Πανεπιστημιο Ιωαννινων
Τμήμα Πληροφορικής

Βάσεις Δεδομένων ΙΙ
Ακαδημαϊκό Έτος 2000-2001

ΤΕΛΙΚΟ ΔΙΑΓΩΝΙΣΜΑ
ΕΞΕΤΑΣΤΙΚΗ ΙΟΥΝΙΟΥ

Καλή Επιτυχία

Θέμα 1: [Μονάδες 15] Κτίζετε ένα σύστημα βάσεων δεδεομένων για μια συγκεκριμένη εφαρμογή όπου οι δοσοληψίες προσπελαύνουν τα δεδομένα με τη σειρά. Συγκεκριμένα, τα δεδομένα είναι διατεταγμένα $x_1, x_2, \ldots x_n$ έτσι ώστε αν μια δοσοληψία προσπελάσει το x_i μετά δε μπορεί να προσπελάσει οποιοδήποτε δεδομένο πριν το x_i. Δηλαδή, αν ένα χρονοπρόγραμμα S περιέχει τις πράξεις $p_k(x_i)$ και $p_k(x_j)$ (όπου το p είναι είτε πράξη εγγραφής είτε πράξη ανάγνωσης) και $i < j$ τότε $p_k(x_i) <$ (δηλαδή προηγείται) $p_k(x_j)$. Ο συνεργάτης σας ισχυρίζεται ότι εξαιτίας αυτής της ιδιότητας τα χρονοπρογράματα είναι πάντα σειριοποιήσιμα με συγκρούσεις και δε χρειάζεται να υλοποιήσετε έλεγχο συνδρομικότητας (concurrency control). Αποδείξτε ότι κάνει λάθος.

Θέμα 2: [Μονάδες 20] Για τα παρακάτω χρονοπρογράμματα δείξτε αν είναι: σειριοποιήσιμα βάσει όψεων, σειριοποιήσιμα βάσει συγκρούσεων, με δυνατότητα ανάκαμψης, χωρίς διάδοση ανακλήσεων, αυστηρά. Αποδείξτε την απάντησή σας.

(i) $r_1(x)r_2(y)w_3(x)r_2(x)r_1(y)$
(ii) $r_1(x)w_2(x)w_1(x)abort(T_2)commit(T_1)$
(iii) $r_1(x)w_2(x)w_1(x)commit(T_2)commit(T_1)$

Θέμα 3: [Μονάδες 15] Στο μάθημα περιγράψαμε τη μέθοδο ανάκαμψης ARIES. Ποια είναι τα τρία βασικά στάδια (φάσεις) της μεθόδου και τι πετυχαίνει κάθε στάδιο (φάση); (η απάντηση σας να μην υπερβαίνει τη 1 σελίδα)

Θέμα 4: [Μονάδες 15] Συγκρίνετε ένα σχεσιακό με ένα αντικειμενο-στρεφές σύστημα διαχείρισης βάσεων δεδομένων (ΣΔΒΔ). Δώστε μια εφαρμογή για την οποία θα διαλέγατε ένα σχεσιακό ΣΔΒΔ και μία για την οποία θα διαλέγατε ένα αντικειμενο-στρεφές και εξηγείστε γιατί. (η απάντηση σας να μην υπερβαίνει τη 1 σελίδα)

Θέμα 5: [Μονάδες 35] Θεωρείστε τις παρακάτω σχέσεις ενός κατανεμημένου ΣΔΒΔ:

Εργαζόμενος(<u>Εργαζόμενος_Αριθμός</u>, Τμήμα_Αριθμός, Μισθός)
Τμήμα(<u>Τμήμα_Αριθμός</u>, Προϊστάμενος_Αριθμός, Προϋπολογισμός)

Το γνώρισμα Προϊστάμενος_Αριθμός είναι ο Εργαζόμενος_Αριθμός του προϊσταμένου του Τμήματος. Κάθε σχέση έχει πλειάδες μήκους 20 bytes και τα πεδία Μισθός και Προϋπολογισμός παίρνουν τιμές που κατανέμονται ομοιόμορφα από το 1 έως το 1.000.000. Η σχέση Εργαζόμενος καταλαμβάνει 100.000 σελίδες και η σχέση Τμήμα 5.000, ο buffer (μνήμη) κάθε επεξεργαστή ενός κόμβου είναι 100 σελίδες των 4.000 bytes η καθεμία. Το κόστος μεταφοράς μιας σελίδας από τη μνήμη στο δίσκο (I/O) είναι t_d και το κόστος μεταφοράς από τον ένα κόμβο στον άλλο t_s.

Θέμα 5(α): Η σχέση Εργαζόμενος είναι αποθηκευμένη στη Νάπολη και η σχέση Τμήμα στο Βερολίνο. Δεν υπάρχουν ευρετήρια για καμία από τις σχέσεις. Θεωρείστε την ερώτηση:

SELECT *
FROM Εργαζόμενος E, Τμήμα T

WHERE Εργαζόμενος.Εργαζόμενος_Αριθμός = Τμήμα.Προϊστάμενος_Αριθμός

Η ερώτηση γίνεται στο Παρίσι και μόνο 1% των εργαζομένων είναι προϊστάμενοι. Υπολογίστε το κόστος υπολογισμού της ερώτησης για καθεμία από τις παρακάτω περιπτώσεις:

(i) Υπολογισμός της ερώτησης στη Νάπολη με μεταφορά της σχέσης Τμήμα στη Νάπολη και μεταφορά του αποτελέσματος στο Παρίσι.
(ii) Υπολογισμός της ερώτησης στη Νάπολη με χρήση ημισυνένωσης και μεταφορά του αποτελέσματος στο Παρίσι.
(iii) Υπολογισμός της ερώτησης στο Παρίσι με μεταφορά και των δύο σχέσεων στο Παρίσι.

Θέμα 5(β): Θεωρείστε ότι το κατανεμημένο ΣΔΒΔ έχει 10 κόμβους. Η σχέση Εργαζόμενος είναι οριζόντια κατατεμαχισμένη στους 10 κόμβους με βάση το Μισθό, έτσι ώστε οι πλειάδες με Μισθός $\leq$ 100.000 είναι στον πρώτο κόμβο, οι πλειάδες με 100.000 $<$ Μισθός $\leq$ 2000.000 στο δεύτερο κοκ. Επίσης, το τεμάχιο Μισθός με συνθήκη $\leq$ 100.000 έχει αντίγραφα σε όλους τους κόμβους. Θεωρείστε μια δοσοληψία που δίνει αύξηση 10% σε όλους. Θεωρείστε επίσης ότι έχουμε σύγχρονο πρωτόκολλο ελέγχου αντιγράφων για τα αντίγραφα του τεμαχίου Μισθός $\leq$ 100.000 και πολιτική ROWA. Περιγράψτε ποιοι κόμβοι προσπελαύνονται και ποια κλειδιά εκχωρούνται. Προσοχή: οι συνθήκες του αρχικού κατατεμαχισμού της σχέσης Εργαζόμενος πρέπει να διατηρηθούν και μετά την αύξηση.

Επιπρόσθετο Θέμα: Αποθήκες Δεδομένων [Μονάδες 25] Δώστε σύντομες απαντήσεις (όχι παραπάνω από 10 γραμμές ανά ερώτηση)

1. Δώστε την αρχιτεκτονική μιας αποθήκης δεδομένων
2. Περιγράψτε μεθόδους για να επιτύχουμε ενημέρωση μιας αποθήκης δεδομένων με αυξητικό υπολογισμό
3. Περιγράψτε δυο πράξεις που γίνονται στους υπερκύβους
4. Τι είναι το σχήμα Star;