Sample solutions to midterm #1:

Page 2, Question 1:
- Although the text is not all the same size, if most fields are almost the maximum size of the field, then it is possible that a fixed length field will take up less space than the field plus the 2 bytes (or whatever) for the length field (and possibly any overhead bytes on the page).

Page 2, Question 2:
- No, it means that 2 writes are required, not one. They may be staggered, but not necessarily.

Page 2, Question 3:
x1 = 00001111 (x6 XOR x2 XOR x4)
x3 = 00110011 (x7 XOR x1 XOR x4)

Page 3
(a) Avg. seek is 1/3 of the max. possible range = 1 ms (setup) + 1/3 * (4500 cyl/300 cyl) * 1 ms = 1 ms + 15/3 ms = 6 ms
... Transfer time = 1 cyl * 15 tracks/cyl * 1 rev/track * 1 min/5400 rev * 60 sec./1 min. * 1000 ms/1 sec. = 166.67 ms
... Sum of the above = 172.67 ms

(b) 1 cyl * 15 tracks/cyl * 256 sectors/track * 1 page / (4096 bytes/ (512 bytes/sector)) = 480 pages/cyl
... thus, ceiling (2000 pages / (480 pages/cyl)) = 5   (OK if you said 4.16 cyl)

Page 4
Q. 3 – Less elapsed time is likely to be needed since there will be far fewer I/O's (pages to transfer) under the compressed setup (and it is reasonable to assume that CPU time << I/O time).

Q. 4 – ceiling(65,000,000 records / (325 records/page)) = 200,000 pages

pass 0 = ceiling(200,000 pages / (400 pages/run)) = 500 sorted runs of 400 pages each
pass 1 = ceiling(500 / (400-1)) = 2 sorted runs (1 is 399*400 pages = 159,600 pages long, and the other is 101 * 400 pages = 40,400 pages long)
pass 2 = ceiling(2 /(400-1)) = 1 sorted run of 200,000 pages

Page 5

# of pages in 1 cyl = (1 MB/cyl) / (4K/page) = (2^20 / 2^12) pages/cyl = 2^8 pages/cyl = 256 pages/cyl

ceiling(204,800 pages / (256 pages/cyl)) = 800 cyl of data

Phase 1:
Transfer time = 2 * 204,800 pages * 0.5 ms/page = 204,800 ms = 204.8 sec.

Seek time = ceiling(800 cyl / (32 cyl/fill)) = 25 fills (do seeks on both input & output)

⇨ 2 * 25 fills * (10 ms for first cyl. + 1 ms for next 31 cylinders)
⇨ 2 * 25 * (10 + 31) ms = 2050 ms for seeks

Therefore, phase 1 time is the sum: 204.8 sec. + 2.05 sec. = 206.85 sec.

Phase 2:
Transfer time = 204.8 sec., as above

Seek time = sum of:
a) 10 ms to find the first sorted run's first cyl. + (1 ms + (32/300)ms) for the next 24 sorted runs' first cylinders)
b) 25 sorted runs * 31 cylinders remaining * 10 ms/seek (assume random from here on)
c) writing output: 800 cyl * 10 ms (assume writes take place at unknowable times)
Thus, sum is (10 + 26.64 + 25(31)(10) + 800(10)) ms = 15.79 sec.

Overall, the sum of Phase 1 and Phase 2, with seeks considered, is 206.85 + 220.59 = 427.44 sec.