

Υλοποίηση των Σχεσιακών Τελεστών

Ένα σχεσιακό ΣΔΒΔ πρέπει να συμπεριλαμβάνει αλγόριθμους για υλοποίηση των διαφορετικών τύπων των σχεσιακών πράξεων (καθώς και άλλων πράξεων) που μπορεί να εμφανισθούν σε μια στρατηγική εκτέλεσης επερώτησης. Οι πράξεις αυτές συμπεριλαμβάνουν τις βασικές και τις επεκταμένες πράξεις της σχεσιακής άλγεβρας και, σε πολλές περιπτώσεις, συνδυασμούς αυτών των πράξεων. Για κάθε τέτοια πράξη ή συνδυασμό πράξεων παρέχονται ένας ή περισσότεροι αλγόριθμοι για την εκτέλεση της πράξης. Όλοι οι αλγόριθμοι ανακτούν πλειάδες από μια ή περισσότερες σχέσεις εισόδου.

Δρόμοι Προσπέλασης (Access Paths)

Ένας δρόμος προσπέλασης είναι ή :

(1) Μια σάρωση ενός αρχείου

ή

(2) Ένα ευρετήριο και μια συνθήκη επιλογής

Ένα ευρετήριο αντιστοιχεί στην συνθήκη επιλογής αν μπορεί να χρησιμοποιηθεί για την ανάκτηση των πλειάδων που ικανοποιούν την συνθήκη.

Μια απλή συνθήκη επιλογής έχει την μορφή

γνώρισμα θ τιμή

$\theta \in \{<, \leq, =, >, \geq, \neq\}$

Η επιλεκτικότητα ενός δρόμου προσπέλασης είναι το πλήθος των σελίδων που πρέπει να προσπελασθούν για να ανακτηθούν οι επιθυμητές πλειάδες

Απόδοση

Στον υπολογισμό του κόστους μας ενδιαφέρει μόνο το κόστος εισόδου/εξόδου. Το κόστος αυτό θα το μετρήσουμε πλήθος σελίδων που μεταφέρονται από/προς τη μνήμη.

Η απόδοση κάθε πράξης μπορεί να εκτιμηθεί για διαφορετικές μεθόδους υλοποίησης.

ΠΡΟΜΗΘΕΥΤΗΣ (ΠΡΜ)

Κ_Προμ	Π_Όνομα	Είδος	Πόλη
22	Ανδρέου	7	Αθήνα
31	Πέτρου	8	Πάτρα
28	Δέδες	12	Λάρισα
58	Παππάς	7	Αθήνα

ΠΡΟΙΟΝ (ΠΡ)

Κ_Πρ	Πρ_Όνομα	Χρώμα	Βάρος
Π35	Οθόνη	Γκρι	8
Π76	DVD	Μαύρο	4
Π14	Κράνος	Άσπρο	1

ΕΡΓΟ (ΕΡ)

Κ_Ερ	Προυπ	Πόλη
E25	245K	Αθήνα
E32	500K	Άργος
E68	1000K	Αθήνα

ΠΑΡΑΓΓΕΛΙΑ (ΠΑΡ)

Κ_Προμ	Κ_Πρ	Κ_Ερ	Ποσοτ
22	Π35	E25	200
31	Π35	E25	100
31	Π76	E32	150
28	Π14	E68	150
22	Π76	E25	450
22	Π35	E32	670
22	Π35	E68	650
58	Π14	E68	670
28	Π14	E32	540
22	Π14	E25	200

Πράξη της Επιλογής (1)

Απλή επιλογή της μορφής:

$$\sigma_{R.A\theta c}(\mathbf{R})$$

Δεν υπάρχει ευρετήριο και τα δεδομένα δεν είναι ταξινομημένα στο γνώρισμα A

Σειριακή αναζήτηση (εξαντλητική αναζήτηση): ανάκτηση κάθε εγγραφής του αρχείου και έλεγχος αν οι τιμές των γνωρισμάτων της ικανοποιούν τη συνθήκη επιλογής.

Το κόστος είναι όσο μια εξαντλητική σάρωση του αρχείου. Αν επομένως το μέγεθος του αρχείου είναι b σελίδες αυτό θα είναι και το κόστος γενικά.

(Πότε μπορεί να είναι λιγότερο;)

σ Ποσοτ>200 (ΠΑΡ)

ΠΑΡΑΓΓΕΛΙΑ (ΠΑΡ)

Κ_Προμ	Κ_Πρ	Κ_Ερ	Ποσοτ
22	Π35	Ε25	200
31	Π35	Ε25	100
31	Π76	Ε32	150
28	Π14	Ε68	150
22	Π76	Ε25	450
22	Π35	Ε32	670
22	Π35	Ε68	650
58	Π14	Ε68	670
28	Π14	Ε32	540
22	Π14	Ε25	200

Δεν υπάρχει ευρετήριο στο γνώρισμα της επιλογής αλλά το αρχείο είναι φυσικά ταξινομημένο στο γνώρισμα της επιλογής

Διαδική αναζήτηση: Αν η συνθήκη επιλογής περιλαμβάνει μια σύγκριση σε ένα γνώρισμα-κλειδί πάνω στο οποίο το αρχείο είναι ταξινομημένο, μπορεί να χρησιμοποιηθεί δυαδική αναζήτηση (η οποία είναι πιο αποτελεσματική από τη σειριακή αναζήτηση).

Με την δυαδική αναζήτηση μπορούμε να εντοπίσουμε την πρώτη πλειάδα που ικανοποιεί την συνθήκη επιλογής και μπορούμε να εντοπίσουμε τις υπόλοιπες σαρώνοντας το αρχείο από το σημείο αυτό.

ΠΑΡΑΓΓΕΛΙΑ (ΠΑΡ)

$\sigma_{\text{Ποσοτ}>200}$ (ΠΑΡ)

Το κόστος της δυαδικής αναζήτησης είναι:
 $O(\log_2 b)$

Και επιπλέον το κόστος της σειριακής αναζήτησης

Κ_Προμ	Κ_Πρ	Κ_Ερ	Ποσοτ
31	Π35	Ε25	100
31	Π76	Ε32	150
28	Π14	Ε68	150
22	Π14	Ε25	200
22	Π35	Ε25	200
22	Π76	Ε25	450
28	Π14	Ε32	540
22	Π35	Ε68	650
58	Π14	Ε68	670
22	Π35	Ε32	670

Χρήση Πρωτεύοντος ευρετηρίου ή κλειδιού κατακερματισμού

**Στην περίπτωση που το γνώρισμα της
συνθήκης επιλογής είναι πρωτεύον κλειδί και η
συνθήκη είναι συνθήκη ισότητας τότε μπορεί
να χρησιμοποιηθεί η οργάνωση του αρχείου (με
το όποιο κόστος της) για την εύρεση της
πλειάδας που ικανοποιεί την συνθήκη.**

**Το κόστος θα είναι το ύψος του ευρετηρίου
συν 1. Αν x το ύψος του ευρετηρίου θα είναι
 $x+1$**

Χρήση ενός πρωτεύοντος ευρετηρίου για ανάκτηση πολλών εγγραφών

Αν η συνθήκη σύγκρισης είναι $>$, \geq , $<$, ή \leq σε ένα πεδίο-κλειδί με πρωτεύον ευρετήριο χρησιμοποιούμε το ευρετήριο για να βρούμε την εγγραφή που ικανοποιεί την αντίστοιχη συνθήκη ισότητας και στη συνέχεια ανακτούμε όλες τις εγγραφές που ακολουθούν στο ταξινομημένο αρχείο.

Για παράδειγμα στην ερώτηση $\sigma_{K_Προμ>28}$ (IPM) μπορεί να χρησιμοποιηθεί το ευρετήριο για την εύρεση της πρώτης εγγραφής και στη συνέχεια να γίνει σάρωση.

Χρήση ευρετηρίων συστάδων για ανάκτηση πολλών εγγραφών

Αν η συνθήκη επιλογής περιλαμβάνει μια συνθήκη ισότητας σε ένα γνώρισμα που δεν είναι κλειδί αλλά έχει ορισθεί σ' αυτό ένα ευρετήριο συστάδων, χρησιμοποιούμε το ευρετήριο συστάδων για την ανάκτηση όλων των εγγραφών που ικανοποιούν τη συνθήκη επιλογής.

Για παράδειγμα η $\sigma_{\text{Ποσοτ}>200}$ (ΠΑΡ) και υπάρχει ευρετήριο συστάδα στην ποσότητα

Το κόστος θα είναι x (ύψος ευρετηρίου) συν την ανάκτηση των s εγγραφών που ικανοποιούν την συνθήκη s/bfr

Χρήση δευτερεύοντος ευρετηρίου (B+-δένδρου) σε μια συνθήκη ισότητας

Αυτή η μέθοδος αναζήτησης μπορεί να χρησιμοποιηθεί για να ανακτηθεί μία μοναδική εγγραφή, αν το γνώρισμα της συνθήκης είναι κλειδί (έχει μοναδικές τιμές) ή για να ανακτηθούν πολλαπλές εγγραφές αν το γνώρισμα της συνθήκης δεν είναι κλειδί. Επιπλέον μπορεί να χρησιμοποιηθεί για ανάκτηση εγγραφών, προκειμένου για συνθήκες που περιλαμβάνουν τους τελεστές $>$, $>=$, $<$, ή $<=$.

Πολύπλοκες Συνθήκες Επιλογής

Γενικά η συνθήκη επιλογής μπορεί να είναι μια λογική παράσταση. Μπορεί να έχει την μορφή:

συνθήκη1 op συνθήκη2 op ... συνθήκηN

όπου το op είναι ο λογικός τελεστής AND ή OR.

Για παράδειγμα

σ Είδος=7 AND Πόλη='Άργος' (ΠΡΜ)

Μια τέτοια παράσταση μπορούμε να την μετατρέψουμε σε συζευκτική μόνο

Συζευκτική επιλογή με χρήση ατομικού ευρετηρίου

Αν ένα γνώρισμα που περιλαμβάνεται σε οποιαδήποτε απλή συνθήκη της συζευκτικής συνθήκης έχει μια διαδρομή προσπέλασης που επιτρέπει τη χρήση μιας από τις μεθόδους όπως και σε απλή συνθήκη, χρησιμοποιούμε αυτή τη συνθήκη για την ανάκτηση των εγγραφών και στη συνέχεια ελέγχουμε αν κάθε ανακτηθείσα εγγραφή ικανοποιεί τις υπόλοιπες απλές συνθήκες της συζευκτικής συνθήκης.

σ Είδος=7 AND Πόλη='Αργος' (IPM)

Αν υπάρχει ευρετήριο στο Είδος ή την Πόλη

Συζευκτική επιλογή με χρήση σύνθετου ευρετηρίου

Αν δύο ή περισσότερα γνωρίσματα συμμετέχουν σε συνθήκες ισότητας στη συζευκτική συνθήκη και υπάρχει σύνθετο ευρετήριο (ή δομή κατακερματισμού) στον συνδυασμό των πεδίων.

σ Είδος=7 AND Πόλη='Άργος' (IPM)

Αν υπάρχει σύνθετο ευρετήριο στο Είδος και Πόλη, το σύνθετο ευρετήριο θα δώσει άμεσα τις πλειάδες που πληρούν και τις δύο συνθήκες

Συζευκτική επιλογή με τομή των δεικτών εγγραφών

Αν υπάρχουν δευτερεύοντα ευρετήρια σε όλα (ή σε μερικά) πεδία που ενέχονται σε συνθήκες σύγκρισης ισότητας της συζευκτικής συνθήκης κι αν τα ευρετήρια περιλαμβάνουν δείκτες εγγραφών (αντί δείκτες μπλοκ) τότε κάθε ευρετήριο μπορεί να χρησιμοποιηθεί για ανάκτηση ενός συνόλου δεικτών εγγραφών που ικανοποιούν τη μεμονωμένη απλή συνθήκη. Η τομή αυτών των συνόλων δεικτών δίνει τους δείκτες εγγραφών που ικανοποιούν τη συζευκτική συνθήκη. Οι δείκτες της τομής χρησιμοποιούνται στη συνέχεια για την άμεση ανάκτηση αυτών των εγγραφών. Αν μερικές μόνον από τις συνθήκες έχουν δευτερεύοντα ευρετήρια, κάθε ανακατωμένη εγγραφή ελέγχεται για να διαπιστωθεί αν ικανοποιεί και τις υπόλοιπες συνθήκες

σ Χρώμα='Μαύρο' AND Βάρος=10 (ΠΡ)

Αν υπάρχει ευρετήριο και στο χρώμα και στο βάρος μπορούν να βρεθούν οι δυο λίστες μπλοκ που ικανοποιούν τις συνθήκες και οι ζητούμενες είναι η τομή τους.

**Η ORACLE χρησιμοποιεί διάφορες τεχνικές για συζευκτικές συνθήκες. Μια είναι αυτή που εφαρμόζει AND σε bitmaps. Μια άλλη χρησιμοποιεί συνένωση με κατακερματισμό (θα την εξετάσουμε παρακάτω)στις δυο λίστες.
Ο SQL Server υλοποιεί συνένωση ευρετηρίων.**

Διαζευκτική Συνθήκη

Μια διαζευκτική συνθήκη (όπου οι απλές συνθήκες συνδέονται με το λογικό τελεστή OR) είναι πολύ πιο δύσκολη σε ό,τι αφορά την επεξεργασία της και τη βελτιστοποίηση.

Μπορούμε να κάνουμε εξαντλητική σάρωση της σχέσης.

Ή μπορεί να χρησιμοποιήσουμε διάφορα ευρετήρια.

Ύστερ κερκρίμένα αν υπάρχουν ευρετήρια για γνωρίσματα της σύζευξης μπορούμε να τα χρησιμοποιήσουμε για περιορίσουμε την αναζήτηση

σ (Χρώμα='Άσπρο OR Πρ_Ονομα='DVD') AND Βάρος=25 (ΠΡ)

Αν κάθε όρος της διάζευξης έχει ευρετήριο μπορούμε να αποφύγουμε την σάρωση της σχέσης. Χρησιμοποιούμε το ευρετήριο κάθε συνθήκης. Βρίσκουμε έτσι την λίστα των διευθύνσεων που ικανοποιούν κάθε συνθήκη και στη συνέχεια βρίσκουμε την ένωση τους.

Ο SQL Server τέτοιες ενώσεις καθώς και bitmaps με διαζευκτικές συνθήκες.

ORACLE

- **Μετατροπή της ερώτησης σε ένωση ερωτήσεων χωρίς OR.**
- **Αν οι συνθήκες της διάζευξης περιέχουν το ίδιο γνώρισμα χρησιμοποίηση της δομής του ευρετηρίου.**
- **Χρήση bitmaps με διαζευκτικές συνθήκες**

Η Sybase χρησιμοποιεί ενώσεις λιστών