[A]

Πίνακας Δήλωση:5000 εγγραφές , 20 εγγραφές/σελίδα. Συνεπώς ο πίνακας Δήλωση έχει 5000/20=250 σελίδες.

Πίνακας Φοιτητής: 1000 εγγραφές και 10 εγγραφές/σελίδα. Συνεπώς ο πίνακας Φοιτητής έχει 1000/10=100 σελίδες.

Β=21 διαθέσιμα πλαίσια στην ενδιάμεση μνήμη.

Υπολογισμός κόστους σύζευξης Δ⨝Φ:

1)Κόστος σύζευξης με εμφώλευση βρόχων κατά σελίδα:

Περίπτωση 1η: Έστω Δ εξωτερικός και Φ εσωτερικός.

Κόστος = Δ+Δ\*Φ

Κόστος = 250 + 250 \*100=25.250 Ι/Ο.

Περίπτωση 2η: Έστω Φ εξωτερικός και Δ εσωτερικός.

Κόστος = Φ+Φ\*Δ

Κόστος= 100 + 100\*250=25.100 Ι/Ο.

2)Κόστος σύζευξης με εμφώλευση βρόχων κατά μπλοκ:

Περίπτωση 1η: Έστω Δ εξωτερικός και Φ εσωτερικός.

Κόστος = Δ+ Ceil(Δ/Β-2) \*Φ, Ceil(Δ/Β-2)=Ceil(250/19)=14.

Κόστος= 250 + 14 \* 100=1650 Ι/Ο.

Περίπτωση 2η: Έστω Φ εξωτερικός και Δ εσωτερικός.

Κόστος = Φ+ Ceil(Φ/Β-2) \*Δ ,Ceil(Φ/Β-2)=Ceil(100/19)=6.

Κόστος = 100+6\*250=1600 Ι/Ο.

3)Κόστος σύζευξης με εμφώλευση βρόχων και χρήση ευρετηρίου:

Περίπτωση 1η: Έστω Δ εξωτερικός και Φ εσωτερικός.

Επειδή έχουμε ευρετήριο στον πίνακα Φοιτητής θα έχουμε τον Δ εξωτερικό και τον Φ εσωτερικό.

Κόστος = Δ+((Δ\* PΔ) \* κόστος εύρεσης εγγραφής) , κόστος εύρεσης εγγραφής =1.

Κόστος = 250 + (250\*20) \* 1

Κόστος =5250 Ι/Ο.

4)Κόστος σύζευξης με ταξινόμηση και συγχώνευση:

Περίπτωση 1η : Απλή έκδοση.

Κόστος= 2 \*Δ \*(1+Ceil(log(Ceil(Δ/Β))) + 2 \*Φ \*(1+Ceil(log(Ceil(Φ/Β))) + (Δ+Φ)

Έχοντας λογάριθμο με βάση Β-1=20.

Ceil(log(Ceil(Δ/Β)))=Ceil(log(12))=1.

Ceil(log(Ceil(Φ/Β))=Ceil(log(5))=1.

Κόστος = 2 \* 250 (1+1) + 2 \* 100( 1+1) + (250+100)

Κόστος = 1000+400+350

Κόστος=1750 Ι/Ο.

Περίπτωση 2η : Βελτιωμένη έκδοση.

Για να επιτυγχάνεται η βελτιωμένη έκδοση πρέπει:

B>=sqrt(Δ) => 21>=sqrt(250)=>21>=16 όπου Δ ο μεγάλος πίνακας.

Οπότε για την βελτιωμένη λύση χρειαζόμαστε τουλάχιστον 16 πλαίσια της ενδιάμεσης μνήμης .

Κόστος = Πέρασμα Δ +Πέρασμα Φ+ Συγχώνευση

Κόστος = 2\*Δ + 2 \* Φ+ (Δ+Φ)

Κόστος = 3\*Δ+ 3\*φ => Κόστος =3 \* 250+3 \* 100=>Κόστος =1050 Ι/Ο.

5)Κόστος σύζευξης με κατακερματισμό:

Για να επιτυγχάνεται η βέλτιστη λύση πρέπει :

B-1>=sqrt(Φ) =>20>=sqrt(100)=>20>=10 , όπου Φ ο μικρός πίνακας.

Οπότε για την βέλτιστη λύση χρειαζόμαστε 10 πλαίσια της ενδιάμεσης μνήμης.

Β=21: Διαμερίσματα Δ: Β-1=20.

Διαμερίσματα Φ: Β-1=20

Ceil(Δ/Β-1)=Ceil(250/20)=13 σελίδες.

Ceil(Φ/Β-1)=Ceil(100/20)= 5 σελίδες.

Κόστος = 2 \*Δ+ 2 \* Φ + Δ+Φ

Κόστος =2\* 250+ 2 \*100+ 250 + 100

Κόστος =1050 Ι/Ο.

Επειδή Β>sqrt(Δ)=> 21= sqrt(250) με το ίδιο κόστος θα μπορούμε να έχουμε και τον Δ εξωτερικό. Παρόλα αυτά είναι καλύτερα να κρατήσουμε τον μικρότερο πίνακα ως εξωτερικό.

Επίσης παρατηρούμε ότι το κόστος σύζευξης ταξινόμησης και συγχώνευσης στην βελτιωμένη έκδοση είναι 1050 Ι/Ο όπως και στην περίπτωση του κόστους σύζευξης με κατακερματισμό. Επιλέγουμε ως καλύτερο πλάνο την σύζευξη με κατακερματισμό διότι για να επιτευχθεί το κόστος των 1050 Ι/Ο χρειάζεται να χρησιμοποιήσει 10 πλαίσια της ενδιάμεσης μνήμης έναντι της βελτιωμένης έκδοσης σύζευξης ταξινόμησης και συγχώνευσης που χρειάζεται 16 πλαίσια για να πετύχει αυτό το κόστος.

Συνεπώς καλύτερη επιλογή η σύζευξη με ευρετήριο κατακερματισμού με κόστος =1050 Ι/Ο.

[B]

Έχουμε το εξής SQL αίτημα :

Select Φ.ΑΜ,Επώνυμο

FROM Δήλωση Δ JOIN Φοιτητής Φ ΟΝ Δ.ΑΜ=Φ.ΑΜ

WHERE Ημερομηνία>=’1/1/2018’ and Επώνυμο LIKE’Π%’

Σύμφωνα με εκφώνησης γνωρίζουμε ότι:

1)Οι μισές εγγραφές του πίνακα Δήλωση είναι σε ημερομηνίες από την 1/1/2018 και μετά:

2500 εγγραφές / (20 εγγραφές/σελίδα) = 125 σελίδες επίσης επειδή από τον πίνακα δήλωση θέλουμε μόνο το ένα πεδίο το ΑΜ, θα έχουμε 125/3=42 ,συνεπώς Τ1=42 σελίδες.

2) Το 10% των φοιτητών έχουν επίθετο που αρχίζει από Π:

10% των εγγραφών έχουν επίθετο που αρχίζει από Π: 1000 εγγραφές \*(10/100)= 100 εγγραφές και 10 εγγραφές/σελίδα , έχουμε Τ2=100/10 =>Τ2=10 σελίδες, ομοίως με πριν επειδή από τον πίνακα Φοιτητής χρειαζόμαστε μόνο δύο πεδία θα έχουμε ότι Τ2=10\*(2/3) =>Τ2=7 σελίδες.

Έχοντας προωθήσει προβολές .

Έχοντας Τ1=42 και Τ2=7 .

1)Κόστος σύζευξης με εμφώλευση βρόχων κατά σελίδα .

Εξωτερικός ο πίνακας Φοιτητής.

Κόστος =Τ1 +Τ1\*Τ2 +(Τ1+Τ2) + (Φ+Δ)

Κόστος =7+7 \*42+ 49+ 350=> Κόστος = 700 Ι/Ο

2) Κόστος σύζευξης με εμφώλευση βρόχων κατά μπλοκ:

B=21 .

Κόστος =Τ2+ Ceil(T2/B-2)\* T1 + (Δ+Φ) + (Τ1+Τ2)

Κόστος =7+ Ceil(7/19) \* 42 + (100+250) + (42+7)

Κόστος = 7+ 1 \* 42 + 350 +49 => Κόστος=448 Ι/Ο.

3)Κόστος σύζευξης με εμφώλευση βρόχων και ευρετήριο:

Κόστος =(Δ+Τ1)+ Τ1 +(Εγγραφές με ημερομηνία >=’1/1/2018’ \* 1) =250+42 +42 +(2500 \* 1) =2834 Ι/Ο.

Σχήμα σχεσιακής άλγεβρας:

πΦ.ΑΜ, Επώνυμο

Ι

Ι

Σύζευξη Φ.ΑΜ=Δ.ΑΜ

Ι

πΑΜ, Επώνυμο

I I

σΕπώνυμο LIKE’Π%’ σΗμερομηνία>=’1/1/2018’

I Ι

Φοιτητής Δήλωση

4) Κόστος σύζευξης με ταξινόμηση και συγχώνευση:

Α)Όχι βελτιωμένη λύση:

Κόστος = 2 \* Τ1 \*(1+Ceil(log(Ceil(T1/B)))+ 2\*T2 \*(1+Ceil(log(Ceil(T2/B))) + (T1+T2)

Έχοντας λογάριθμο με βάση το Β-1=20.

Ceil(log(Ceil(T1/B))=Ceil(log(2))=1.

Ceil(log(Ceil(T2/B))=Ceil(log(1))=0.

Κόστος =2 \* 42 (1+1)+ 2 \* 7 (1+0) + (42+7)

Κόστος = 231.

Επιπλέον Δ+Τ1=250+42=292 και Φ+Τ2=100+7=107.

Άρα τελικό κόστος εκτέλεσης = 231+ 292+107=630 Ι/Ο.

Β)Βελτιωμένη λύση:

Ομοίως με την άσκηση Α:

Για να επιτυγχάνεται η βελτιωμένη έκδοση πρέπει:

B>=sqrt(Δ) => 21>=sqrt(250)=>21>=16 όπου Δ ο μεγάλος πίνακας.

Οπότε για την βελτιωμένη λύση χρειαζόμαστε τουλάχιστον 16 πλαίσια της ενδιάμεσης μνήμης .

Κόστος = (T1+T2) + (Φ+Δ) + 3\* (Τ1+Τ2)

Κόστος=49 + 350 +3 \* 42 => Κόστος=546 Ι/Ο.

5)Κόστος σύζευξης με ευρετήριο κατακερματισμού:

Για να επιτυγχάνεται η βέλτιστη λύση πρέπει :

B-1>=sqrt(Φ) =>20>=sqrt(100)=>20>=10 , όπου Φ ο μικρός πίνακας.

Οπότε για την βέλτιστη λύση χρειαζόμαστε 10 πλαίσια της ενδιάμεσης μνήμης.

Κόστος = (Τ1+Τ2)+ (Φ+Δ+) + 3\*(Τ1+Τ2) Κόστος= 49 +350 + 3 \* 49 =>

Κόστος = 546 Ι/Ο.

Συνεπώς η καλύτερη λύση είναι η σύζευξη με εμφώλευση βρόχων κατά μπλοκ με

Κόστος =448 Ι/Ο.

Το δέντρο σχεσιακής άλγεβρας :

πΦ.ΑΜ, Επώνυμο

Ι

Ι

Σύζευξη Φ.ΑΜ=Δ.ΑΜ

Ι Ι

Ι Ι

Ι Ι

πΑΜ, Επώνυμο πΑΜ

Ι Ι

σΕπώνυμο LIKE’Π%’ σΗμερομηνία>=’1/1/2018’

I Ι

Φοιτητής Δήλωση