

## Συμπεράσματα και περιγραφή των τρόπων οργάνωσης των αρχείων:

Ονοματεπώνυμο: Κωνσταντίνος Νικολός  
ΑΜ: 2019030096

### Α τρόπος Οργάνωσης:

Στο συγκεκριμένο τρόπο οργάνωσης χρησιμοποιούμε ένα αρχείο ,με blocks μεγέθους 128bytes αποτελούμενα από 4 records (Κλάση MainRecord) μεγέθους 32bytes.Κάθε record περιέχει στην 1<sup>η</sup> γραμμή το κλειδί key της εγγραφής(4bytes) και την πληροφορία σε ένα letter(28bytes).

### Β τρόπος οργάνωσης:

Στο συγκεκριμένο τρόπο οργάνωσης χρησιμοποιούμε δύο αρχεία όπου στο 1<sup>ο</sup> αποθηκεύονται οι εγγραφές μεγέθους 8 byte (κλάση PointerRecord) για την αποθήκευση των ζευγών κλειδιά-δείκτες , ενώ στο δεύτερο αποθηκεύονται εγγραφές 32bytes που περιέχουν το κλειδί στην 1<sup>η</sup> γραμμή και τα δεδομένα (κλάση MainRecord ).

### Γ-Δ τρόπος οργάνωσης:

Η δομή των αρχείων παραμένει ίδια με τον Α τρόπο οργάνωσης με βασική διαφορά ότι αφού αποθηκευτούν τα αρχεία ταξινομούνται με βάση τα κλειδιά τους.Το ίδιο ισχύει και για τον Δ τρόπο καθώς η δομή του παραμένει η ίδια με τον Β τρόπο οργάνωσης ενώ ταξινομείται το αρχείο των κλειδιών.

## Συμπεράσματα

Παράδειγμα αποτελεσμάτων:

Μέθοδος	Α τρόπος οργάνωσης αρχείου	Β τρόπος οργάνωσης αρχείου	Γ τρόπος οργάνωσης αρχείου	Δ τρόπος οργάνωσης αρχείου	Απόδοση Εξωτερικής Ταξινόμησης τρόπος Γ	Απόδοση Εξωτερικής Ταξινόμησης τρόπος Δ
Disk Reads	31391	7373	217+ 38162 (ex.sorting)	222+8274 (ex.sorting)	38162	8274
Disk Writes	2500	3125	2500+35000 (ex.sorting)	3125+7500 (ex.sorting)	35000	7500
Average Disk Reads	1570	369	11	11	-----	-----

Μπορούμε να πούμε ότι οι τρόποι οργάνωσης Β,Δ είναι πιο αποτελεσματικοί στην αναζήτηση αρχείων καθώς κατά την διάρκεια της αναζήτησης στην ουσία αποθηκεύουμε προσωρινά στο αρχείο ευρετηρίου μεγαλύτερο αριθμό εγγραφών σε κάθε block αφού η

εγγραφή του κάθε PointerRecord είναι 8 byte έναντι των 32 byte για κάθε MainRecord επομένως μειώνουμε τα disk accesses στο  $\frac{1}{4}$  αφού κάθε block του μικρού αρχείου αποτελείται από 16 εγγραφές επομένως διαβάζονται 16 εγγραφές ανά read στην μνήμη σε αντίθεση με τους Α,Γ τρόπους όπου διαβάζονται 4 εγγραφές ανά read. Ακόμη μπορούμε να προσθέσουμε ότι οι Γ,Δ τρόποι οργάνωσης που έχουν ταξινομηθεί είναι πιο αποτελεσματικοί παρότι τα αποτελέσματα δείχνουν να έχουν παραπάνω disk accesses διότι μιλάμε για έναν μικρό αριθμό κλειδιών που ζητήθηκε να αναζητηθούν, σε περίπτωση μεγαλύτερων ή πιο συγκεκριμένων σε εύρος τιμής κλειδιών αναζητήσεων τα disk accesses των Γ,Δ είναι λιγότερα. Επίσης αξίζει να αναφερθεί ότι μια υλοποίηση της sort με περισσότερα buffers θα ήταν πιο αποτελεσματική με μικρότερο disk I/O υπερτερώντας των μεθόδων Α,Β. Τέλος για τους τρόπους αναζήτησης η BinarySearch κάνει  $O(\log n)$  disk accesses, ενώ η κανονική search  $O(n)$  disk accesses όπου  $n$  ο αριθμός block του αρχείου.