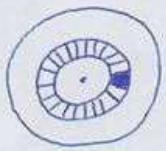
Database Systems: 1st Mid

Date: 9th September 2016 **Duration: 1.5 hrs**

1. **No clarifications during the exam.**
2. **Make *reasonable assumptions* and *clearly state* them to answer *ambiguous* questions.**
3. **Show your steps. Be concise and organized.**
4. **Calculators allowed. Sharing of calculators *not* allowed.**

1) The Megatron 777 disk has the following characteristics:

1. There are 10 surfaces, with 10,000 tracks each.
2. Tracks hold an average of 1000 sectors of 512 bytes each.
3. 20% of each track is used for gaps.
4. The disk rotates at 10,000 rpm.
5. The time it takes the head to move n tracks is $1+0.001n$ milliseconds.



(a) What is the capacity of the disk?
(b) If all tracks hold the same number of sectors, what is the density of bits in the sectors of a track?
(c) What is the maximum seek time?
(d) What is the maximum rotational latency?
(e) If a block is 16,384 bytes (i.e. 32 sectors), what is the transfer time of a block? **[10]**

2) We have a 1 GB sized relation R of 10,000,000 tuples. Each tuple of 100 bytes has several fields, one of which is the *sort key* field, which may not be a primary key. The machine on which sorting occurs has one Megatron 777 disk (described above in Q1) and 50 MB of main memory available. Disk blocks are 4096 bytes. How long would it take to sort R using 2-phase, multiway merge sort. **[10]**

3) Suppose we are using RAID level 4 (i.e. 1 redundant disk for parity), with 4 data disks and 1 redundant disk. For simplicity, assume blocks are a single byte.
(a) Give the block of the redundant disk if the corresponding blocks of the data disks are: 01010110, 11000000, 00111011, and 11111011.
(b) Data disk 1 has failed. Recover the block of that disk if the contents of disks 2 through 4 are:
01010110, 11000000, and 00111011, while the redundant disk holds 11111011. **[10]**

4) Suppose we swizzle all pointers automatically, we can perform the swizzling in half the time it would take to swizzle each one separately. If the probability that a pointer in main memory will be followed at least once is p, for what values of p is it more efficient to swizzle automatically than on demand? **[10]**

5) Suppose blocks hold either 3 records, or 10 key-pointer pairs. As a function of n, the number of records, how many blocks do we need to hold a data file and:
(a) A dense index? (b) A sparse index? **[10]**