

ΤΕΛΙΚΟ ΔΙΑΓΩΝΙΣΜΑ
Εξεταστική Ιουνίου

Διάρκεια Εξέτασης: 3 ώρες
Καλή Επιτυχία

ΘΕΜΑ 1 (Διαχείριση Συναλλαγών) [1.5 Μονάδα]

(α) Δώστε τους ορισμούς των τεσσάρων σημαντικών ιδιοτήτων μιας συναλλαγής, δηλαδή της ατομικότητας, (atomicity) συνέπειας (consistency) απομόνωσης (isolation) και μονιμότητας (duration). Ποια/ποιες ιδιότητες εξασφαλίζει ένα πρωτόκολλο ελέγχου ταυτοχρονισμού και ποια/ποιες οι τεχνικές ανάκαμψης (επαναφοράς από βλάβη).

(β) Για το παρακάτω χρονοπρόγραμμα

$W_2(X) W_1(Y) R_2(Y) W_2(Z) W_3(X)$

- (i) προσθέστε πράξεις Commit για τις συναλλαγές ώστε το χρονοπρόγραμμα που προκύπτει να *μην* έχει δυνατότητα ανάκαμψης (non-recoverable)
- (ii) προσθέστε πράξεις Commit για τις συναλλαγές ώστε το χρονοπρόγραμμα που προκύπτει να έχει δυνατότητα ανάκαμψης (recoverable) αλλά με διάδοση ανακλήσεων (cascade aborts)
- (iii) προσθέστε πράξεις Commit για τις συναλλαγές ώστε το χρονοπρόγραμμα που προκύπτει να έχει δυνατότητα ανάκαμψης (recoverable), να μη διαδίδει τις ανακλήσεις αλλά να *μην* είναι αυστηρό
- (iv) προσθέστε πράξεις Commit για τις συναλλαγές ώστε το χρονοπρόγραμμα που προκύπτει να είναι αυστηρό.

Εξηγείστε την απάντησή σας.

ΘΕΜΑ 2 (Έλεγχος Ταυτοχρονισμού) [2 Μονάδες]

(α) Έστω τέσσερις συναλλαγές T1, T2, T3, και T4. Η T1 παίρνει χρονόσημα (timestamp) 2, η T2 9, η T3 6 και η T4 7 από τον αλγόριθμο ταυτοχρονισμού με χρονοσήματα. Υποθέστε ότι και οι τέσσερις επικυρώνονται. Δώστε 3 διαφορετικούς πιθανούς γράφους σειριοποιησιμότητας.

(β) Για το καθένα από τα παρακάτω χρονοπρογράμματα απαντήστε αν μπορούν να προκύψουν από 2PL (κλείδωμα 2 φάσεων) πρωτόκολλο. Εξηγείστε την απάντησή σας.

- (i) $W_1(X) R_2(Y) R_1(Y) R_2(X)$
- (ii) $R_1(X) R_2(Y) W_3(X) R_2(X) R_1(Y)$

ΘΕΜΑ 3 (Τεχνικές Ανάκαμψης) [2 Μονάδες]

(α) Τι σημαίνει μια πολιτική NO FORCE/NO STEAL; Σε περίπτωση αποτυχιών, σε ποια ιδιότητα των δοσοληψιών προκαλεί προβλήματα μια τέτοια πολιτική και γιατί;

(β) Πότε ένα checkpoint (σημείο ελέγχου) λέγεται ασαφές; Συγκρίνεται πως η διαφορά του να παίρνουμε ασαφές και μη ασαφές checkpoint επηρεάζει την απόδοση του συστήματος (i) στην κανονική του λειτουργία (όταν δεν έχουμε αποτυχίες), (ii) το χρόνο που απαιτείται για ανάκαμψη από αποτυχία συστήματος (iii) το χρόνο που απαιτείται για ανάκαμψη από αποτυχία δίσκου

(γ) Στο πρωτόκολλο ARIES έχει σημασία η σειρά με την οποία θα γίνουν REDO οι πράξεις εγγραφής και γιατί;

ΘΕΜΑ 4 (Παράλληλες και Κατανεμημένες Βάσεις Δεδομένων) [2 μονάδες]

(α) Τι είναι η ημι-συνένωση και πότε χρησιμοποιείται;

(β) Θεωρείστε τη σχέση:

Τμήμα(did, mgrid, Προϋπολογισμός)

Κάθε πλειάδα της σχέσης είναι 20 bytes, κάθε γνώρισμα και οι τιμές του γνωρίσματος Προϋπολογισμός είναι ομοιόμορφα κατανεμημένες μεταξύ του 0 και του 1.000.000. Η σχέση Τμήμα καταλαμβάνει 5.000 σελίδες. Κάθε επεξεργαστής έχει 100 καταχωρητές 4.000 bytes ο καθένας. Το I/O κόστος μιας σελίδας είναι t_d και το κόστος μεταφοράς στο δίκτυο μιας σελίδας είναι t_s .

Έστω η ερώτηση:

Το τμήμα (did) με το μεγαλύτερο προϋπολογισμό.

Υπολογίστε το *συνολικό* χρόνο εκτέλεσής της ερώτησης σε όλους τους κόμβους καθώς και το *απαιτούμενο* (elapsed) χρόνο. Επίσης, δώστε το speed-up με βάση τον καλύτερο σειριακό αλγόριθμο.

(i) Θεωρείστε ένα παράλληλο ΣΔΒΔ με 100 επεξεργαστές, και ότι η σχέση έχει διαμεριστεί με round-robin

(ii) Θεωρείστε ένα παράλληλο ΣΔΒΔ με 100 επεξεργαστές, ότι η σχέση έχει διαμεριστεί με διαστήματα τιμών με βάση το γνώρισμα προϋπολογισμός.

ΘΕΜΑ 5 (Εξόρυξη Δεδομένων) [1.5 μονάδα]

(α) Ορίστε την υποστήριξη και την εμπιστοσύνη ενός συνδυαστικού κανόνα.

(β) Θεωρείστε ένα δέντρο αποφάσεων ύψους h , με n' κόμβους και n ($n \leq n'$) φύλλα. Πόσους κανόνες κατηγοριοποίησης μοντελοποιεί;

(γ) Στον αλγόριθμο συγκρότησης (clustering) BIRCH από τι καθορίζεται ο μέγιστος αριθμός k των συγκροτημάτων: Για να μειωθεί ο αριθμός των συγκροτημάτων η μέγιστη ακτίνα των συγκροτημάτων πρέπει να αυξηθεί ή να ελαττωθεί;

ΘΕΜΑ 6 (Βάσεις Δεδομένων και το Διαδίκτυο) [1 μονάδα]

Να παρουσιαστεί το αποτέλεσμα ενός αρχείου υπογραφών με εύρος 4 δυαδικά ψηφία (bits) για τα παρακάτω 5 έγγραφα. Κατασκευάστε τη δική σας συνάρτηση κατακερματισμού η οποία να απεικονίζει τους όρους σε αντίστοιχες δυαδικές θέσεις στη κάθε μια υπογραφή. Στη συνέχεια, χρησιμοποιείστε το αρχείο υπογραφών που έχετε δημιουργήσει για να απαντήσετε στο “Pitt AND Troy”.

Έγγραφο (αρχείο)	Λέξεις Κλειδιά
1	movie Troy Pitt Achilles
2	Homer Troy poem Iliad Odyssey
3	actor Pitt Aniston Missouri
4	Achilles Iliad Hector Menelaus Paris Troy
5	movie Snatch Pitt