2004 Paper 12 Question 1

Data Structures and Algorithms

It is proposed to store a large number of records on a disk using Larsen's method so that any lookup can be done using only one disk transfer. All the records are of length 200 bytes and each contains a 20 byte key. The data is to be held on a single disk preformatted to contain 100,000,000 sectors each of size 4096 bytes. Reading multiple consecutive sectors is regarded as a single transfer.

(*a*) Describe Larsen's algorithm in detail and, for the records and disk specified above, state the disk block size, the signature size and the amount of main memory that you would choose to use.   [10 marks]

(*b*) Carefully estimate the maximum number of records that could reasonably be stored on the disk assuming the sizes you gave in part (*a*).   [6 marks]

(*c*) Discuss the advantages and disadvantages of different signature sizes.
                                                                          [4 marks]