

Ομάδα εργασίας:

Ονοματεπώνυμο, Αριθμός Μητρώου:

- 1. Γεώργιος Τσίρης, 1115201700173**
- 2. Ιωάννης Σκλιάμης, 1115201700141**

Σχεδιαστικές Επιλογές:

1. Το πρώτο μπλοκ ενός αρχείου κατακερματισμού (hash file) διακρίνεται από την συμβολοσειρά "hash" εκεί όπου αρχίζουν τα δεδομένα του. Αμέσως μετά είναι αποθηκευμένο το πλήθος των κουβάδων (buckets_num) που διαθέτει το αρχείο. Ο υπόλοιπος χώρος του πρώτου μπλοκ δεν χρησιμοποιείται.
2. Από το δεύτερο μπλοκ και για όσα μπλοκ χρειαστεί εκτείνεται το ευρετήριο (map). Κάθε μπλοκ ευρετηρίου αρχίζει με έναν μετρητή κουβάδων (bucket_counter). Στην συνέχεια βρίσκεται ένας ακέραιος που είναι ο αριθμός του επόμενου μπλοκ ευρετηρίου (next_map_block), αν υπάρχει αλλιώς έχει την τιμή NO_LINK. Από εκεί και πέρα, όλος ο υπόλοιπος χώρος του μπλοκ είναι για αποθήκευση κουβάδων (συγκεκριμένα ακεραίων αριθμών που αντιστοιχούν στο πρώτο μπλοκ με εγγραφές του εκάστοτε κουβά). Αρχικά κανένας κουβάς δεν έχει μπλοκ που να του αντιστοιχούν, οπότε έχει την τιμή EMPTY_BUCKET.
3. Τα υπόλοιπα μπλοκ (πλην του πρώτου και των μπλοκ ευρετηρίου), αρχίζουν με ένα ακέραιο (record_counter), ο οποίος ανα πάσα στιγμή έχει καταμετρημένο το τρέχον πλήθος εγγραφών που διαθέτει το συγκεκριμένο μπλοκ. Στην συνέχεια βρίσκεται ένας ακέραιος που είναι ο αριθμός του επόμενου μπλοκ στην αλυσίδα υπερχειλίσης (overflow_block), αν υπάρχει αλλιώς έχει την τιμή NO_LINK. Από εκεί και πέρα, όλος ο υπόλοιπος χώρος είναι για την αποθήκευση εγγραφών του εκάστοτε κουβά στον οποίο ανήκει το μπλοκ.
4. Ο κουβάς στον οποίο οδηγείται μια εγγραφή εξαρτάται από την τιμή της συνάρτησης κατακερματισμού για είσοδο το id της εγγραφής. Στην συγκεκριμένη περίπτωση, η συνάρτηση κατακερματισμού υλοποιεί το υπόλοιπο της διαίρεσης του id της εγγραφής προς το συνολικό πλήθος κουβάδων (buckets_num) που διαθέτει το αρχείο.
5. Η κατασκευή του ευρετηρίου δεν υλοποιείται στην συνάρτηση HT_CreateIndex(), αλλά στην πρώτη κλήση της συνάρτησης HT_InsertEntry() για το αρχείο, μιας και τότε είναι η πρώτη φορά που θα χρειαστεί η δομή του ευρετηρίου. Η HT_CreateIndex(), απλώς δεσμεύει το πρώτο μπλοκ και το αρχικοποιεί κατάλληλα.
6. Η συνάρτηση HT_DeleteEntry() αφού βρεί ποια εγγραφή πρέπει να διαγράψει (κάνοντας hash το δοσμένο id), την κάνει overwrite με την τελευταία εγγραφή του συγκεκριμένου μπλοκ και μειώνει τον μετρητή εγγραφών (record_counter) κατά 1.
7. Η HT_InsertEntry() αφού βρεί σε ποιον κουβά πρέπει να εισάγει την εγγραφή (κάνοντας hash το δοσμένο id), αρχίζοντας από το πρώτο μπλοκ που αντιστοιχεί στον κουβά αυτό αναζητά την πρώτη διαθέσιμη θέση στην αλυσίδα υπερχειλίσης του για να προχωρήσει στην εισαγωγή.
8. Ο global εμβέλειας πίνακας file_descs περιέχει τους αναγνωριστικούς αριθμούς (file descriptors) των αρχείων κατακερματισμού που έχουν ανοιχθεί μέχρι εκείνη την στιγμή. Όλες οι θέσεις του αρχικοποιούνται σε EMPTY_DESC μέσω της HT_Init().

to compile and link:

- make hp: Δοσμένη main (ht_main.c)
- make ht1: Πολλά insert/delete εγγραφών σε ένα ανοιχτό αρχείο (ht_main1.c)
- make ht2: Πολλά insert/delete εγγραφών σε πολλά ανοιχτά αρχεία με διαφορετικά buckets_num (ht_main2.c)

to run: ./build/runner (Προσοχή: Αρχεία με ίδιο όνομα με αυτά που θα παραχθούν δεν θα πρέπει να προϋπάρχουν! Ως εκ τούτου θα πρέπει να σβήνονται μετά από κάθε εκτέλεση.)