

Σχεδιασμός Βάσεων Δεδομένων

Εργασία 1

Τοσκολλάρι Ρόναλντ

3160244

Άσκηση 1

Α) Ένα μπλόκ αποθηκεύει 512B άρα ο αριθμός εγγραφών σε ένα block είναι: $512B/113B = 4.53$. Όπου 113B είναι το μέγεθος μίας εγγραφής.

Επειδή οι εγγραφές πρέπει να είναι ολόκληρες σε ένα μπλοκ το 4.53 γίνεται 4. Άρα η τελική απάντηση είναι 4 εγγραφές ανα μπλόκ.

Β) Συνολικά έχω 20.000 εγγραφές, άρα $20.000/4 = 5.000$ το πλήθος των μπλοκ που χρειάζομαι.

Γ) Για να βρω τι μέσο χρόνο για 5.000 μπλοκ πρέπει να βρω τον χρόνο που χρειάζομαι για 2.500 block.

ι) Στην σειριακή αναζήτηση έχω seek time και rotation delay μόνο στο πρώτο μπλοκ και για τα άλλα έχω μόνο transfer time.

$$\text{Avg.RandomDelay} = 30/2.400 = 0,0125s = 12,5\text{msec}$$

Κάθε μπλοκ είναι 512B, συνολικά έχω 20 μπλοκ ανα ίχνος άρα σε κάθε ίχνος έχω $20*512B = 10.240B$ ανα ίχνος.

$$\begin{aligned} \text{Για να διαβάσω 512B θέλω } & 512/(10240/(60/2400)) = \\ & 512/(409.600/1000) = 1.25\text{msec}.(1) \end{aligned}$$

Άρα ο τελικός χρόνος για να διαβασώ 2500 μπλοκ είναι $T = \text{Avg. Seek Time} + \text{Avg. Rotation Delay} + (2500*1.25\text{msec}) = 30\text{msec} + 12,5\text{msec} + 3.125\text{msec} = 3,167.5\text{msec}$

ii) Όταν διαβάζω τυχαία δεδομένα στο δίσκο το Avg. Seek Time και Avg. Delay Rotation υπάρχει κάθε φορά. Άρα έχω $T = 2500 * (\text{Avg. Seek Time} + \text{Avg. Rotation Delay} + 1.25\text{msec}) = 109.375 \text{ msec}$.

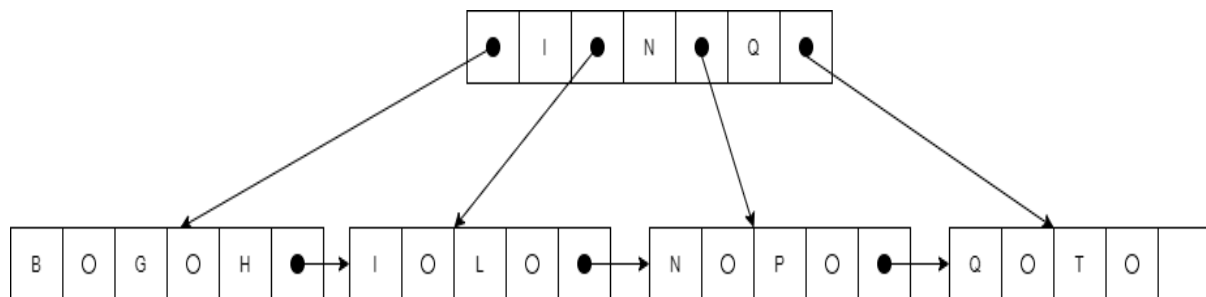
Δ) Εδώ δεν έχουμε σειριακή αναζήτηση αλλά δυαδική, ο χρόνος εκτέλεσης την δυαδικής αναζήτησης είναι $O(\log_2 n)$. Εμείς έχουμε συνολικά έναν πίνακα 5000 μπλοκ αφού τόσες χρειάζονται τα μπλόκ.

Άρα ο χρόνος που χρειαζόμαστε είναι $T = O(\log_2 5000) * (T\text{-Random}) = 12,2 * 1.25 = 15,25 \text{ msec} = 0,01525 \text{ sec}$

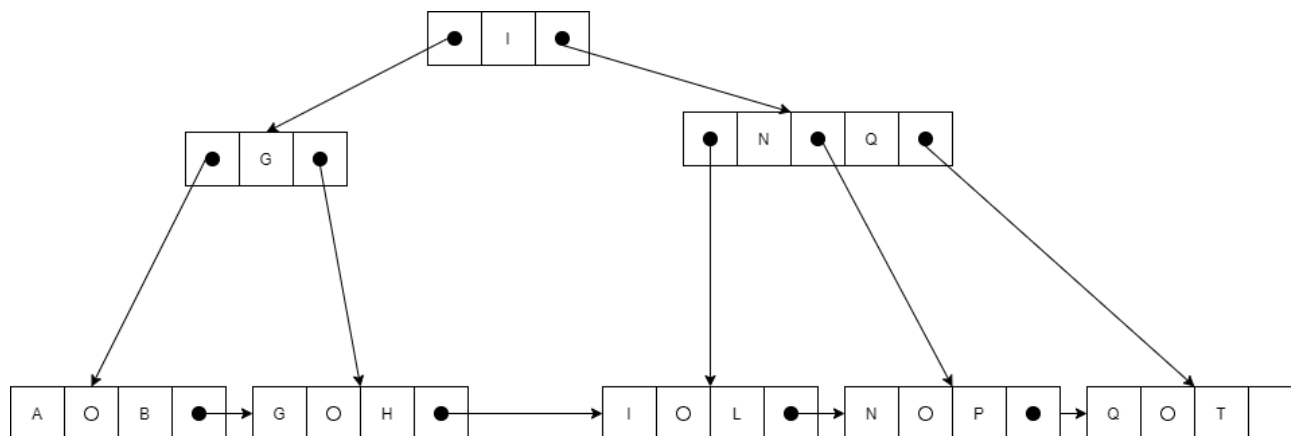
Άσκηση 2

A)

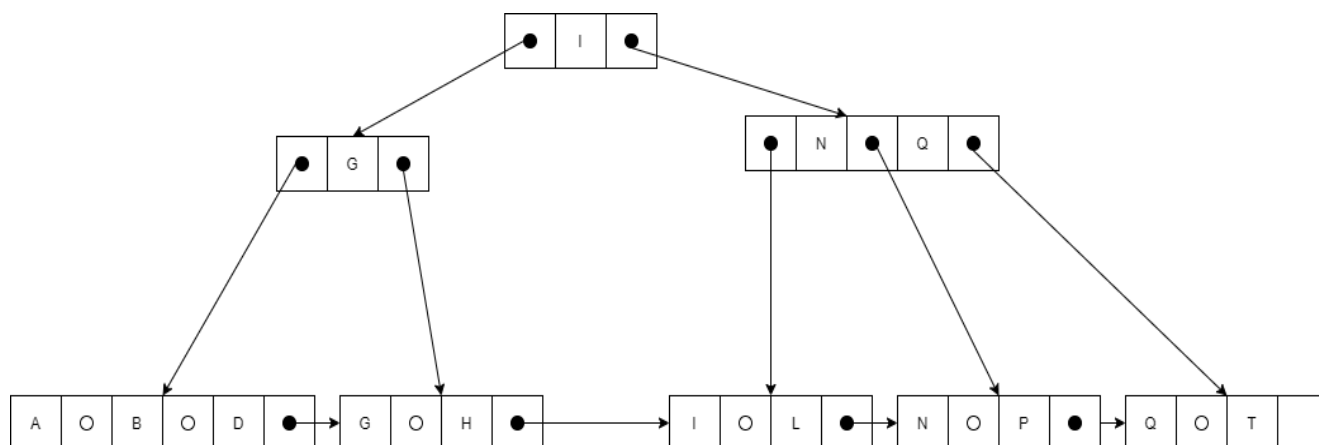
Εισαγωγή του G. Το G είναι μικρότερο από το I οπότε πάει στο αριστερό παιδί, και η εισαγωγή του θα γίνει μετά το B.



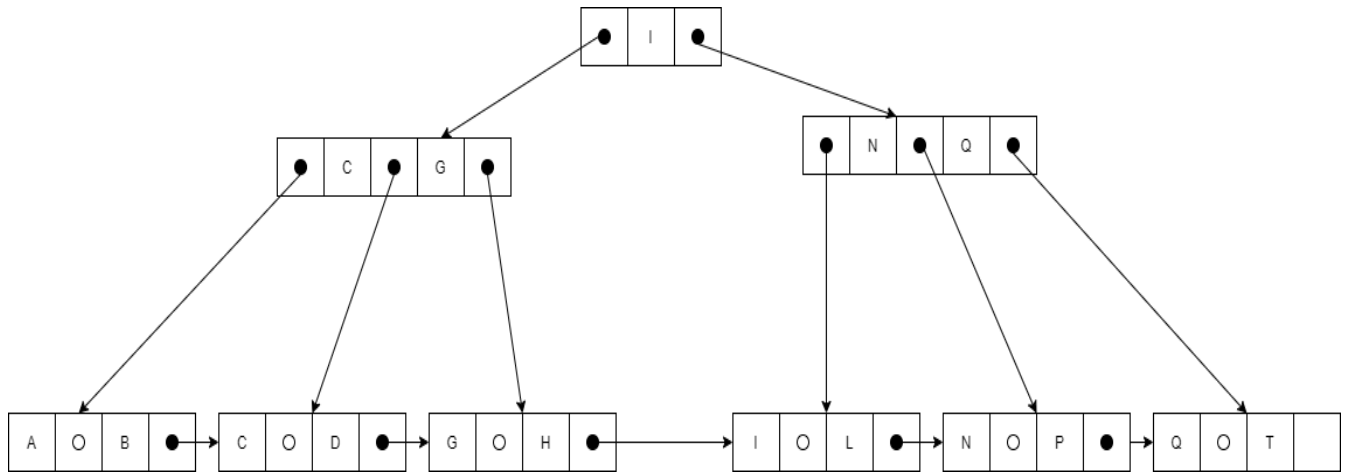
Εισαγωγή του A. Το A πρέπει να μπει αριστερά από το B, σε αυτό τον κόμβο όμως έχουμε ήδη 3 κλειδιά οπότε πρέπει να σπάσει σε 2 κομμάτια όπου το πρώτο θα έχει το A, B και το δεύτερο το G, H. Ο κόμβος πατέρα θα σπάσει σε δύο νέους όπου ο πρώτος θα έχει το κλειδί G και θα είναι πατέρας των A, B και G, H. Ο δεύτερος θα έχει τα κλειδιά N, Q και θα είναι πατέρας των άλλων κόμβων. Τέλος οι δύο νέοι κόμβοι θα έχουν πατέρα έναν νέο κόμβο με τιμή I.



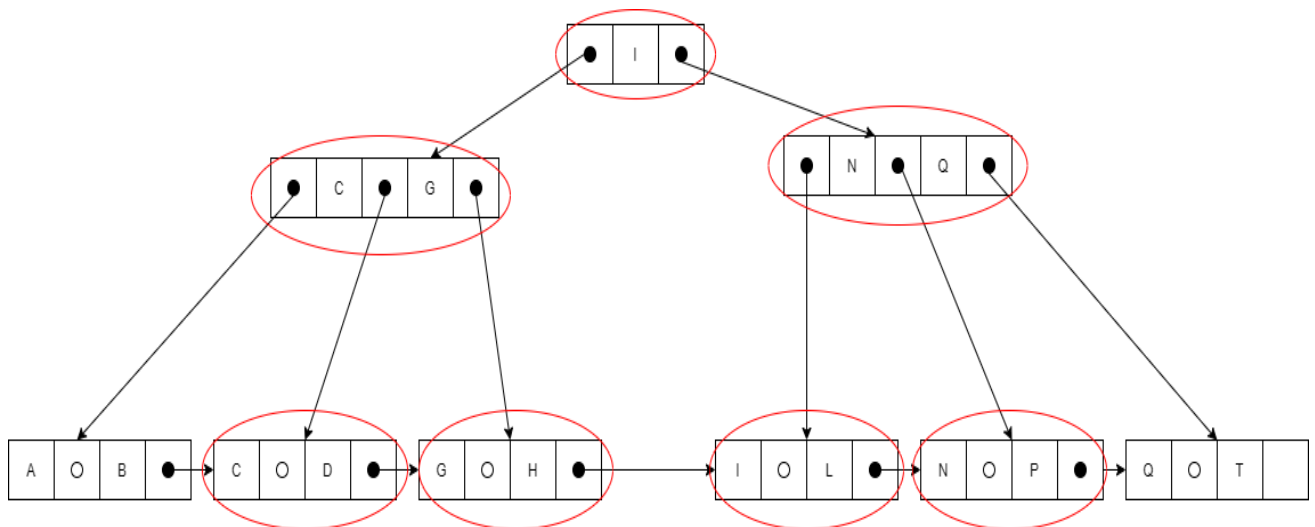
Εισαγωγή του D. Το $D < I$, επίσης το $D < G$, ο κόμβος έχει χώρο και επίσης $B < D$. Άρα έχουμε το παρακάτω δέντρο.



Εισαγωγή του C. Το $C < \text{επίσης το } C < G$, εδώ ο κόμβος δεν έχει χώρο άρα πρέπει να σπάσει. Έτσι έχουμε 2 νέους κόμβους που ο πρώτος έχει το A,B και ο δεύτερος το C,D. Βάζουμε στο κόμβο πατέρα το μικρότερο στοιχείο το οποίο είναι αυτό που μόλις βάλαμε και προκύπτει το δέντρο παρακάτω.



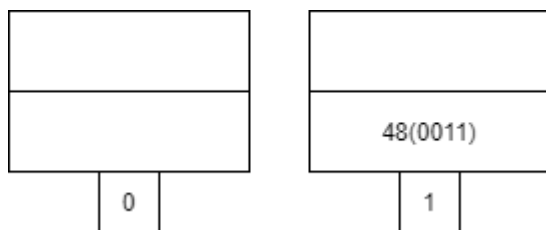
B)



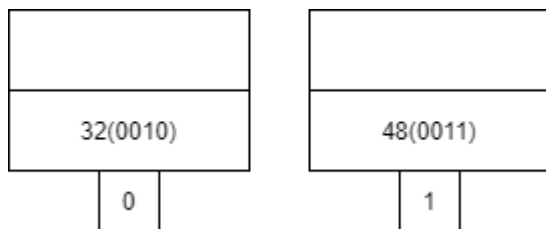
Με κόκκινο είναι οι κόμβοι που θα περάσει το range query $key \geq "C"$ AND $key \leq "P"$.

Άσκηση 3

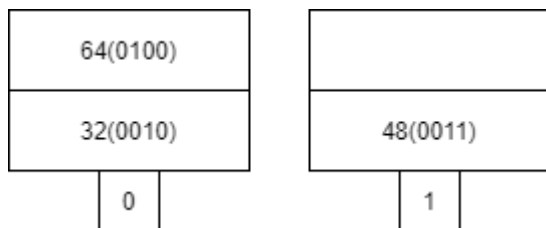
A) Το πρώτο δίνεται απο εκφώνηση.



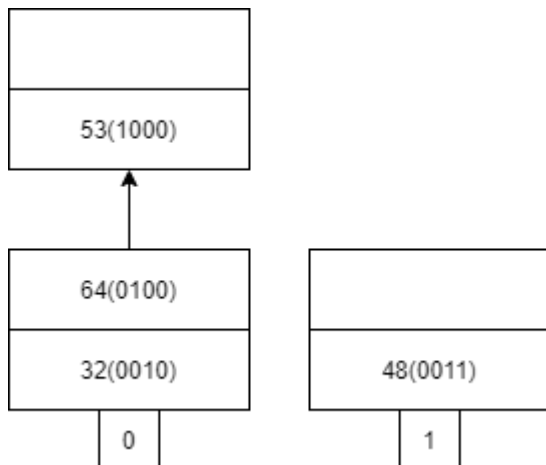
Εισαγωγή 32, $h(32) = 2$ (0010), $U < 80\%$



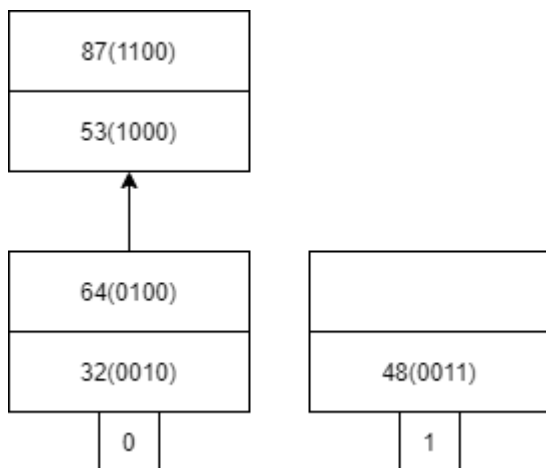
Εισαγωγή 64, $h(64) = 4$ (0100), $U < 80\%$



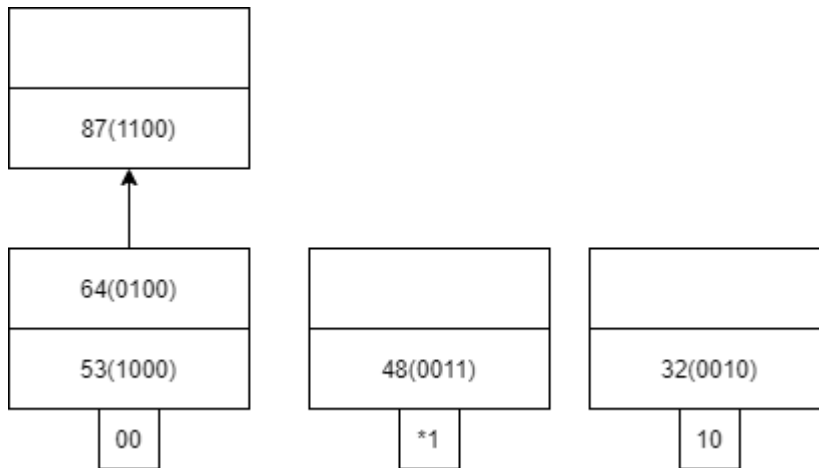
Εισαγωγή 53, $h(53) = 8$ (1000), $U < 80\%$



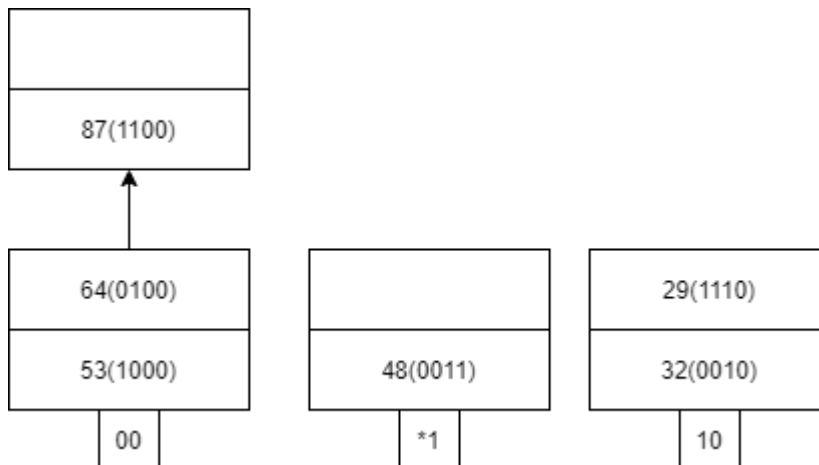
Εισαγωγή 81, $h(81) = 12$ (1100), $U < 80\%$



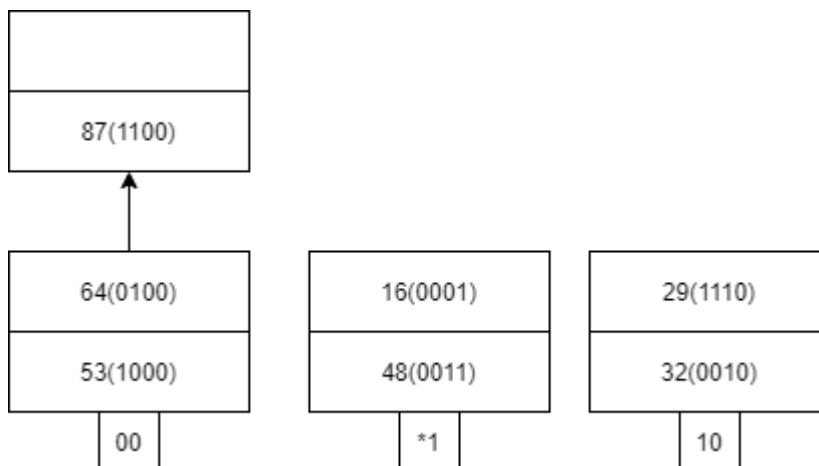
Μετά την εισαγωγή έχω $U > 80\%$, αρα $m \rightarrow 10$



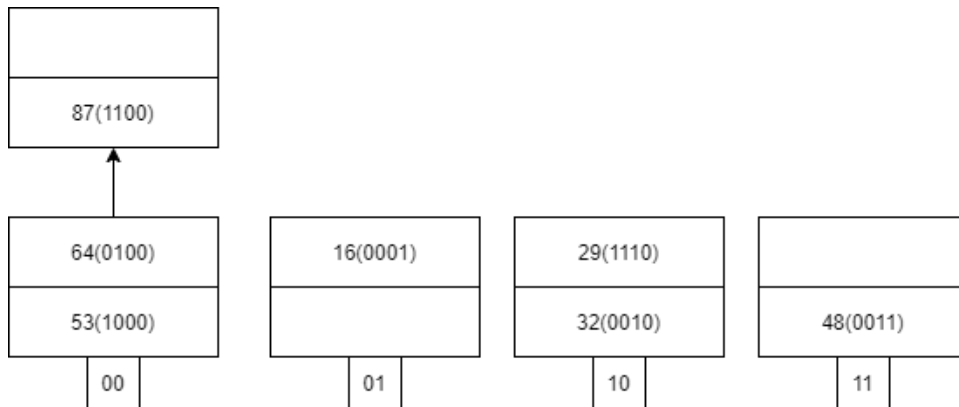
Εισαγωγή 29, $h(29) = 14$ (1110), $U < 80\%$



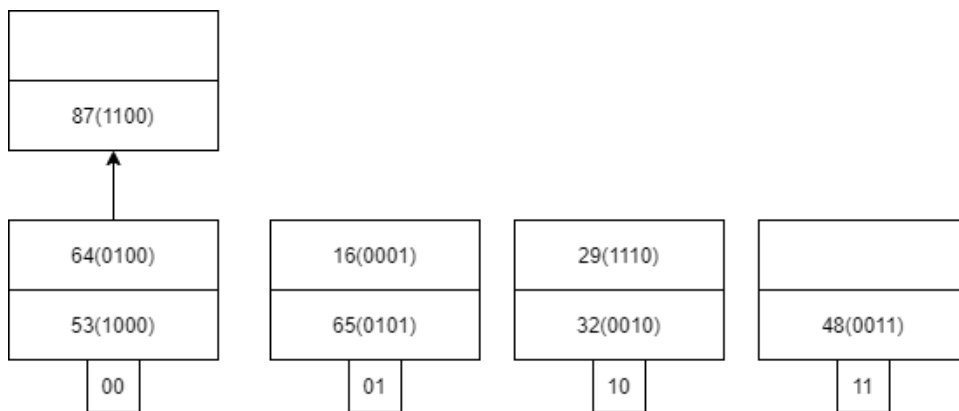
Εισαγωγή 16, $h(16) = 1$ (0001), $U < 80\%$



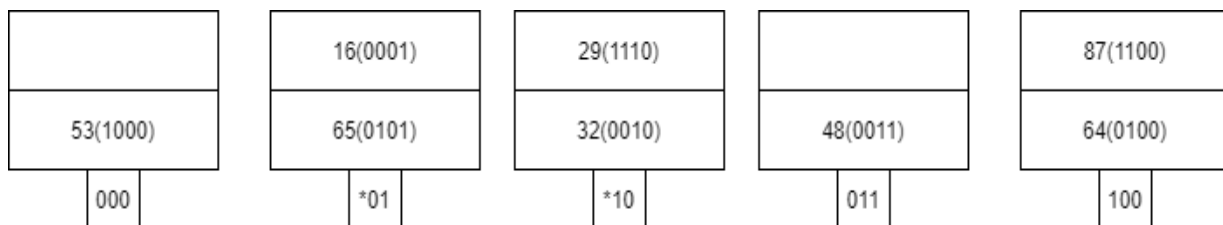
Μετά την εισαγωγή έχω $U > 80\%$, άρα $m \rightarrow 11$



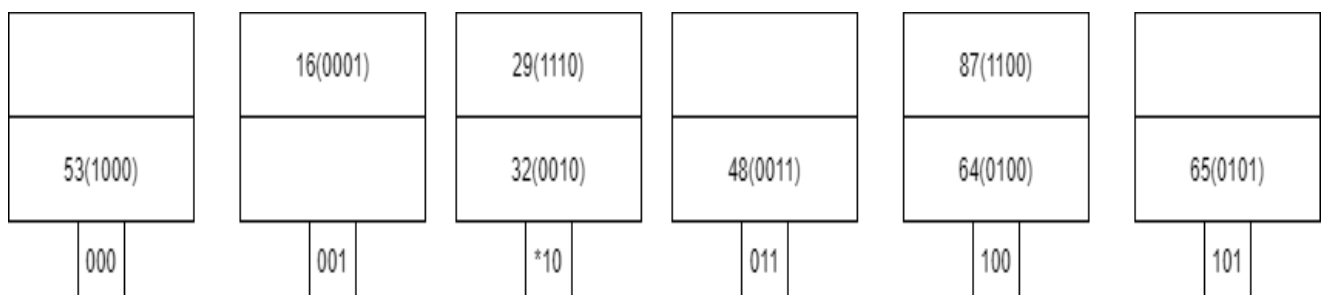
Εισαγωγή 65, $h(65) = 5$ (0101), $U < 80\%$



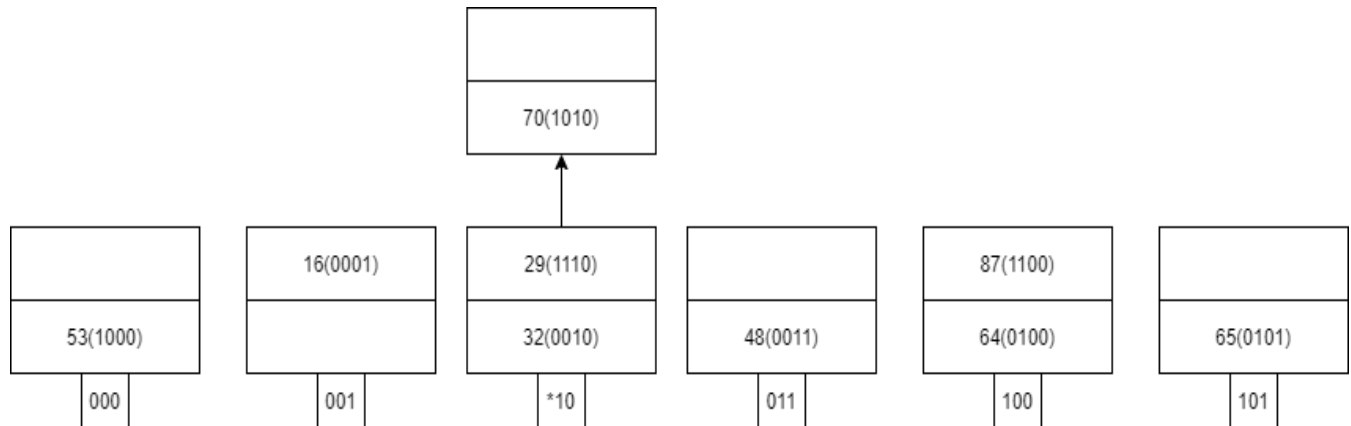
Μετά την εισαγωγή έχω $U > 80\%$, άρα $m \rightarrow 100$



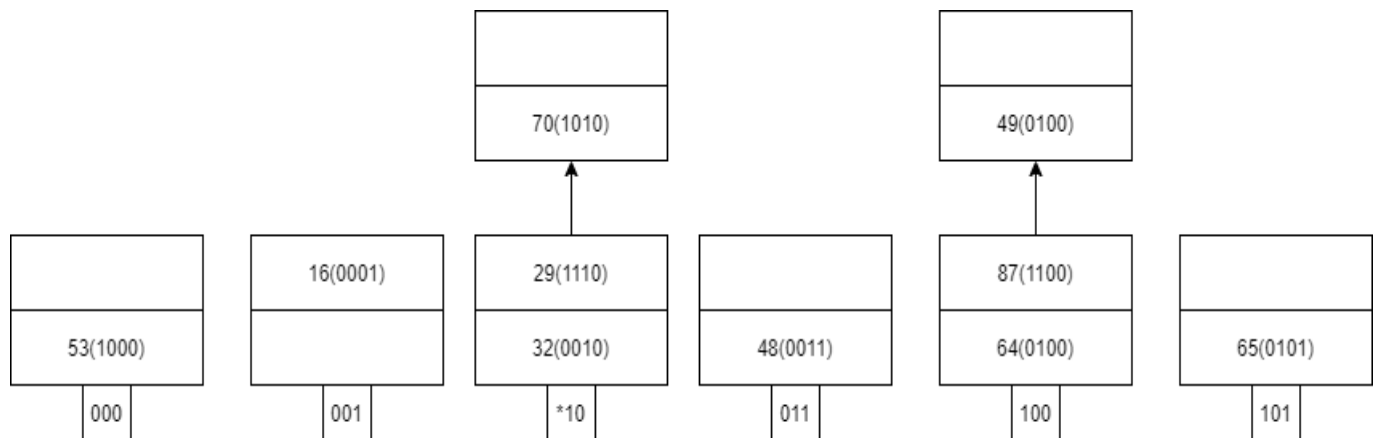
Μετά την διάσπαση έχω ακόμη $U > 80\%$, άρα $m \rightarrow 101$



Εισαγωγή 70, $h(70) = 10$ (1010), $U < 80\%$



Εισαγωγή 49, $h(49) = 4$ (0100), $U < 80\%$



B)

I) 8 εγγραφές θέλουν 1 προσπέλαση και 2 εγγραφές θέλουν 2 προσπελάσεις, άρα χρειαζόμαστε $8 \cdot (1/10) + 2 \cdot (2/10) = 1.2$

II) Στην χειρότερη περίπτωση θα ψάξω κάθε bucket, άρα θα είναι $O(n)$ για κάθε περίπτωση. Τώρα στή μέση περίπτωση θα είναι $O(n/2)$. Δηλαδή $O(5/2) = 2.5$.

