
Προηγμένα Θέματα Βάσεων Δεδομένων

Διδάσκων: Άγγελος Μιχάλας

Περιεχόμενα

- Βασικά Τμήματα του ΣΔΒΔ
- Τύποι Εγγραφών
- Τύποι Σελίδων
- Ομαδοποίηση Εγγραφών
- Διαχείριση Χώρου Δίσκου
- Διαχείριση Απομονωτικής Μνήμης

ΣΔΒΔ και λειτουργίες φυσικού επιπέδου

- Το λειτουργικό σύστημα είναι υπεύθυνο για τη διαχείριση των αρχείων σε χαμηλό επίπεδο.
- Περιορισμοί που επηρεάζουν τα ΣΔΒΔ.
 - Σε σύστημα 32 Bit το μεγαλύτερο μέγεθος αρχείου μπορεί να είναι μέχρι 4GB. Εμπορικές εφαρμογές μπορεί να απαιτούν μεγαλύτερα αρχεία
 - Τα ΣΔΒΔ πρέπει να λειτουργούν σε διαφορετικά ΛΣ → δεν πρέπει να προσαρμόζουν το λογισμικό στις ιδιαιτερότητες κάθε ΛΣ.
- Λύση
 - Τα **ΣΔΒΔ παρέχουν** κάποιες **βασικές λειτουργίες χαμηλού επιπέδου** → εκτελούν λειτουργίες φυσικού επιπέδου

Τα Βασικά Τμήματα του ΣΔΒΔ

- Αποτελείται από πολλά υποσυστήματα
 - Μερικές από τις **βασικές απαιτήσεις του ΣΔΒΔ καλύπτονται** από τις **λειτουργίες του ΛΣ**
 - **Οι πολυπλοκότερες διεργασίες καλύπτονται από το ΣΔΒΔ** χρησιμοποιώντας ως βάση τις λειτουργίες του ΛΣ.
- **Μεταγλωττιστής ερωτημάτων DML (INSERT, UPDATE, DELETE, SELECT, MERGE, CALL, EXPLAIN PLAN, LOCK TABLE)**
 - Είναι υπεύθυνο για τη μεταγλώττιση του ερωτήματος του χρήστη.
 - Το ερώτημα δίνεται σε γλώσσα χειρισμού δεδομένων και μετατρέπεται σε εσωτερική αναπαράσταση κατάλληλη για επεξεργασία.
 - Οι μεταγλωττισμένες εντολές προωθούνται στη μηχανή εκτέλεσης.

Τα Βασικά Τμήματα του ΣΔΒΔ

- **Μεταγλωττιστής εντολών DDL (CREATE, ALTER, DROP, TRUNCATE, COMMENT, RENAME)**
 - Είναι υπεύθυνος για τη μεταγλώττιση των εντολών της γλώσσας ορισμού δεδομένων
 - Η εκτέλεση των εντολών δημιουργεί το σχήμα της ΒΔ.
 - Οι μεταγλωττισμένες εντολές προωθούνται στη μηχανή εκτέλεσης.
- **Μηχανή εκτέλεσης**
 1. Προσδιορίζεται ο τρόπος απάντησης του ερωτήματος του χρήστη
 2. Πριν την επεξεργασία του ερωτήματος προηγείται η διαδικασία της βελτιστοποίησης
 - Εναλλακτικοί τρόποι αναπαράστασης του ερωτήματος → Επιλέγεται αυτός που έχει το μικρότερο κόστος
 - Λαμβάνονται υπόψη παράγοντες όπως ο αναμενόμενος αριθμός προσπελάσεων στο δίσκο, το κόστος επεξεργασίας για ταξινόμηση των αποτελεσμάτων

Τα Βασικά Τμήματα του ΣΔΒΔ

- Έλεγχος ταυτοχρονισμού
 - Πολλά ερωτήματα αναφέρονται στο ίδιο σύνολο δεδομένων για ανάγνωση ή εγγραφή.
 - Το υποσύστημα προσδιορίζει αν το ερώτημα πρέπει να εκτελεστεί άμεσα ή πρέπει να καθυστερήσει μέχρι να εκτελεστούν άλλες λειτουργίες.
 - Χρησιμοποιεί μηχανισμούς κλειδώματος

Τα Βασικά Τμήματα του ΣΔΒΔ

- Διαχειριστής επανάκτησης και Ημερολογίου
 - Υπάρχουν περιπτώσεις όπου ο τερματισμός μιας συναλλαγής **δεν πραγματοποιείται ομαλά**
 - Μετά την επαναφορά του συστήματος απαιτείται έλεγχος για την ακεραιότητα των δεδομένων
 - Το ΣΔΒΔ καταχωρεί όλες τις ενέργειες σε ημερολόγιο για να μπορεί να επαναλάβει ή να απορρίπτει ορισμένες.

Τα Βασικά Τμήματα του ΣΔΒΔ

- Διαχειριστής συναλλαγών
 - Μία συναλλαγή αποτελείται από μία ή περισσότερες λειτουργίες που πρέπει να ολοκληρωθούν.
 - Αν μια από τις λειτουργίες αποτύχει, πρέπει να ακυρωθεί ολόκληρη συναλλαγή.
- Διαχειριστής αρχείων
 - Ο Διαχειριστής Αρχείων είναι υπεύθυνος για την τακτοποίηση των εγγραφών.

Τα Βασικά Τμήματα του ΣΔΒΔ

- Διαχειριστής απομονωτικής μνήμης
 - Το υποσύστημα υποστηρίζει τα τμήματα της ΒΔ που θα βρίσκονται στη κύρια μνήμη.
 - Ο Διαχειριστής Μνήμης Απομόνωσης είναι υπεύθυνος για την κατάτμηση της κύριας μνήμης.

Τα Βασικά Τμήματα του ΣΔΒΔ

- Διαχειριστής δίσκου/αποθήκης
 - Ο Διαχειριστής Αποθήκης είναι υπεύθυνος για την **ανάγνωση (read)** και την **αποθήκευση (write)** σελίδων στο δίσκο.
 - Είναι υπεύθυνος για τον **τρόπο φυσικής αποθήκευσης των δεδομένων στο δίσκο** και τον έλεγχο της μετακίνησης από και προς το δίσκο.

Τα Βασικά Τμήματα του ΣΔΒΔ

Η ΒΔ σε φυσικό επίπεδο αποτελείται από 4 διαφορετικά τμήματα δεδομένων που είναι

- **Αρχεία δεδομένων**

- Είναι ο κύριος χώρος αποθήκευσης των δεδομένων της ΒΔ.
Εδώ αποθηκεύονται οι πίνακες

- **Κατάλογοι**

- Περιλαμβάνουν μεθόδους προσπέλασης για τη διευκόλυνση της διαδικασίας αναζήτησης των δεδομένων.
 - Π.χ. Δενδρικές δομές προσπέλασης
 - Π.χ. Κατακερματισμός

Τα Βασικά Τμήματα του ΣΔΒΔ

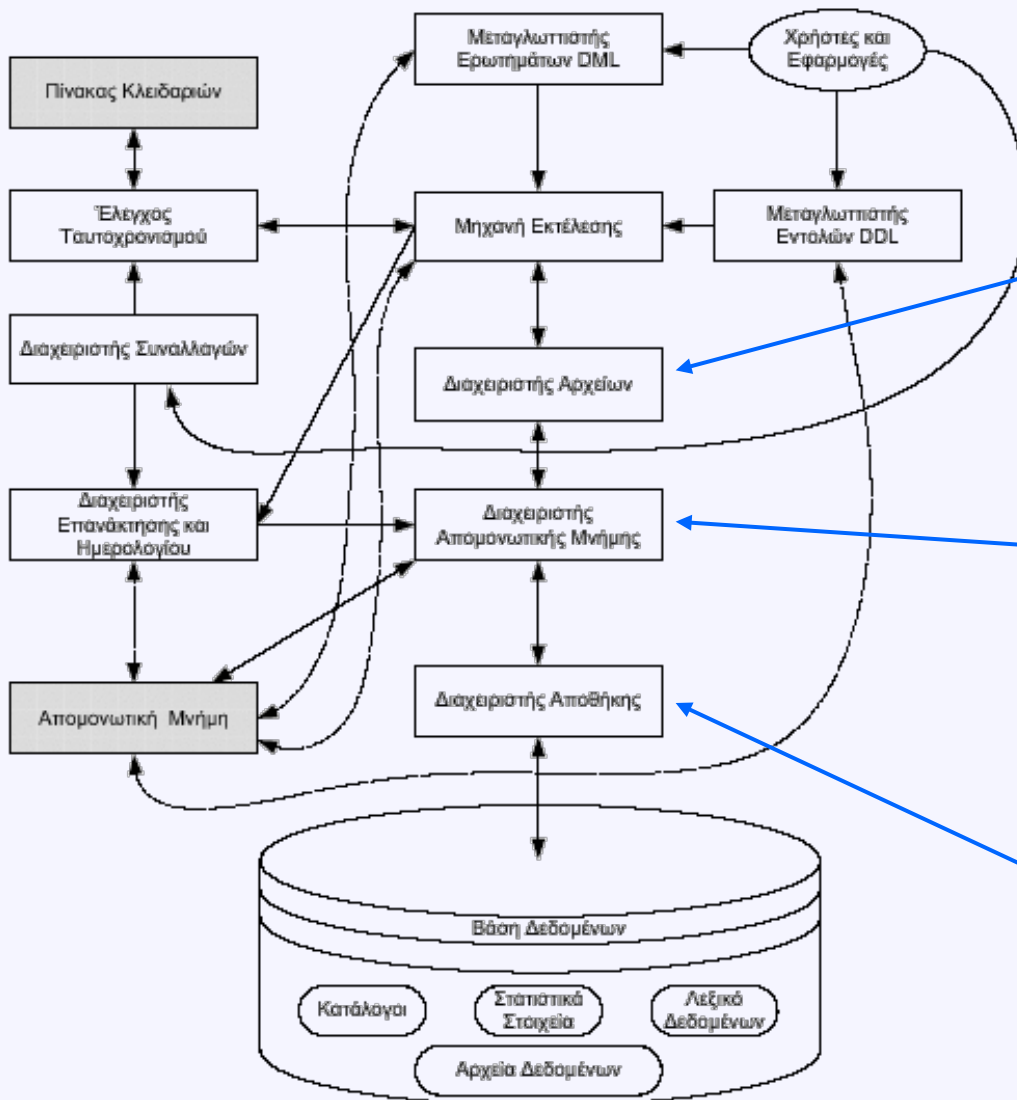
- Στατιστικά στοιχεία

- Είναι δεδομένα που αποθηκεύει το ίδιο το ΣΔΒΔ με στόχο την εκτίμηση του κόστους των ερωτημάτων

- Λεξικό δεδομένων

- Το αποτέλεσμα της μετάφρασης των εντολών της γλώσσας ορισμού δεδομένων αποθηκεύεται σε ειδικό αρχείο που λέγεται λεξικό δεδομένων.

Τα Βασικά Τμήματα του ΣΔΒΔ

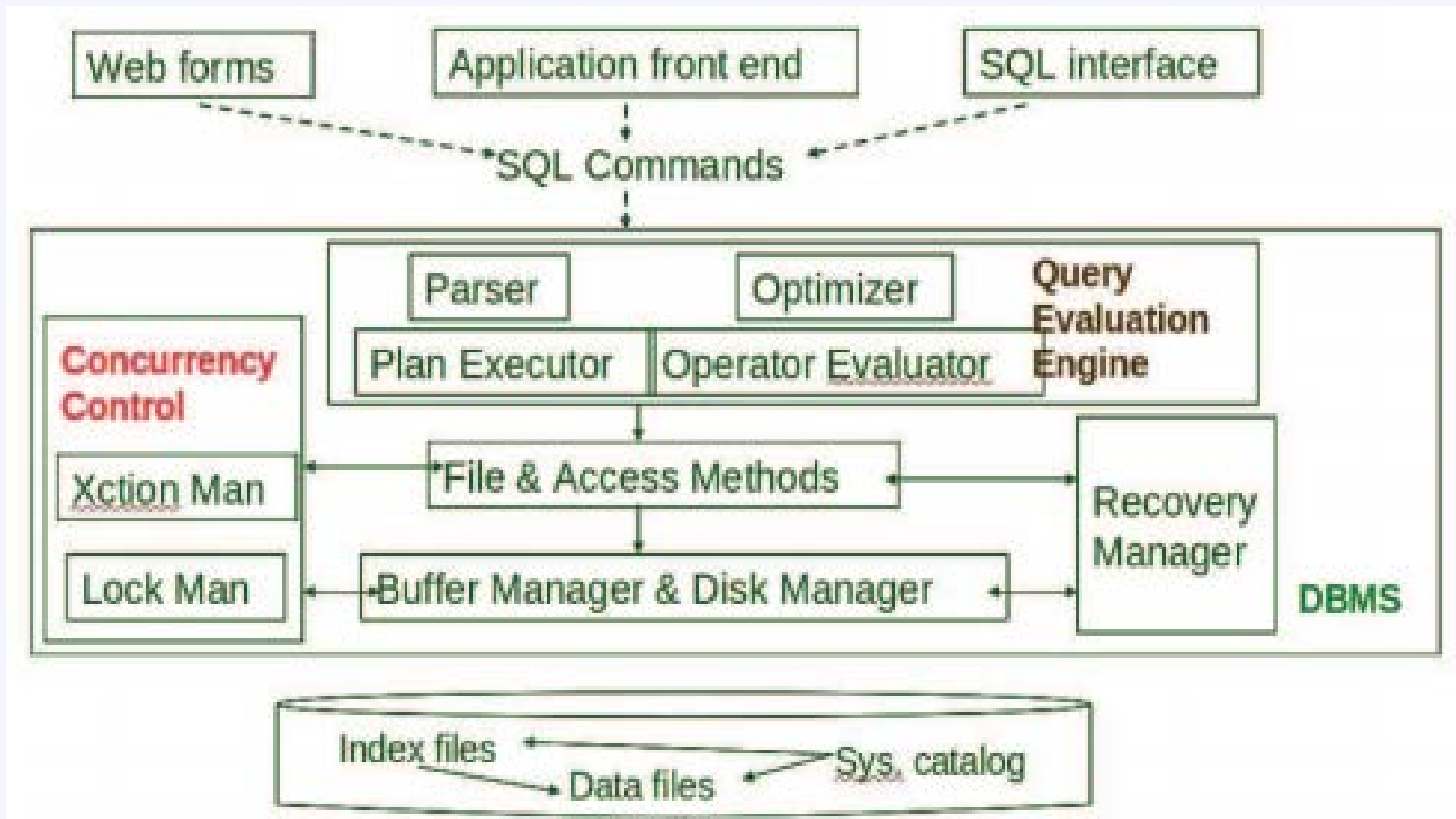


Ο Διαχειριστής Αρχείων είναι υπεύθυνος για την τακτοποίηση των εγγραφών.

Ο Διαχειριστής Μνήμης Απομνημονεύσιμης Μνήμης είναι υπεύθυνος για την κατάτμηση της κύριας μνήμης.

Ο Διαχειριστής Αποθήκης είναι υπεύθυνος για την ανάγνωση και την αποθήκευση σελίδων στο δίσκο.

Τα Βασικά Τμήματα του ΣΔΒΔ



Τύποι Εγγραφών

Τύποι Εγγραφών

- Η γραφική παράσταση της οργάνωσης των πεδίων μιας εγγραφής λέγεται **γραμμαγράφηση** (layout).
 - Εγγραφές με ίδια και ισομήκη πεδία αλλά διαφορετική διάταξη δεν έχουν την ίδια γραμμογράφηση.
- Οι εγγραφές μιας οντότητας έχουν ίδια μορφή και λέγονται **εγγραφές σταθερού μήκους**.
 - Συνήθως στα **σχεσιακά ΣΔΒΔ**

Όνομα	Διεύθυνση	Κωδικός	Τμήμα	Μισθός
-------	-----------	---------	-------	--------

|—30 bytes—|—40 bytes—|—4 bytes—|—3 bytes—|—4 bytes—|

Τύποι Εγγραφών

- Ο υπολογισμός μιας διεύθυνσης για την επεξεργασία του πεδίου μιας εγγραφής γίνεται όπως και κατά τον υπολογισμό διευθύνσεων σε πίνακες.
- Μερικές φορές στα σχεσιακά μοντέλα καθώς και στα αντικειμενοστραφή συστήματα προκύπτουν **εγγραφές μεταβλητού μήκους**.

Πλεονεκτήματα Εγγραφών Μεταβλητού Μήκους

- Τα πλεονεκτήματα ως προς την αποθήκευση είναι ιδιαίτερα σημαντικά όταν:
 - υπάρχει μεγάλη απόκλιση των μηκών των εγγραφών από το μέσο μήκος εγγραφής,
 - τα αρχεία είναι ογκώδη,
 - είναι μεγάλη η συχνότητα χρήσης,
 - το υλικό είναι ακριβό.

Μεταβλητότητα Μήκους Εγγραφών

Οφείλεται σε 3 λόγους (για τα σχεσιακά συστήματα ενδιαφέρον παρουσιάζει η 1^η περίπτωση)

- Πεδία μεταβλητού μήκους

- **Για κάθε εγγραφή δεσμεύεται χώρος ίσος με το μήκος του μεγαλύτερου στιγμιότυπου της εγγραφής**

- Στα συστήματα που δεν υποστηρίζουν εγγραφές σταθερού μήκους

- ➔ Τα συστήματα μεταβλητού μήκους **εξοικονομούν χώρο** σε πεδία που το μήκος ποικίλει πολύ.

- Συναντάται συνήθως σε πεδία τύπου συμβολοσειρές, όπως ονόματα, διευθύνσεις κλπ.

Μεταβλητότητα Μήκους Εγγραφών

- Εγγραφές διαφορετικής μορφής
 - Είναι πιθανό 2 εγγραφές να έχουν το ίδιο μήκος αλλά να είναι διαφορετικής μορφής
 - Λύση: **Η χρήση ενός επιπλέον πεδίου που να δηλώνει τη μορφή της εγγραφής**
 - Ένα αρχείο μπορεί να περιέχει εγγραφές από δύο πίνακες διαφορετικού σχήματος → εγγραφές διαφορετικού μήκους
 - Λύση: **επιπλέον πεδίο που δηλώνει πίνακα και εγγραφή**

Τύπος=Ω	όνομα	διεύθυνση	τμήμα	ωρομίσθιο	ώρες
Τύπος=M	όνομα	διεύθυνση	τμήμα	μισθός	

Μεταβλητότητα Μήκους Εγγραφών

- Επαναλαμβανόμενες ομάδες πεδίων
 - Το φαινόμενο συμβαίνει όταν ένα ή περισσότερα χαρακτηριστικά επαναλαμβάνονται περισσότερες από 2 φορές.
 - Τα σχεσιακά ΣΔΒΔ δεν επιτρέπουν την ύπαρξη επαναλαμβανόμενων πεδίων

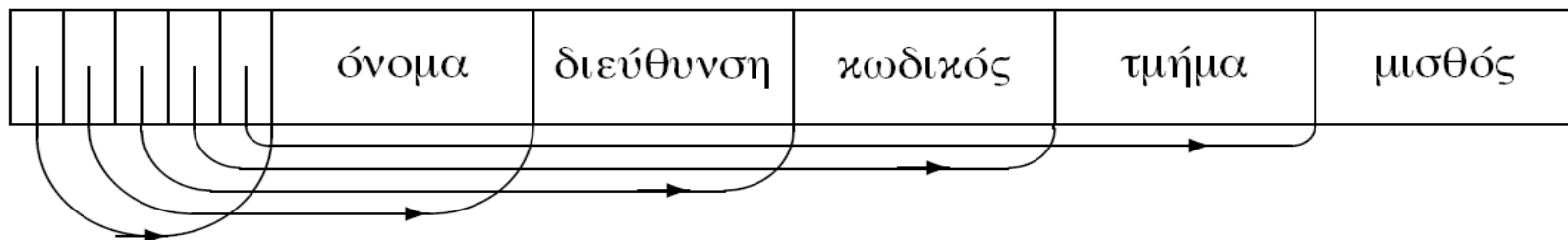
όνομα	κωδικός	όνομα προστατευόμενου	ημερομηνία γέννησης	όνομα προστατευόμενου	...
-------	---------	--------------------------	------------------------	--------------------------	-----

Διαχείριση Εγγραφών Μεταβλητού Μήκους

- Χρήση ειδικών διαχωριστών
 - Απαιτεί σειριακή σάρωση του αρχείου για τον εντοπισμό του πεδίου

όνομα	\$	διεύθυνση	\$	κωδικός	\$	τιμήμα	\$	μισθός
-------	----	-----------	----	---------	----	--------	----	--------

- Χρήση ενός μικρού καταλόγου στην αρχή της εγγραφής όπου για κάθε πεδίο αποθηκεύεται η απόστασή του από την αρχή της εγγραφής
 - Απαιτεί περισσότερο χώρο από την 1^η. Δίνει τη δυνατότητα άμεσης προσπέλασης κάθε πεδίου.



Πλεονεκτήματα Καταλόγου

1. Απαιτεί περισσότερο χώρο αλλά είναι χρονικά αποτελεσματικότερη γιατί προσφέρει άμεση προσπέλαση κάθε πεδίου
2. Δεν απαιτείται η αποθήκευση συγκεκριμένου συμβόλου για την τιμή NULL (μη ύπαρξη πραγματικής τιμής) κάποιου πεδίου, αλλά απλώς η απόσταση του επόμενου πεδίου είναι ίση με την απόσταση του συγκεκριμένου πεδίου.

Μειονεκτήματα Εγγραφών Μεταβλητού Μήκους

Κυριότερο πρόβλημα: αύξηση μεγέθους εγγραφής κατά την ενημέρωση ενός πεδίου.

Συνέπειες:

1. ανάγκη **τακτοποίησης του περιεχομένου της σελίδας** με την κατάλληλη **μετακίνηση** (shifting) **μερικών εγγραφών**.
2. ανάγκη **αποθήκευσης** της συγκεκριμένης **εγγραφής σε άλλη σελίδα** λόγω ανεπάρκειας χώρου.
 - αν η εγγραφή μετακινηθεί, τότε στη θέση της αποθηκεύεται η νέα διεύθυνση.