Insert era, ban, bat, kin:



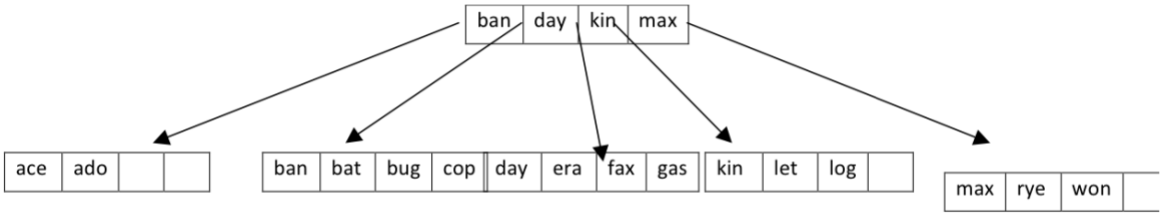
Insert day, log, rye, max:



Insert won, ace, ado, bug:



Insert cop, gas, let, fax:



2. The utilization of the tree is 20 / 24 = 0.833