



Αποθήκευση Δεδομένων

ΜΕΡΟΣ Β': Το «εσωτερικό» ενός ΣΔΒΔ

Εισαγωγή



Δομή ενός ΣΔΒΔ

Η (εσωτερική) αρχιτεκτονική ενός ΣΔΒΔ είναι σε επίπεδα

Τυπικά, κάθε σχέση σε ένα αρχείο στο δίσκο

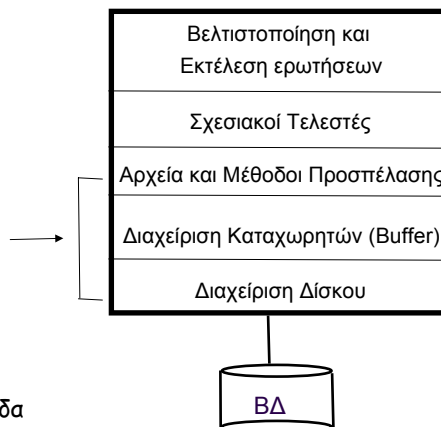
Σήμερα θα δούμε:

Αποθήκευση

Δομή αρχείων

Στη συνέχεια

Τα παραπάνω επίπεδα





Η βάση δεδομένων θα πρέπει να αποθηκευτεί σε κάποιο αποθηκευτικό μέσο

Ιεραρχία αποθήκευσης

πρωτεύουσα αποθήκευση (primary storage)

κύρια μνήμη (main memory) - κρυφή μνήμη (cache)

- άμεση προσπέλαση από την κύρια ΚΜΕ (CPU)
- γρήγορη προσπέλαση
- περιορισμένη χωρητικότητα αποθήκευσης



Δευτερεύουσα αποθήκευση

(μαγνητικοί δίσκοι, ταινίες, δισκέτες, κλπ)

- για την επεξεργασία των δεδομένων απαιτείται η μεταφορά των δεδομένων στην πρωτεύουσα αποθήκευση
- πιο αργή προσπέλαση
- μεγάλη χωρητικότητα
- μικρότερο κόστος (για την ίδια ποσότητα χώρου η κύρια μνήμη 100 φορές ακριβότερη από τη δευτερεύουσα)



Οι περισσότερες βάσεις δεδομένων αποθηκεύονται σε δευτερεύουσες αποθηκευτικές μονάδες κυρίως σε δίσκους

- πολύ μεγάλες (10-100 TB) \Rightarrow μεγάλο κόστος (\$1/GB - 100\$/GB)
- μόνιμη αποθήκευση (nonvolatile storage)

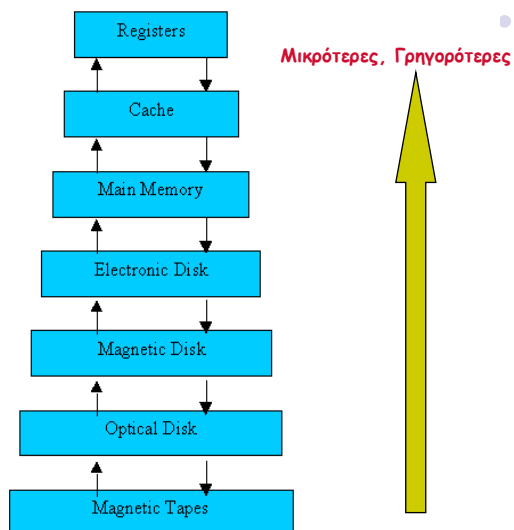
Μαγνητικές ταινίες για

- τήρηση εφεδρικών αντιγράφων
- αρχειοθέτηση (archiving) (δεδομένα που θέλουμε να κρατήσουμε για πολύ καιρό αλλά η προσπέλαση τους είναι σπάνια)



Ιεραρχία Αποθήκευσης

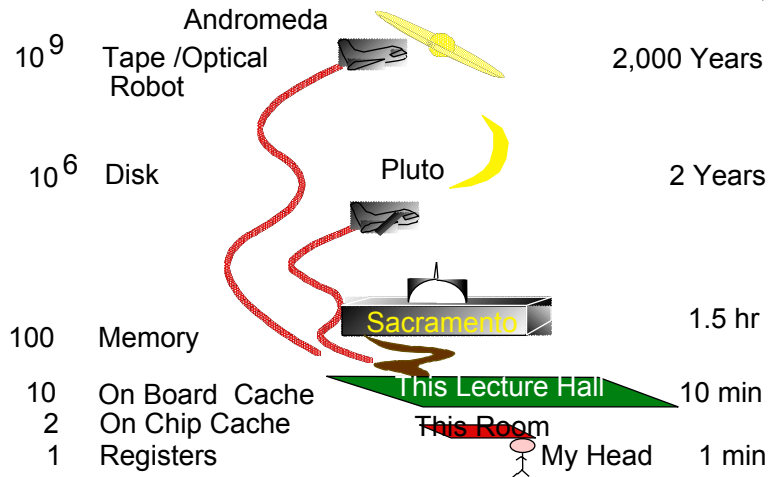
- Κύρια μνήμη (RAM) για τα δεδομένα σε χρήση
- Δίσκοι για τη βδ (δευτερεύουσα αποθήκευση)
- Ταινίες για παλιές εκδοχές δεδομένων (tertiary storage).



Source: Operating Systems Concepts 5th Edition

Αποθηκευτικές Μονάδες

Η αναλογία του Jim Gray's για latency (χρόνο απόκρισης) αποθήκευσης



Βάσεις Δεδομένων 2009-2010

Ευαγγελία Πιτουρά

7

Μαγνητικοί Δίσκοι

Μαγνητικοί Δίσκοι

- Μαγνητισμός μιας περιοχής του δίσκου κατά ορισμένο τρόπο ώστε 1 ή 0
- Χωρητικότητα (capacity) σε Kbyte - Mbyte - Gbyte
- Μαγνητικό υλικό σε σχήμα κυκλικού δίσκου



- Απλής και διπλής όψης

Βάσεις Δεδομένων 2009-2010

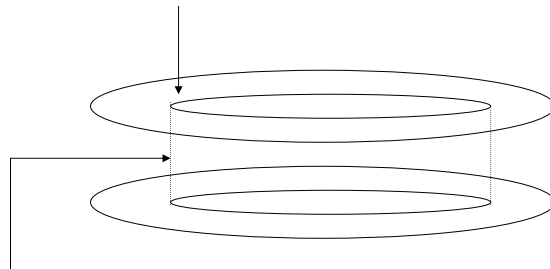
Ευαγγελία Πιτουρά

8



Σε πακέτα δίσκων

Οι πληροφορίες σε ομόκεντρους κύκλους διαφορετικής διαμέτρου: **άτρακτοι track** (συνήθως κάθε άτρακτος την ίδια ποσότητα πληροφορίας)

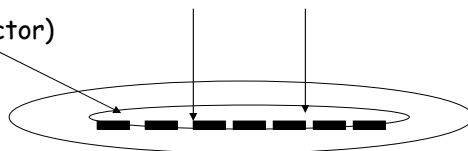


Ομόκεντροι κύκλοι σε διαφορετικές επιφάνειες: **κύλινδρος (cylinder)**



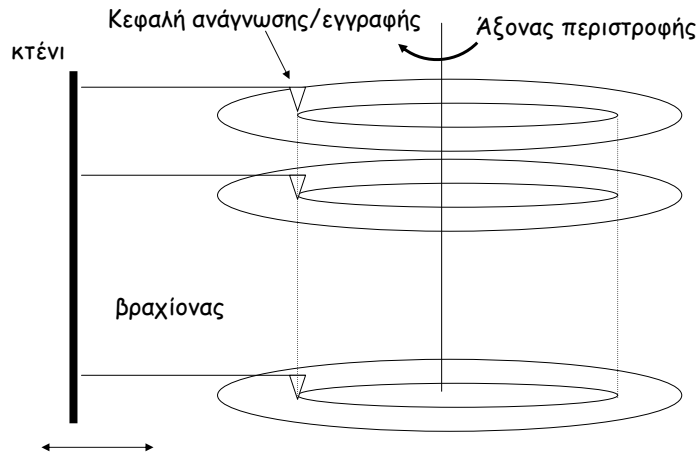
Block (μονάδα μεταφοράς)

Τομέας (sector)



Κάθε άτρακτος χωρίζεται σε τόξα που ονομάζονται **τομείς (sectors)** και είναι χαρακτηριστικό του κάθε δίσκου και δε μπορεί να τροποποιηθεί

Το μέγεθος ενός block τίθεται κατά την αρχικοποίηση του δίσκου και είναι κάποιο πολλαπλάσιο του τομέα



χρόνος εντοπισμού (seek time) Τοποθέτηση κεφαλής στη σωστή άτρακτο 0.3 - 10

χρόνος περιστροφής (rotational delay ή latency) Όσπου η αρχή του σωστού block να βρεθεί κάτω από την κεφαλή

χρόνος μεταφοράς block (block transfer time) χρόνος μεταφοράς δεδομένων από το δίσκο στη μνήμη

Χρόνος προσπέλασης = χρόνος εντοπισμού + χρόνος περιστροφής + χρόνος μεταφοράς

Μεταφορά αρκετών γειτονικών block

Μαγνητικοί Δίσκοι



Παράδειγμα **IBM Deskstar 14GPX** **Seagate Barracuda 7200.9**

Χωρητικότητα: 14.4 GB

80 - 500 GB

(μέσος) Χρόνος Εντοπισμού: 9.1 msec **11ms**

(2.2 για γειτονικά - 15.5 μέγιστο)

(μέσος) Χρόνος Περιστροφής: 4.17 msec **4.16ms**

5 διπλής όψης κυκλικούς δίσκους - 7,200 περιστροφές το λεπτό **7,200**

Χρόνος Μεταφοράς 13MB ανά sec

300MB ανά sec
(σειριακός)

Χρόνος προσπέλασης από το δίσκο ~ 10 msec (micro 10^{-6}) ενώ για θέσης μνήμης 60 nanosecond (nano 10^{-9})

Μαγνητικοί Δίσκοι



Συνήθως μόνο μία κεφαλή τη φορά

Disk controller

- λειτουργίες εγγραφής/ανάγνωσης
- υπολογισμός αθροίσματος ελέγχου (checksum)



Για διαμοιραζόμενους δίσκους: χρόνος στην ουρά του controller



Συμπεράσματα

1. Τα δεδομένα πρέπει να βρίσκονται στη μνήμη
2. Η μονάδα μεταφοράς από το δίσκο στη μνήμη είναι ένα block . Το διάβασμα ή γράψιμο ενός block ονομάζεται λειτουργία Εισόδου/Εξόδου (Input/Output - I/O)
3. Ο χρόνος προσπέλασης (εγγραφής ή ανάγνωσης) ενός block **διαφέρει και εξαρτάται από τη θέση του block**
χρόνος προσπέλασης = χρόνος εντοπισμού + χρόνου περιστροφής + χρόνος μεταφοράς



Μαγνητικές Ταινίες

- Δίσκοι *τυχαίας προσπέλασης* (random access)
- Ταινίες *σειριακής προσπέλασης* (serial access) για να διαβάσουμε το n -οστό block πρέπει να ξεκινήσουμε από την αρχή και να διαβάσουμε και τα $n-1$ blocks



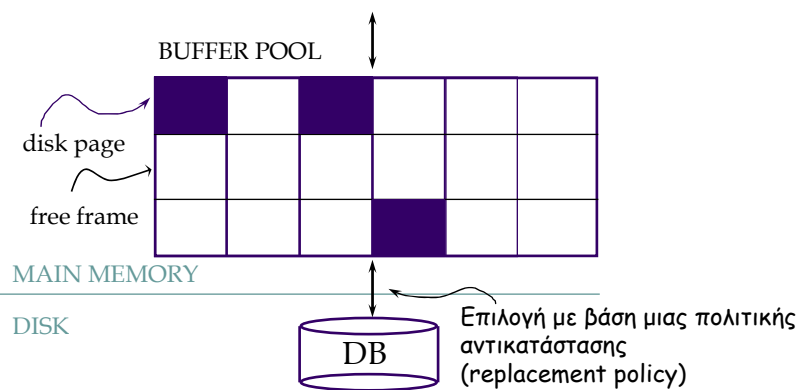
Μεταφορά block σε ενδιάμεση μνήμη

Ενώ γίνεται η μεταφορά των δεδομένων από την δευτερεύουσα στην κύρια μνήμη - παράλληλα και ανεξάρτητα η ΚΜΕ μπορεί να επεξεργάζεται δεδομένα

Ένας ανεξάρτητος επεξεργαστής Εισόδου/Εξόδου ή πολλαπλοί επεξεργαστές



Αιτήματα για σελίδες από τα υψηλότερα επίπεδα





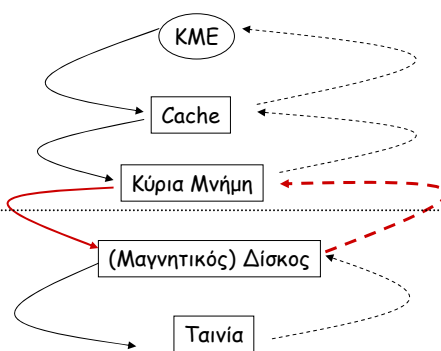
RAID: πλεονάζουσες συστοιχίες ανεξάρτητων δίσκων (καταμερισμός δεδομένων και πλεονασμός)



Πρωτεύουσα
Αποθήκευση

Δευτερεύουσα
Αποθήκευση

Αίτηση για
δεδομένα
→



Δεδομένα που
ικανοποιούν την
αίτηση
←