

Προχωρημένα Θέματα Τεχνολογίας και Εφαρμογών Βάσεων Δεδομένων



Εισαγωγικά του μαθήματος

Πάνος Βασιλειάδης

pvassil@cs.uoi.gr

Φεβρουάριος 2025

www.cs.uoi.gr/~pvassil/courses/db_III/

Θεματολόγιο

- Τι θα μας απασχολήσει στο μάθημα
- Διαδικαστικά του μαθήματος
- Η συνταγή της επιτυχίας

Θεματολόγιο

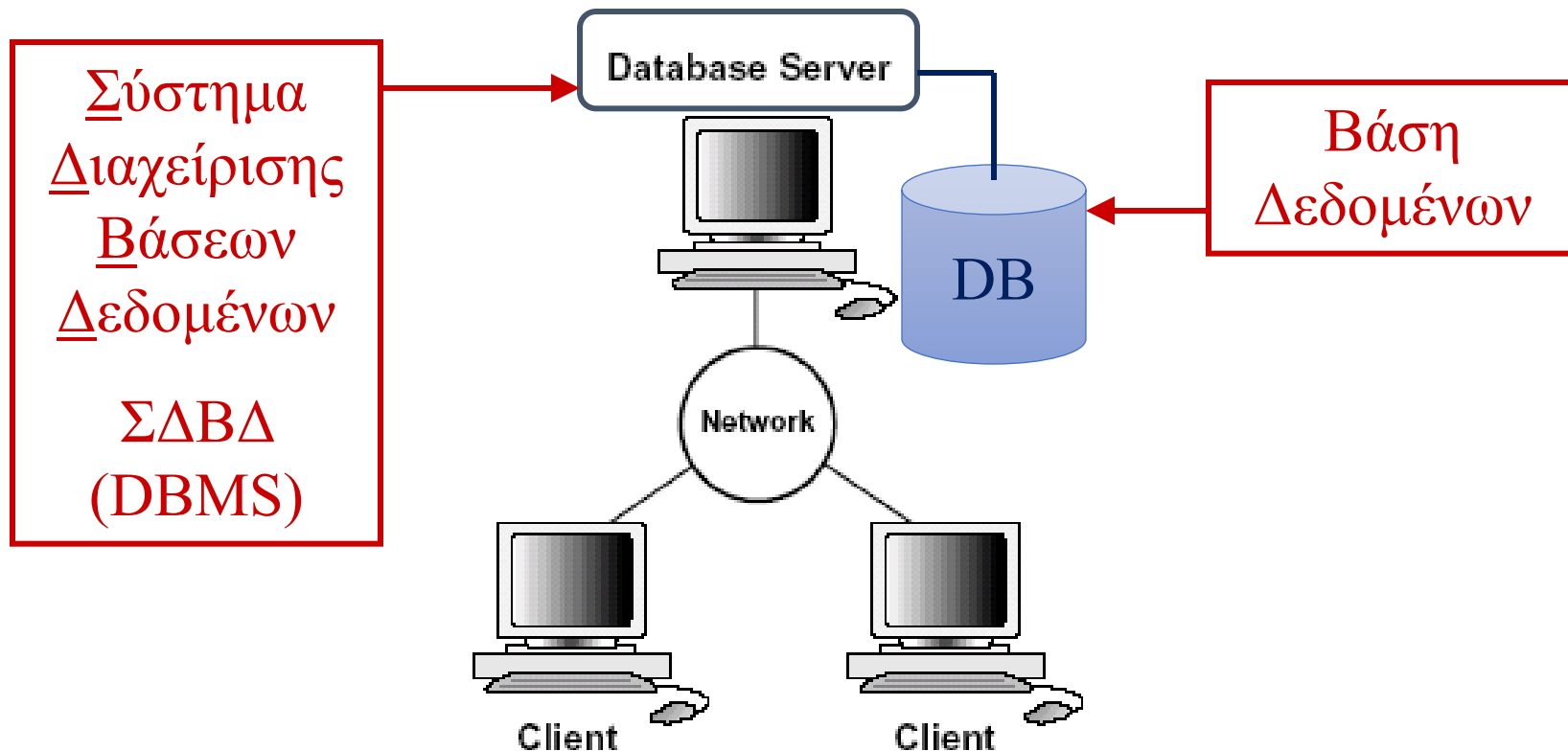
- Τι θα μας απασχολήσει στο μάθημα
- Διαδικαστικά του μαθήματος
- Η συνταγή της επιτυχίας

Τι θα μας απασχολήσει στο μάθημα

- ΜΕΡΟΣ Ι: Εσωτερική δομή ενός ΣΔΒΔ
 - Εσωτερική αρχιτεκτονική ΣΔΒΔ
 - Επεξεργασία και βελτιστοποίηση ερωτήσεων
 - Διαχείριση συναλλαγών και έλεγχος ταυτοχρονισμού
 - Ανάνηψη από αποτυχίες
- ΜΕΡΟΣ ΙΙ: Διαχείριση και Εφαρμογές βάσεων δεδομένων
 - Φυσική σχεδίαση και ρύθμιση μιας βάσης δεδομένων
 - Διαχείριση χρηστών και ασφάλεια
 - Αποθήκες δεδομένων

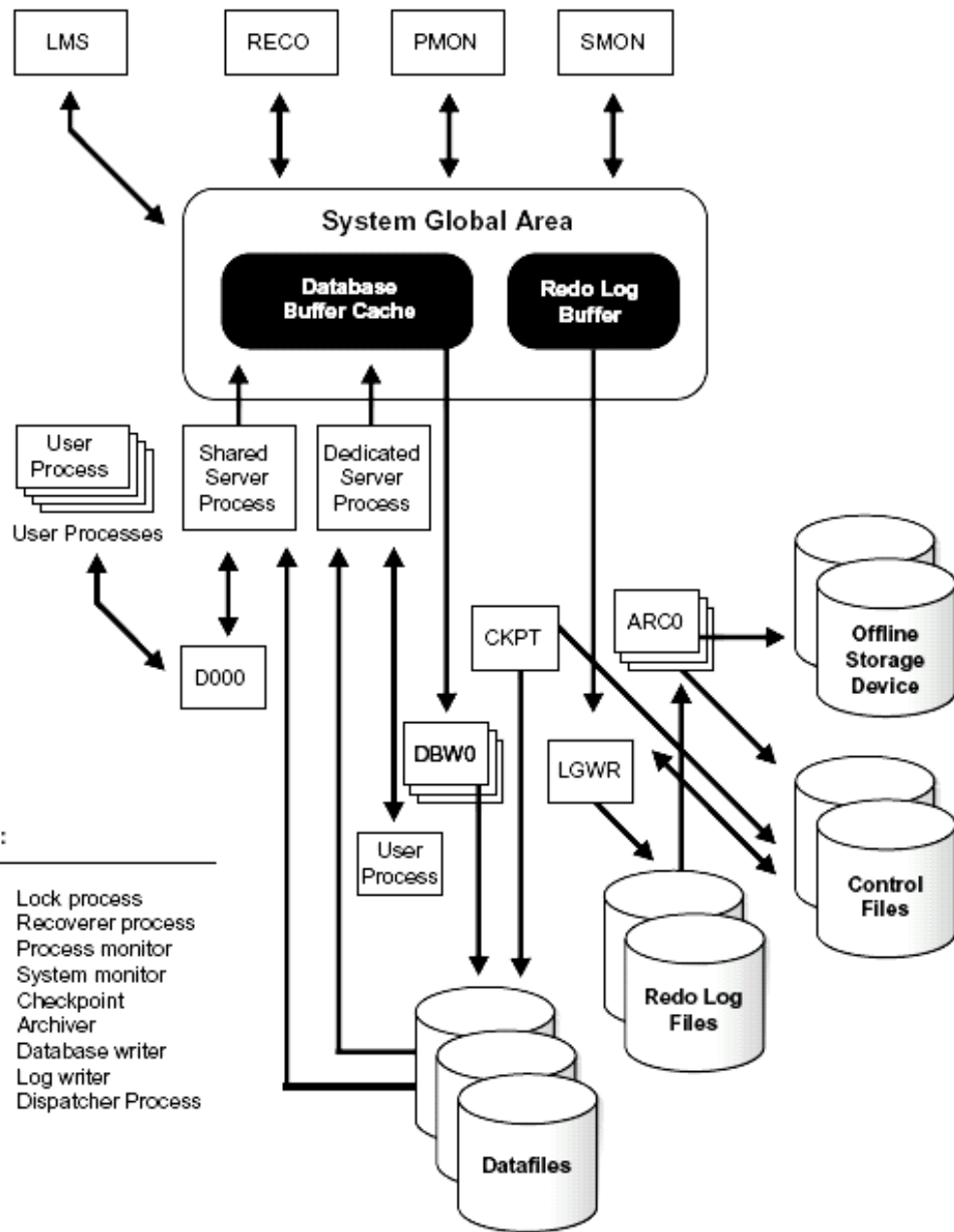
[όσο προλάβουμε ...]

Βάσεις Δεδομένων και ΣΔΒΔ

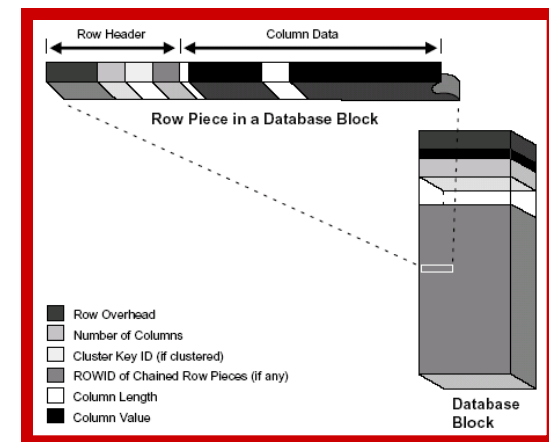


Η αρχιτεκτονική ενός DBMS

- Πώς δουλεύει εσωτερικά ένα DBMS?
- Από ποιες διαδικασίες / δομές αποτελείται?
- Γιατί είναι σημαντικό να το ξέρουμε?



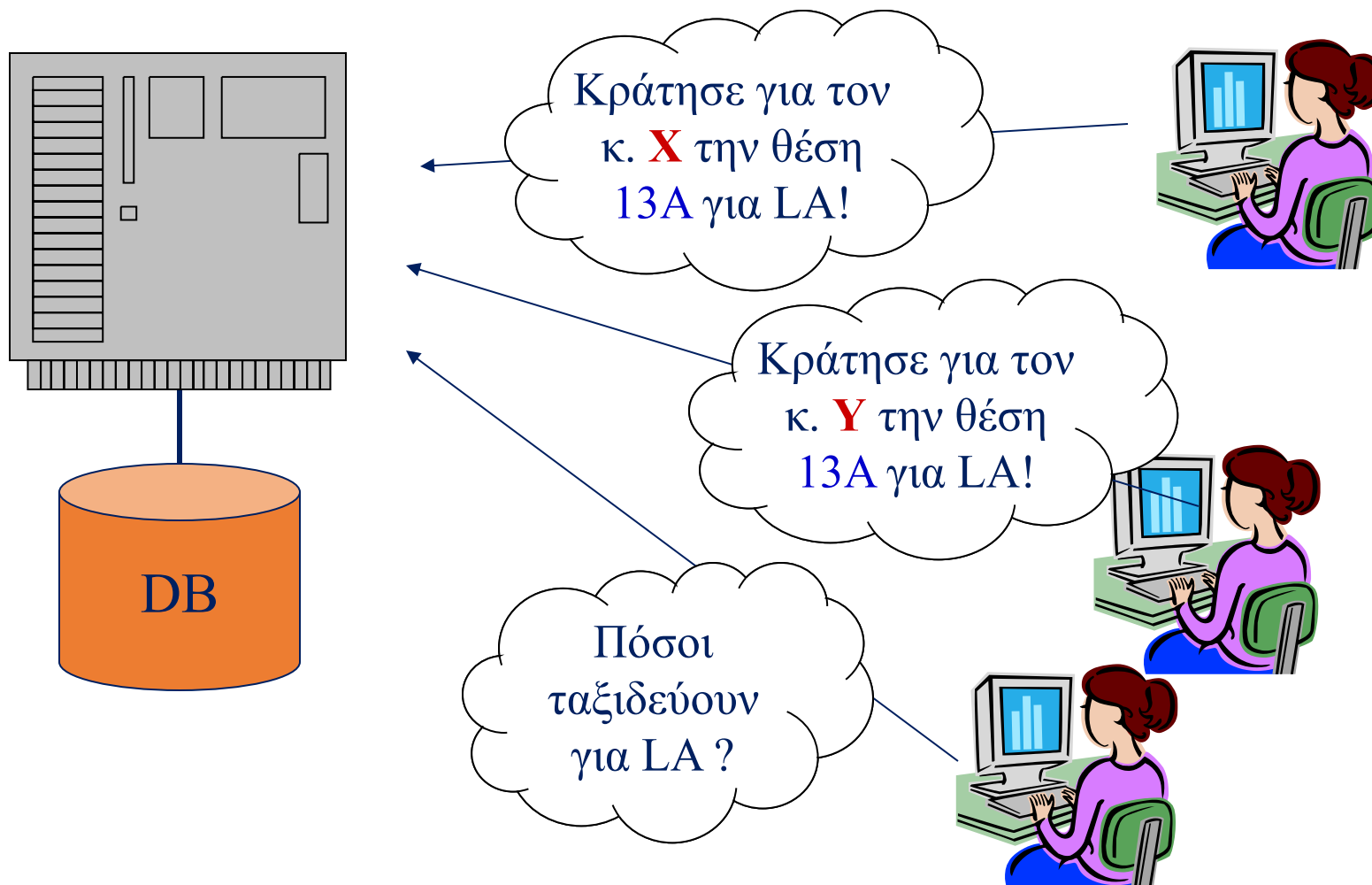
Τι πρέπει να
ξέρουν οι
DBA's και οι
application
developers
για τα
εσωτερικά
ενός DBMS?



Ταυτόχρονη πρόσβαση

- Πώς γίνεται ένας server να υποστηρίξει την ταυτόχρονη πρόσβαση πολλών χρηστών σε ένα DBMS?
- Ποια είναι τα προβλήματα που προκύπτουν?
- Γιατί είναι αυτό σημαντικό?

Συναλλαγές & Ταυτοχρονισμός



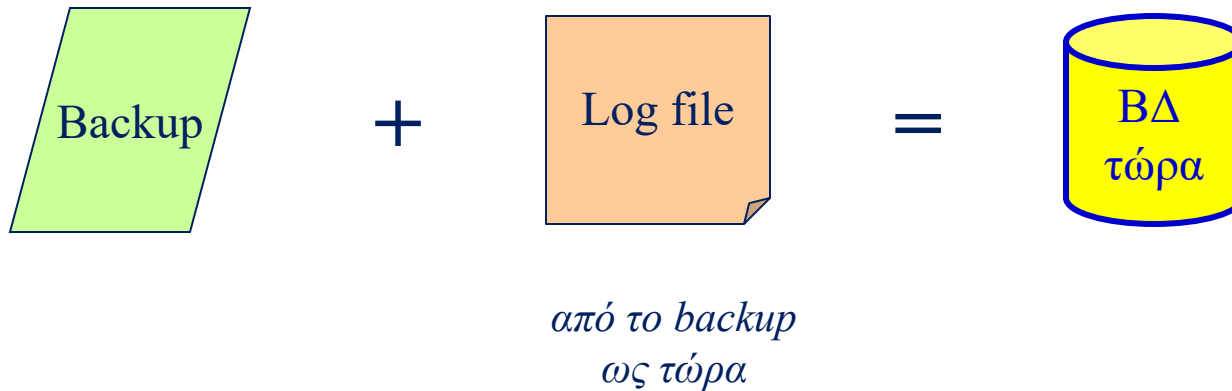
Συναλλαγές και SQL

- Στην SQL-92 κάθε συναλλαγή έχει:
 - Μέθοδο πρόσβασης: **READ ONLY** vs. **READ WRITE**
 - Ένα από 4 επιτρεπτά επίπεδα απομόνωσης:
 - **READ UNCOMMITTED**
 - **READ COMMITTED**
 - **REPEATABLE READ**
 - **SERIALIZABLE**
- Παράδειγμα χρήσης:
SET ISOLATION LEVEL SERIALIZABLE READ WRITE

Αξιοπιστία

- Πώς γίνεται ένα DBMS να εγγυάται ότι τα δεδομένα δε χάνονται?
- Πώς φέρνουμε το σύστημα σε λειτουργία μετά από αποτυχίες?
- Ποια είναι τα προβλήματα που προκύπτουν?
- Γιατί είναι αυτό (εξαιρετικά) σημαντικό?

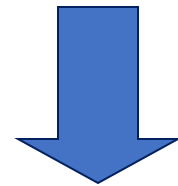
Η λύση στο πρόβλημα της αξιοπιστίας



Log Record τύπου “Write”

- LSN (μονοσήμαντο αναγνωριστικό του log entry)
- Transaction ID (TID)
- Σελίδα που κάνουμε update
- Offset στην εν λόγω σελίδα
- Μήκος σε bytes που αλλάζουμε
- Παλιά τιμή
- Νέα τιμή

**T1: UPDATE EMP
SET ID = 30
WHERE ID = 3**

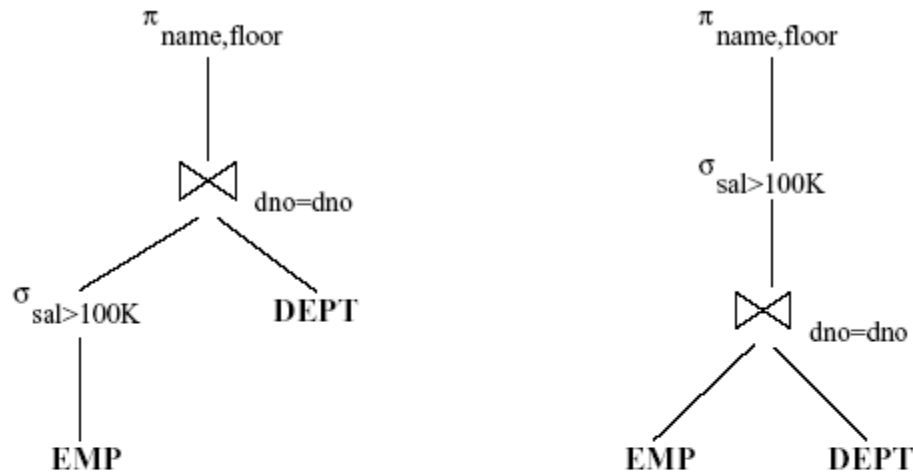


Log Entry: LSN12, T1, PAGE32, 0xFFF32, 8, 3, 30

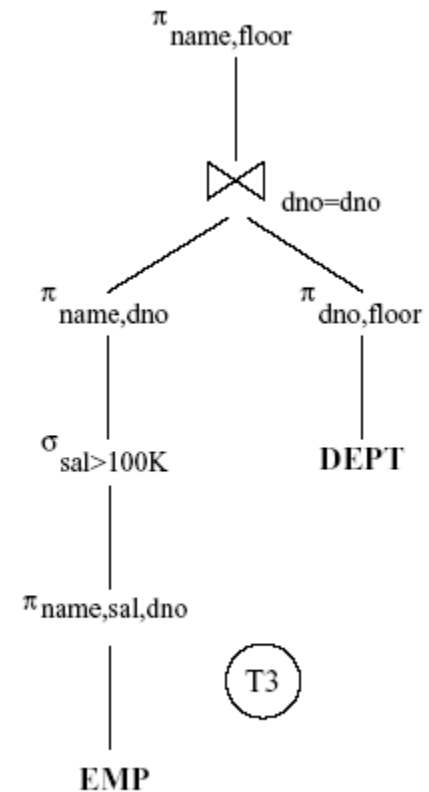
Επεξεργασία ερωτήσεων

- Πώς απαντά ένα DBMS ερωτήσεις?
- Τι γίνεται όταν υποβάλλεται μια ερώτηση στο DBMS?
- Ποιοι είναι οι διαθέσιμοι αλγόριθμοι?
- Γιατί είναι αυτό (εξαιρετικά) σημαντικό για developers και administrators?

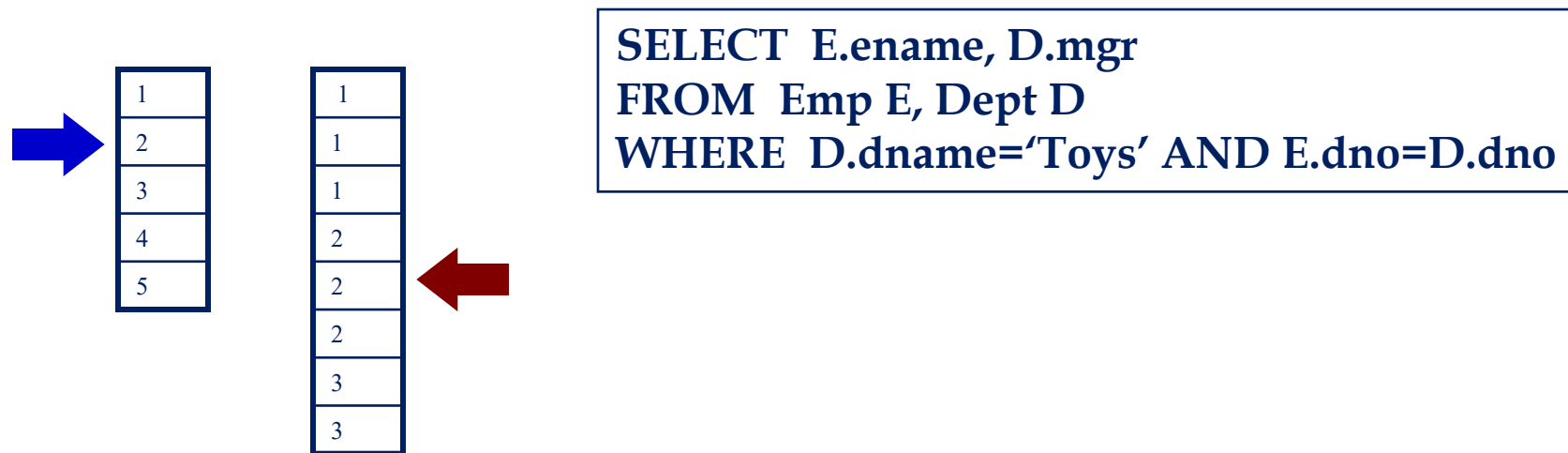
Παράδειγμα ισοδύναμων, εναλλακτικών πλάνων



```
select name, floor
from emp, dept
where emp.dno = dept.dno
and sal > 100K
```



Επεξεργασία Ερωτήσεων



Ταξινόμησε τα Emp, Dept με βάση το πεδίο Dno

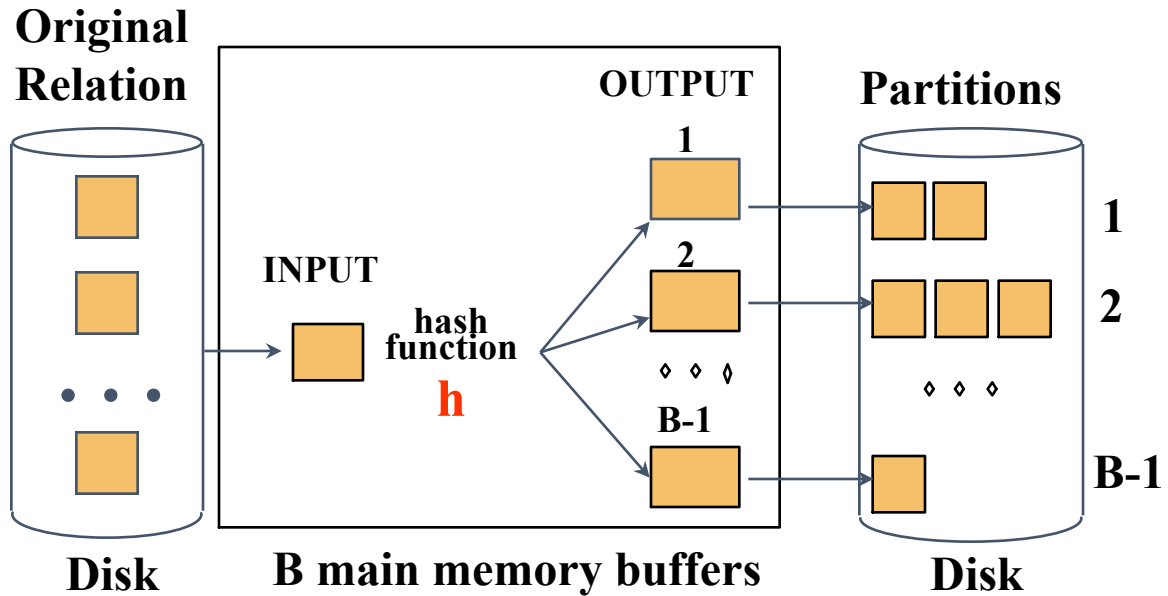
Για κάθε $D \in \text{Dept}$ [& Dname='Toys']

Για κάθε $E \in \text{Emp}$ με ίδιο Dno

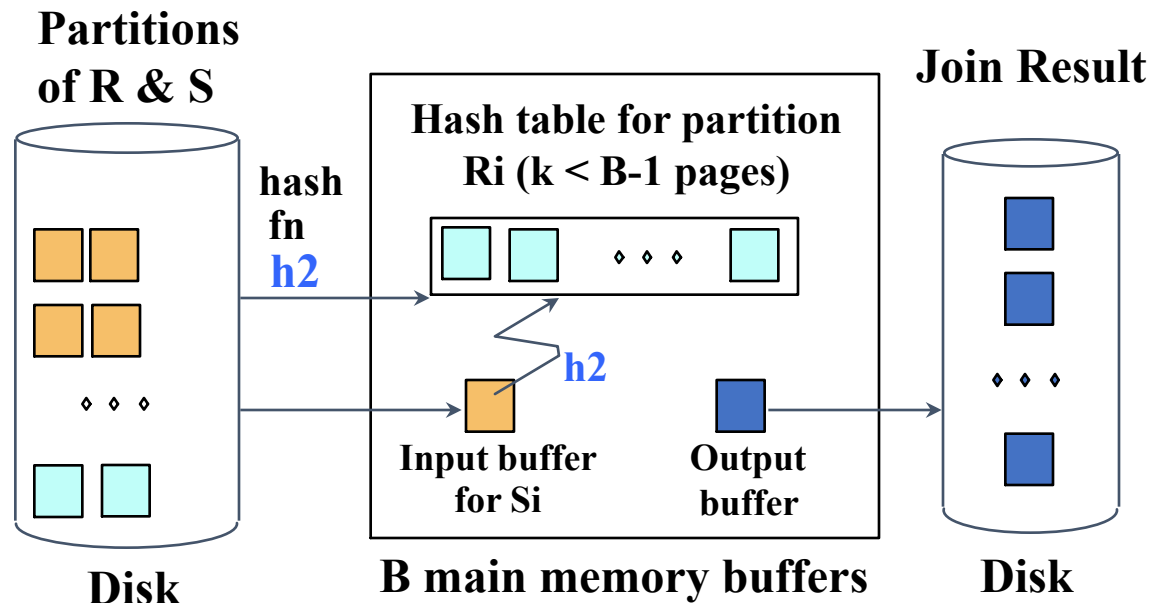
Επέστρεψε E.ename, D.mgr

Hash-Join

- Σπάσε και τις δύο σχέσεις μέσω της hash function **h**: οι εγγραφές της R στο partition **i** θα αντιστοιχούν μόνο με τις εγγραφές του partition **i** της S.



- ❖ Διάβασε ένα partition της R, κάνε του hash με μια συνάρτηση **h2** ($\neq h$!). Ψάξε το αντίστοιχο partition της S, και βρες τις εγγραφές που κάνουν match.



Άλλα θέματα

- Ρύθμιση της σχεδίασης
- Αποθήκες δεδομένων

Θεματολόγιο

- Τι θα μας απασχολήσει στο μάθημα
- Διαδικαστικά του μαθήματος
- Η συνταγή της επιτυχίας

Διαδικαστικά

- Ώρες μαθήματος
- Web
- Διδακτικά Βιβλία
- Διαγώνισμα
- Προγραμματιστική Εργασία
- Βαθμολογία
- ...

Διαδικαστικά

- Διδάσκων: Πάνος Βασιλειάδης
 - Γραφείο: B.2
 - E-mail: pvassil@cs.uoi.gr
- Web: www.cs.uoi.gr/~pvassil/courses/db_III/
 - Περιέχει ανακοινώσεις, ασκήσεις, βαθμολογίες, ...
- Αίθουσα: I2

MS Teams για το 2025: **7eqzqhe**

- Ώρες μαθήματος: Τρίτη, 17.00 – 20.00

Video wall

- Μπορείτε να δείτε κάποια videos στο [Panos Vassiliadis YouTube channel](https://www.youtube.com/channel/UCcAMEsKdL7fJcUewJiLD26A) (<https://www.youtube.com/channel/UCcAMEsKdL7fJcUewJiLD26A>) σε δύο playlist
- (α) [ΜΥΕ030: ΠΘΤΕΒΔ](https://www.youtube.com/playlist?list=PL3G-N7ZzyiDcpX2N7o4Gaf4uTC1qqon3P) <https://www.youtube.com/playlist?list=PL3G-N7ZzyiDcpX2N7o4Gaf4uTC1qqon3P>
για τις διαλέξεις και τις ασκήσεις του μαθήματος
- (β) [Κώδικας για διαχείριση δεδομένων](https://www.youtube.com/playlist?list=PL3G-N7ZzyiDesUARYDi0BupgywPzKEzwSw): κυρίως για το project του μαθήματος (και της Τεχν. Λογισμικού)
<https://www.youtube.com/playlist?list=PL3G-N7ZzyiDesUARYDi0BupgywPzKEzwSw>
την οποία θα συμπληρώνω, κατά το δυνατόν, με προγραμματιστικά θέματα

Βιβλιογραφία

- **Θεμελιώδεις αρχές συστημάτων βάσεων δεδομένων**, Συγγραφείς: Elmasri Ramez, Navathe Shamkant B., **Κωδικός Βιβλίου στον Εύδοξο: 22683637**, Έκδοση: 6η Έκδοση Αναθεωρημένη/2012, ISBN: 978-960-531-281-7, Διαθέτης (Εκδότης): ΔΙΑΥΛΟΣ Α.Ε. ΕΚΔΟΣΕΙΣ ΒΙΒΛΙΩΝ
- **Συστήματα διαχείρισης βάσεων δεδομένων**, Συγγραφείς: Ramakrishnan Raghu, Gehrke Johannes, **Κωδικός Βιβλίου στον Εύδοξο: 133031084**, Έκδοση: 3η Έκδοση/2024 (βελτιωμένη), ISBN: 978-618-221-098-7, Διαθέτης (Εκδότης): ΕΚΔΟΣΕΙΣ Α. ΤΖΙΟΛΑ & ΥΙΟΙ Α.Ε.
- **Συστήματα Βάσεων Δεδομένων**, Συγγραφείς: Abraham Silberschatz, Henry F. Korth, S. Sudarshan, **Κωδικός Βιβλίου στον Εύδοξο: 102070677**, Έκδοση: 7η έκδ./2021, ISBN: 978-960-512-743-5, Διαθέτης (Εκδότης): Χ. ΓΚΙΟΥΡΔΑ & ΣΙΑ ΕΕ

Η συνταγή της επιτυχίας ...

- Βαθμολογία :
 - Διαγώνισμα: 50% [βαθμός ≥ 5]
 - Εργασία: 30% [βαθμός ≥ 5 , **υποχρεωτική**]
 - Πιθανό bonus ως 10% σε καλές εργασίες
 - Ασκήσεις: 20% [προαιρετικές]
- Διαγώνισμα: ανοιχτό το βιβλίο του μαθήματος και μόνο αυτό

Η συνταγή της επιτυχίας ...

- Κατοχύρωση: τίποτε δεν κατοχυρώνεται για επόμενη χρονιά!
- Ο σκοπός του μαθήματος είναι να συνδυάσετε
 - αφομοίωση θεμελιωδών αρχών +
 - εξάσκηση των προγραμματιστικών ικανοτήτων σας +
 - κριτική σκέψημε τρόπο συγκροτημένο
- Για το λόγο αυτό:
 - η προγραμματιστική εργασία **θα τελειώσει νωρίς** (ώστε να μην επιβαρύνει την εξεταστική),
 - το διαγώνισμα θα απαιτήσει σίγουρα κατανόηση των **θεμελιωδών εννοιών + κριτική σκέψη** εκ μέρους

Προγραμματιστική Εργασία

- Η προγραμματιστική άσκηση για το μάθημα είναι **υποχρεωτική**
- Αφορά τη σχεδίαση, υλοποίηση και ρύθμιση ενός **ολοκληρωμένου πληροφοριακού συστήματος**
 - προδιαγραφές και ανάλυση
 - κατασκευή βάσης δεδομένων
 - κατασκευή διαπροσωπείας
 - ρύθμιση λειτουργίας
- **Είναι υποχρεωτικό η εργασία να ολοκληρωθεί πλήρως!!**

Η εργασία σε ομάδες από 1 ως και 3 ατόμων (ναι, μπορείτε και ατομικά)

Ασκήσεις εξάσκησης

- Θα δοθούν **3 σειρές ασκήσεων** (ανάλογα με το χρόνο) για να σας βοηθήσουν με τις έννοιες του μαθήματος.
 - *Για να βοηθήσουν, θα πρέπει να τις κάνετε και όχι να τις αντιγράψετε...*
 - *Επιμένω: αν τις κάνετε, και διαβάσετε σωστά για το σκοπό αυτό μέσα στο εξάμηνο, θα έχετε καλύψει σημαντικό τμήμα της ύλης του διαγωνίσματος!*

Θεματολόγιο

- Τι θα μας απασχολήσει στο μάθημα
- Διαδικαστικά του μαθήματος
- Η συνταγή της επιτυχίας

Προαπαιτούμενη γνώση

- Δεν μπορείτε να καταλάβετε το μάθημα αν **δεν έχετε πάρει/ασχοληθεί/...** το «**Βάσεις Ι**»
- Αν **δεν έχετε περάσει** το «**Βάσεις Ι**», αλλά το **έχετε διαβάσει/παρακολουθήσει/...**, μπορείτε να περάσετε το μάθημα
- Θα χρειαστεί να αναπτύξετε **κώδικα σε ένα Web ή OO περιβάλλον ανάπτυξης**, ο οποίος να περιέχει **SQL** ερωτήσεις
- *Μερικές βασικές γνώσεις είναι θεμελιώδεις ...*

Απαιτήσεις

- Όπως ξέρετε, ισχύει το αξίωμα «το εξάμηνο είναι για εσάς, η εξεταστική για τον διδάσκοντα»
 - Στη διάρκεια του εξαμήνου, υπάρχει ως συνήθως η απαίτηση **να διαβάζετε** την ύλη όπως προχωρούμε, και, να κάνετε τις **ασκήσεις εξάσκησης** (όχι το προηγούμενο βράδυ από την προθεσμία)
 - Η παρακολούθηση του μαθήματος βοηθά, αλλά δεν αντικαθιστά το παραπάνω 😊
- Υποχρεωτική απαίτηση για την επιτυχή παρακολούθηση του μαθήματος είναι η **ολοκληρωμένη** διεκπεραίωση της προγραμματιστικής εργασίας

Φόρτος

- Από πλευράς ύλης και φόρτου ασκήσεων εξάσκησης ο φόρτος δεν είναι ιδιαίτερα μεγάλος
- Το project έχει ένα σημαντικό φόρτο, πολλώ δε μάλλον, αν στο πρόγραμμά σας υπάρχουν και άλλα μαθήματα με υψηλό φόρτο
 - Για το λόγο αυτό, το project κοινοποιείται νωρίς, έχει στάδια και εκπονείται σε ομάδες

Οφέλη

- Η κατανόηση των θεμελιωδών αρχών της λειτουργίας των DBMSs ολοκληρώνει σε μεγάλο βαθμό την εκπαίδευσή σας στο πεδίο των Βάσεων Δεδομένων και επί της αρχής, αλλά και πρακτικά:
 - Οι θεμελιώδεις αρχές της λειτουργίας των DBMSs είναι λίγο πολύ σταθερές τα τελευταία 30 χρόνια (αν και είναι ορατή η πιθανότητα αλλαγών).
 - Η πρακτική σημασία της κατανόησης της εσωτερικής λειτουργίας των DBMSs είναι μεγάλη και για τους application developers και για τους administrators
- Η πρακτική εξάσκηση σε θέματα προγραμματισμού γύρω από Βάσεις Δεδομένων είναι
 - Μια επιπλέον ευκαιρία για πρακτική εμπειρία προγραμματισμού (και μάλιστα από τις τελευταίες στη διάρκεια των προπτυχιακών σας σπουδών)
 - Μια ευκαιρία να δείτε συνολικά τη σχεδίαση, τον προγραμματισμό εφαρμογών και τη διαχείριση των δεδομένων σε μια Βάση Δεδομένων

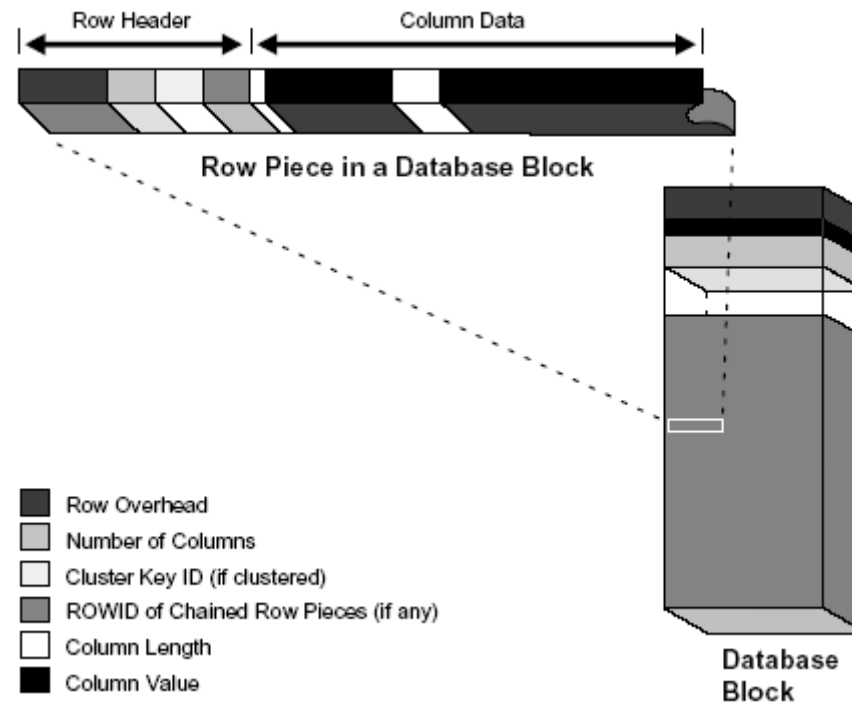
Βοηθητικές διαφάνειες

- Επιλεγμένες διαφάνειες από τις διαλέξεις...

Βάσεις Δεδομένων και ΣΔΒΔ

- Μια **Βάση Δεδομένων** είναι μια οργανωμένη συλλογή θεματικά συναφών δεδομένων (συνήθως σε μεγάλη κλίμακα)
- Ένα **Σύστημα Διαχείρισης Βάσεων Δεδομένων (DataBase Management System – DBMS)** είναι ένα πακέτο λογισμικού που σκοπό έχει τη διαχείριση βάσεων δεδομένων

Εσωτερική Οργάνωση των Δεδομένων σε μια εγγραφή



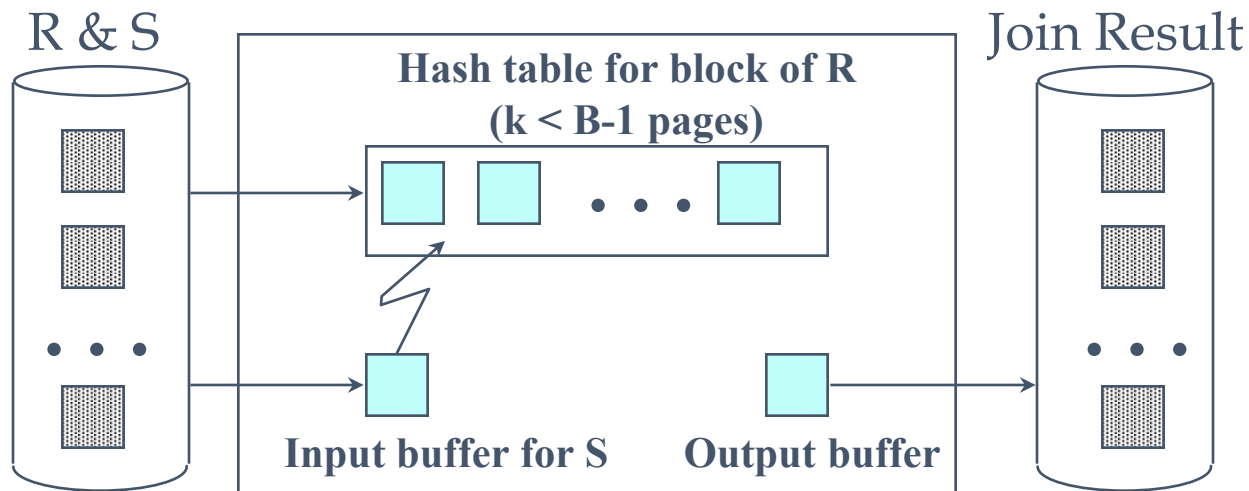
Block Nested Loops Join

L0: Διάβασε ένα block της R (R-block)

Για κάθε σελίδα της S (S-page)

Αν μια εγγραφή r στο R-block και μια εγγραφή s στην
S-page κάνουν match, output $\langle r, s \rangle$

Αν η R δεν εξαντλήθηκε, Goto L0 //διάβασε το επόμενο R-block, scan S, κ.ο.κ.



Συναλλαγή

- **Συναλλαγή** είναι
 - μια σειρά από ενέργειες, οι οποίες
 - διαβάζουν ή γράφουν
 - αντικείμενα της βάσης
- στα αγγλικά “transaction”
- σειρά: διατεταγμένο σύνολο, **λίστα**

Για προχωρημένους: στην θεωρία [της πληροφορικής γενικά, και των ΒΔ ειδικά], οι λεξούλες τύπου «σειρά» έχουν σημασία ... (π.χ., σειρά $\neq$ σύνολο)

Παράδειγμα συναλλαγής

T_0 : μεταφορά 50€ από
το λογαριασμό A στο
λογαριασμό B

```
read (A) ;  
A := A - 50 ;  
write (A) ;  
read (B) ;  
B := B + 50 ;  
write (B) .
```

Προβληματισμός

- Δύο είναι τα βασικά προβλήματα με τις συναλλαγές:
 - Τι θα γίνει αν κατά τη διάρκεια της εκτέλεσης, πέσει το σύστημα?
 - Τι θα γίνει αν δύο συναλλαγές επιχειρούν να μεταβάλλουν το ίδιο αντικείμενο ταυτόχρονα;

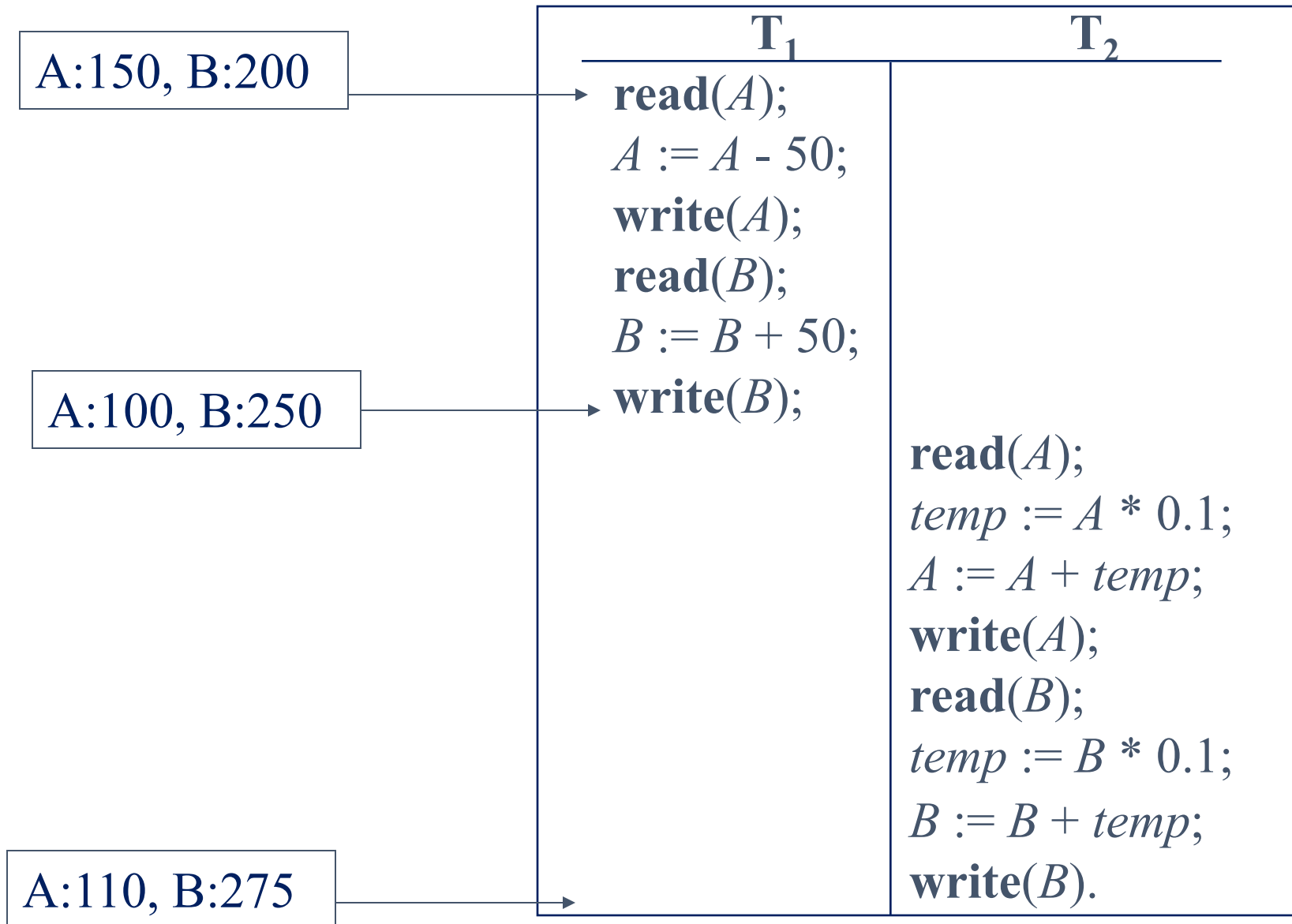
Ορολογία: Στο εξής το σύστημα δεν «πέφτει», αλλά «αποτυγχάνει» 😊

Ταυτοχρονισμός

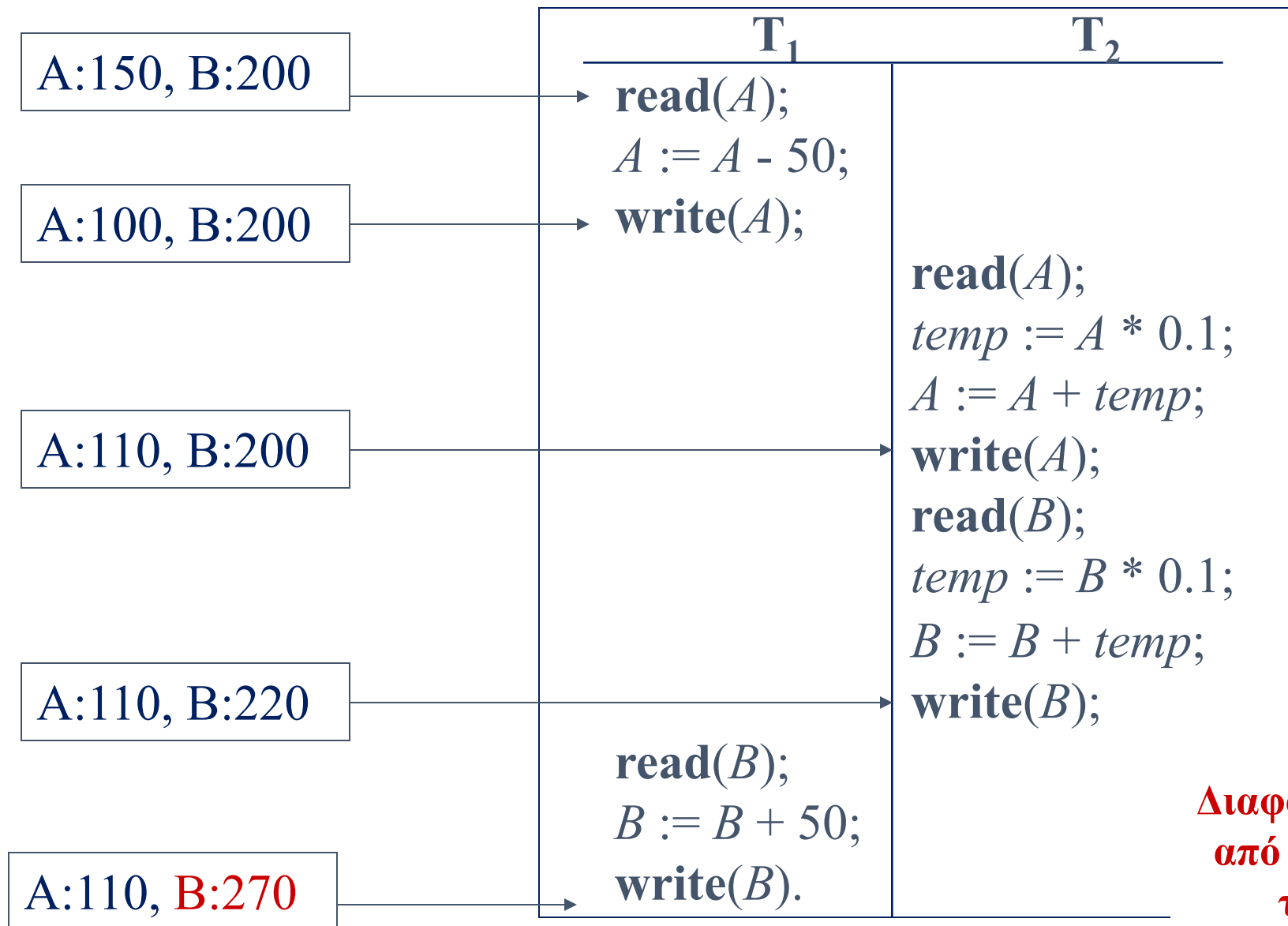
Έστω δύο συναλλαγές που τρέχουν σειριακά:

- T1: μεταφέρει 50 € από τον A στον B
- T2: κάνει αύξηση στον A και το B κατά 10%

Ταυτοχρονισμός σειριακά

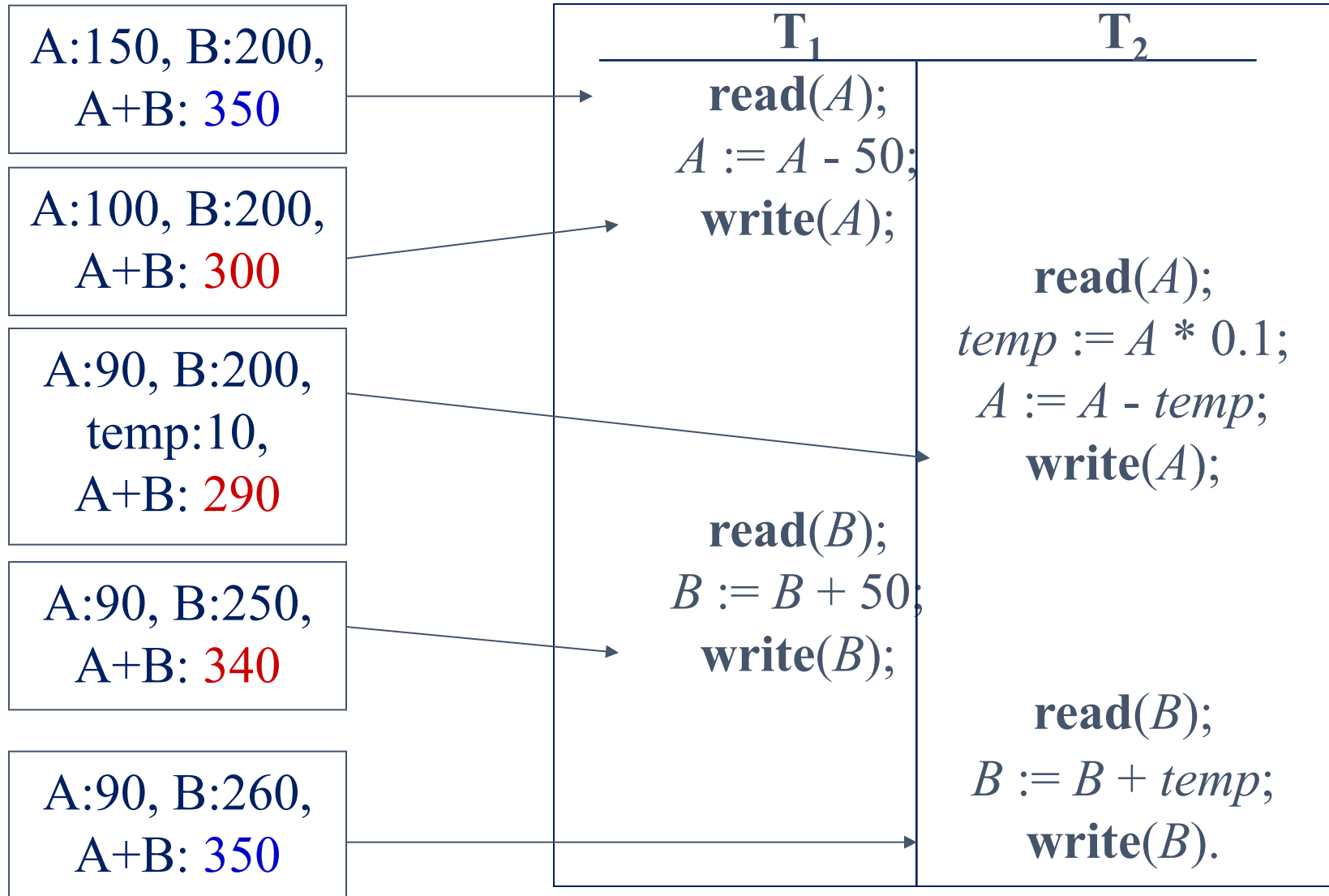


Ταυτοχρονισμός με πολυπλεξία συναλλαγών



Διαφορετικό
από το 275
του
σειριακού!!!

Μη Σειριακό Χρονοπρόγραμμα

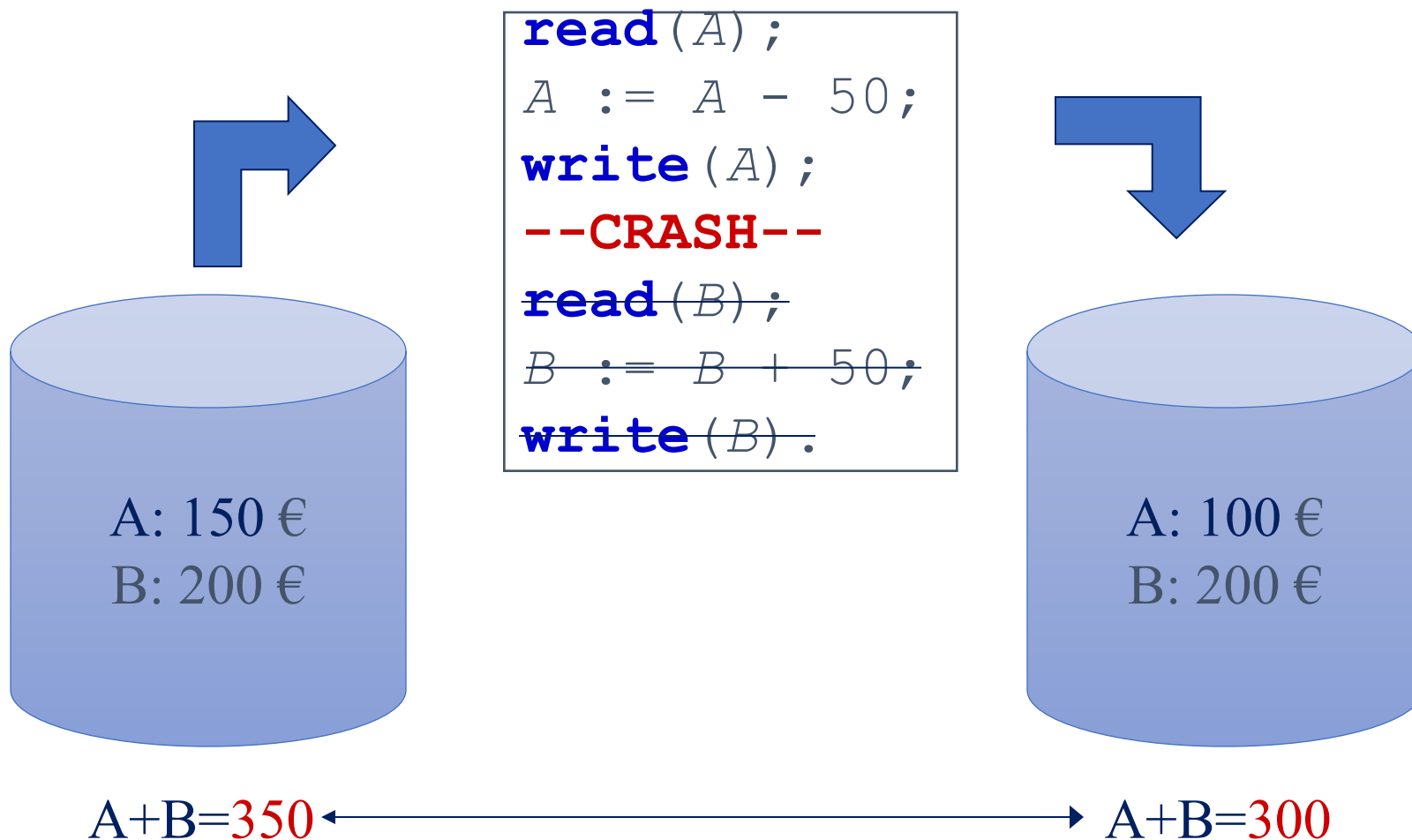


Διαχειριστής κλειδαριών

- Το τμήμα του DBMS που διαχειρίζεται τα κλειδώματα
- Αν το DBMS υποστηρίζει 2PL → **2PL scheduler**
- **Πίνακας κλειδαριών:**
 - ObjectID
 - TransactionID, LockType
 - Queue με συναλλαγές εν αναμονή

	x	y	z
T1		S	
T2		S	
T3		W	

Αποτυχία σε μεταφορά από ένα τραπεζικό λογ/μο σε άλλο: πώς θα γίνει να μη χαθούν λεφτά?



Αστοχίες του συστήματος

- Κακό πρόγραμμα (με λάθη, δηλ.)
- Αστοχία συναλλαγής (αδιέξοδο, abort από τον χρήστη, κλπ)
- Αστοχία του συστήματος (πτώση ρεύματος, αδιέξοδο λειτουργικού συστήματος)
- Αστοχία υλικού (καταστροφή σκληρού δίσκου)

Failure: αστοχία ή αποτυχία

Write Ahead Log

Προτού γράψεις οτιδήποτε στη ΒΔ, καταχώρησε την αντίστοιχη εγγραφή στο log

- Για να γράψεις μια updated σελίδα από το buffer πίσω στο δίσκο, πρέπει στο log (στο δίσκο) να έχουν περαστεί οι παλιές τιμές για τα records της
- Για να κάνεις commit μια δοσοληψία πρέπει στο log (στο δίσκο) να έχουν γραφτεί όλες οι σχετικές log records

Σε περίπτωση αποτυχίας ...

- Σκοπός είναι να διατηρήσουμε τη συνέπεια του συστήματος.
- Πρέπει να επαναλάβουμε (**REDO**) όλες τις δοσοληψίες που έκαναν commit
- Πρέπει να αναιρέσουμε (**UNDO**) όσες παρενέργειες επέφεραν οι δοσοληψίες που δεν πρόλαβαν να κάνουν commit

Log (Ιστορικό)

- **Log** (Ιστορικό/Ιχνος/Ημερολόγιο): ένα αρχείο στο σκληρό δίσκο που καταγράφει όλη την ιστορία των ενεργειών που εκτελέστηκαν από το DBMS
- **Ενέργειες:**
 - BOT/EOT (Begin/End Of Transaction)
 - INS/UPD/DEL (ήτοι, write) ένα record
 - COMMIT/ABORT μια δοσοληψία
 - UNDO/REDO μια ενέργεια εγγραφής (write)

Checkpoints – Σημεία ελέγχου

- Περιοδικά, το σύστημα κάνει τις εξής ενέργειες:
 - Σταματά κάθε άλλη ενέργεια
 - Καταγράφει το σύνολο των ενεργών δοσοληψιών
 - Γράφει (**flush**) όλους τους buffers με log records, στο δίσκο
 - Γράφει (**flush**) όλους τους buffers με records της ΒΔ, στο δίσκο
 - Γράφει στο log μια εγγραφή **CHK** (checkpoint)

Ανάνηψη αν έχουμε WAL

- Έστω ότι το σύστημα αποτυγχάνει και πρέπει να το επαναφέρουμε (ήτοι, να επαναφέρουμε τη ΒΔ σε συνεπή μορφή).
- Η διαδικασία αυτή ονομάζεται **ανάνηψη (recovery)** ή **ανάκαμψη** ή **επαναφορά**
- Αν έχουμε χρησιμοποιήσει **WAL** κατά την κανονική λειτουργία του συστήματος, η ανάνηψη έχει **3 φάσεις**:
 1. Ανάλυση
 2. UNDO
 3. REDO

Ρύθμιση της σχεδίασης

Emp

<u>Eno</u>	Ename	Dno	Sal	Hobby
103	Demis	1	10K	Music
202	Vassilis	2	8K	Food
208	Ilia	2	10K	Coffee
123	Theo	2	5K	sports

Το σχήμα των πινάκων είναι ίδιο

Η ένωση των πλειάδων δίνει τις πλειάδες του αρχικού πίνακα

Emp_Rich

Sal >= 10K

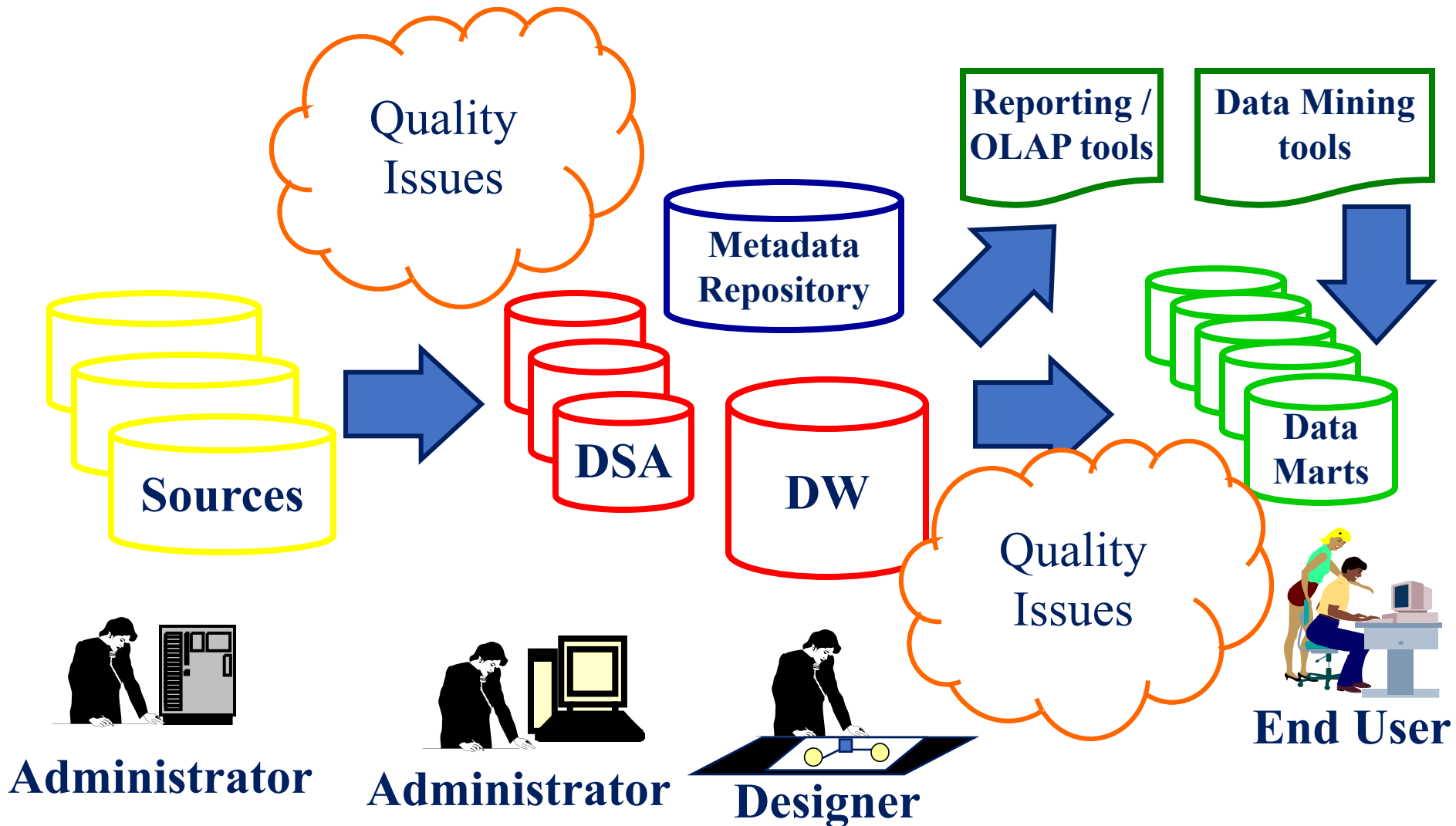
<u>Eno</u>	Ename	Dno	Sal	Hobby
103	Demis	1	10K	Music
208	Ilia	2	10K	Coffee

Emp Poor

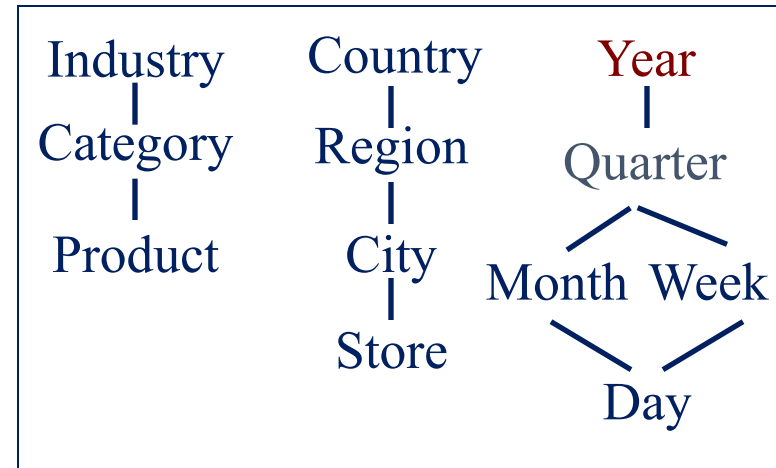
Sal < 10K

<u>Eno</u>	Ename	Dno	Sal	Hobby
202	Vassilis	2	8K	Food
123	Theo	2	5K	sports

Αποθήκες Δεδομένων (Data Warehouses)

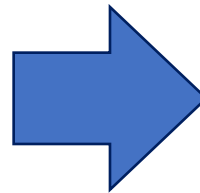


Roll up



Sales volume			
	Products	Store1	Store2
Q1	Electronics	\$5,2	\$5,6
	Toys	\$1,9	\$1,4
	Clothing	\$2,3	\$2,6
	Cosmetics	\$1,1	\$1,1
Q2	Electronics	\$8,9	\$7,2
	Toys	\$0,75	\$0,4
	Clothing	\$4,6	\$4,6
	Cosmetics	\$1,5	\$0,5

Χρόνος: Επίπεδο Quarter

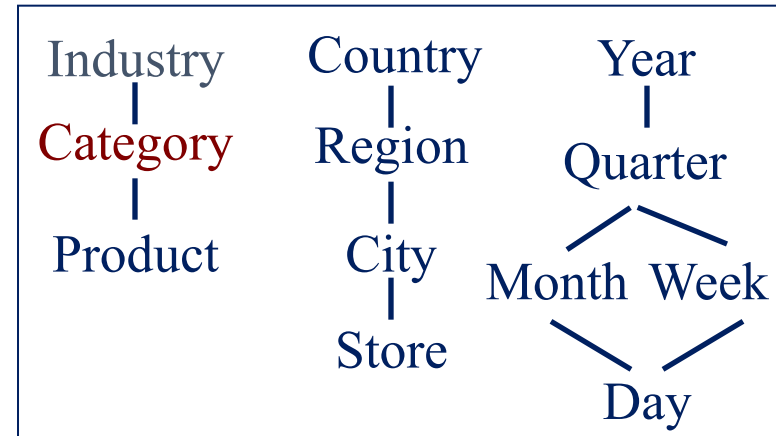


Sales volume			
	Products	Store1	Store2
Year 1996	Electronics	\$14,1	\$12,8
	Toys	\$2,65	\$1,8
	Clothing	\$6,9	\$7,2
	Cosmetics	\$2,6	\$1,6

Χρόνος: Επίπεδο Year

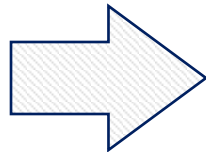
SUM(Sales volumes)

Drill down



Sales volume			
Products		Store1	Store2
Q1	Electronics	\$5,2	\$5,6
	Toys	\$1,9	\$1,4
	Clothing	\$2,3	\$2,6
	Cosmetics	\$1,1	\$1,1
Q2	Electronics	\$8,9	\$7,2
	Toys	\$0,75	\$0,4
	Clothing	\$4,6	\$4,6
	Cosmetics	\$1,5	\$0,5

Item: Επίπεδο Industry



Sales volume			
Electronics		Store1	Store2
Q1	VCR	\$1,4	\$1,4
	Camcorder	\$0,6	\$0,6
	TV	\$2,0	\$2,4
	CD player	\$1,2	\$1,2
Q2	VCR	\$2,4	\$2,4
	Camcorder	\$3,3	\$1,3
	TV	\$2,2	\$2,5
	CD player	\$1,0	\$1,0

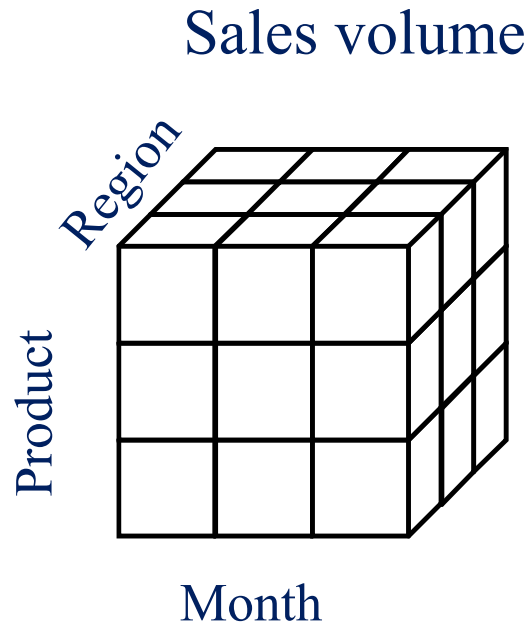
Item: Επίπεδο Category

Top N Queries

```
SELECT P.pid, P.pname, S.sales  
FROM Sales S, Products P  
WHERE S.pid=P.pid AND S.locid=1 AND S.timeid=3  
ORDER BY S.sales DESC  
OPTIMIZE FOR 10 ROWS
```

- Το **OPTIMIZE FOR** clause καλύπτει την απαίτηση προγραμματιστικά. Δεν είναι μέρος της SQL:1999 αλλά υπάρχει π.χ., στην DB2!

Πολυδιάστατο μοντέλο δεδομένων



Διαστάσεις: Product, Region, Date

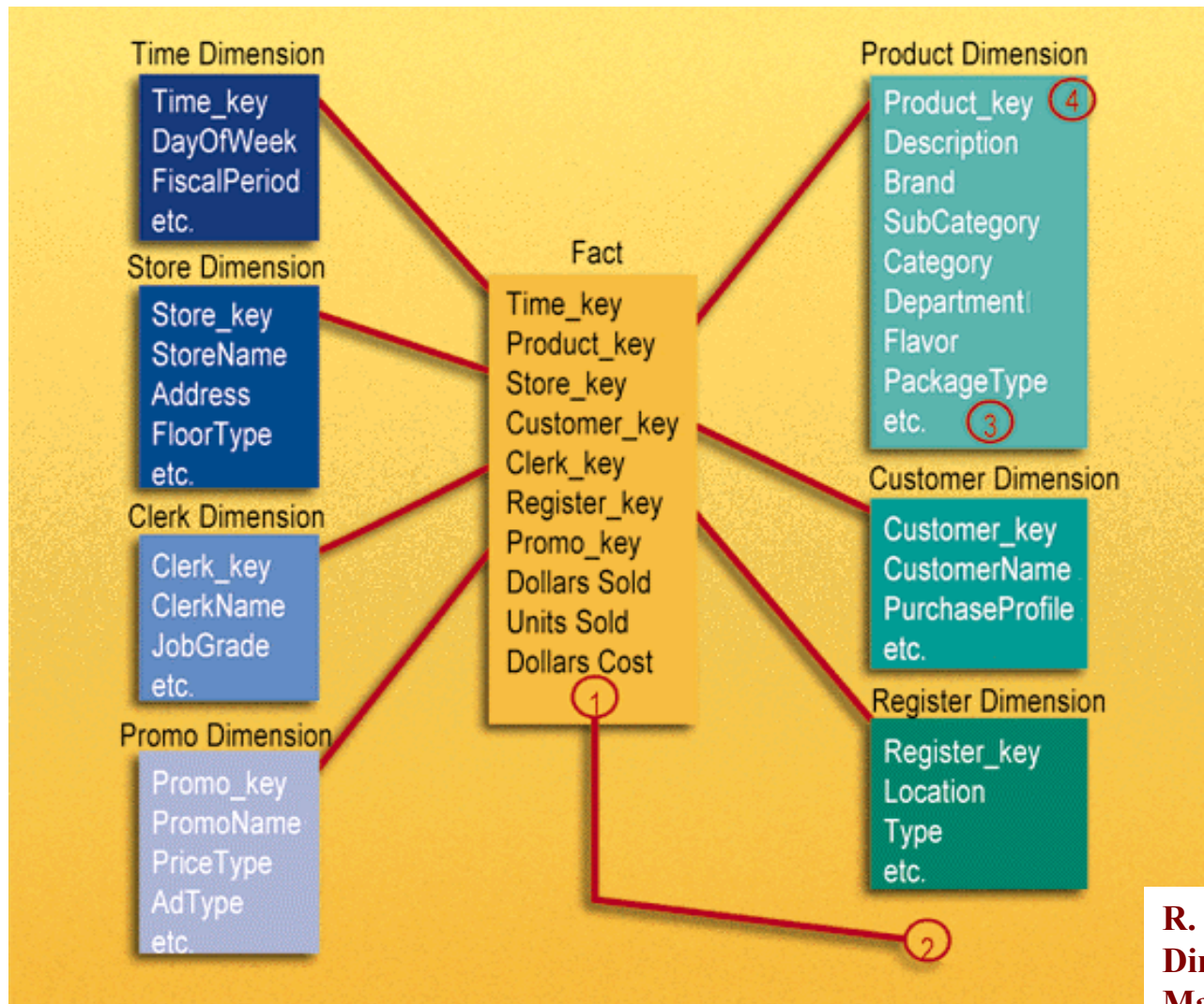
Ιεραρχίες διαστάσεων:

Industry
|
Category
|
Product

Country
|
Region
|
City
|
Store

Year
|
Quarter
/ \
Month Week
\< /
Day

Παράδειγμα σχήματος στο Μοντέλο Διαστάσεων



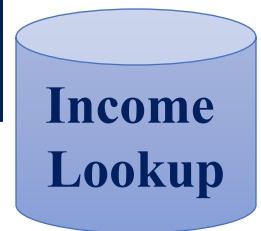
**R. Kimball, A
Dimensional Modeling
Manifesto, DBMS
Magazine, Aug. 1997**

Αποκανονικοποίηση - Denormalization

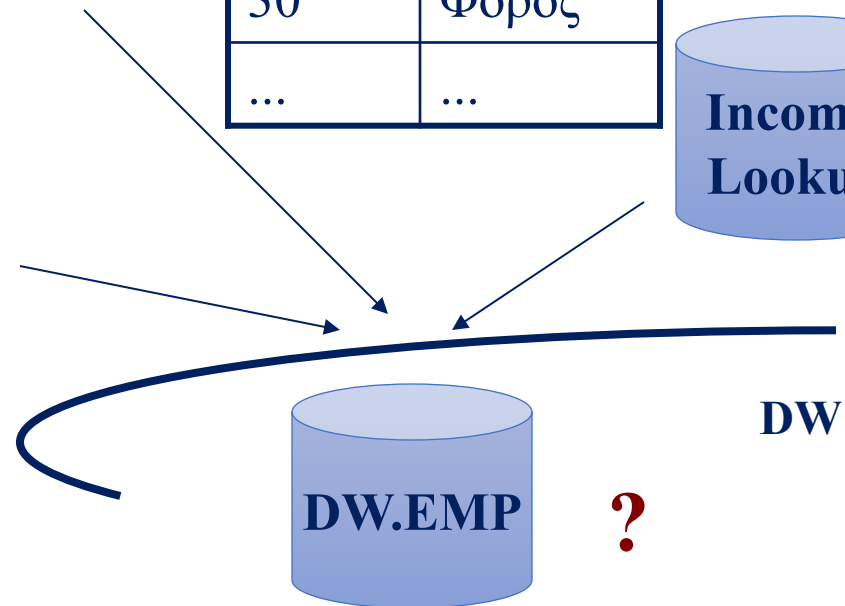
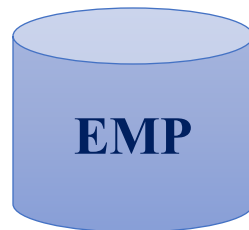
<u>EMP ID</u>	<u>IL ID</u>	Amount
110	10	1500
110	30	300



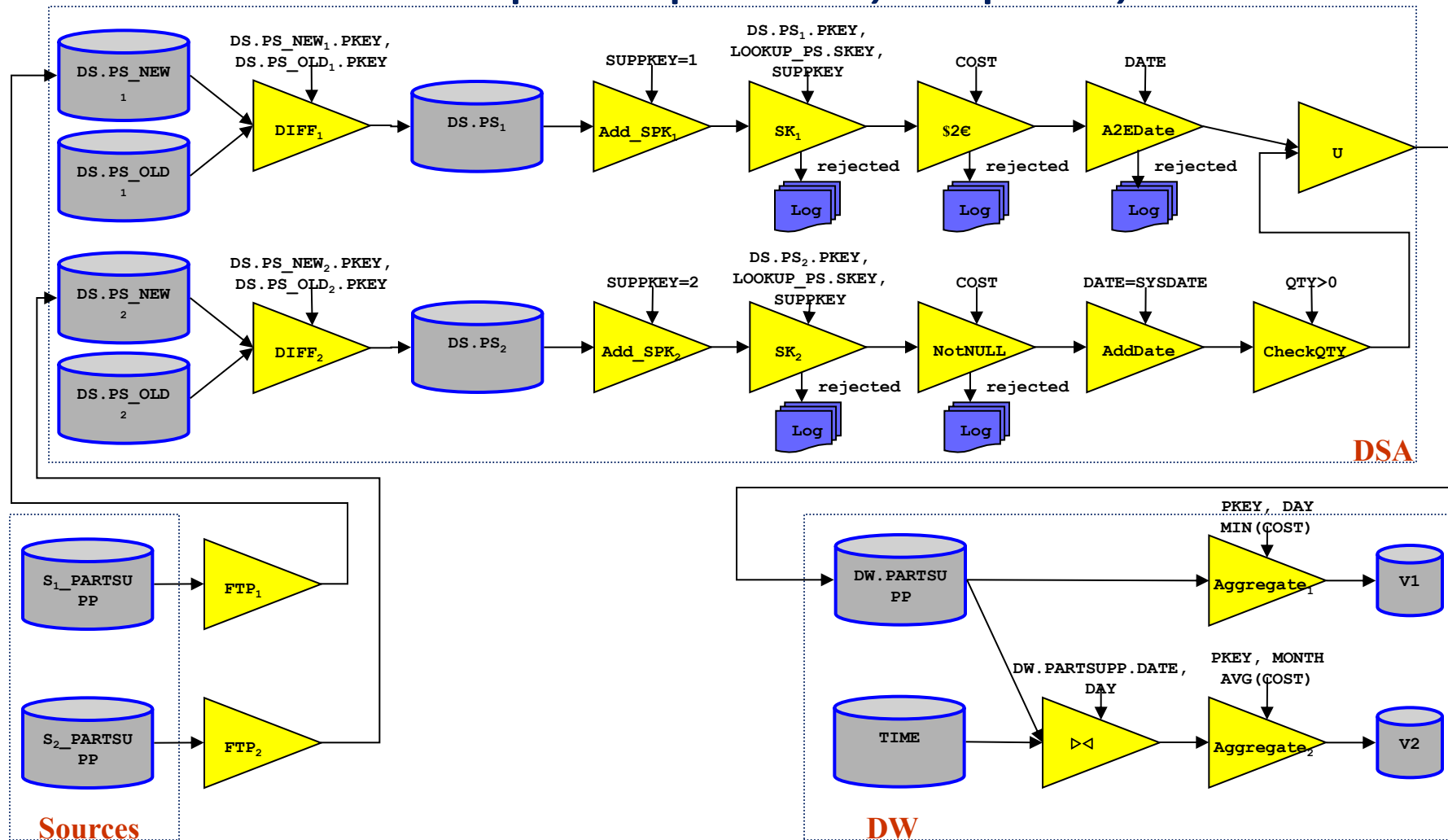
<u>IL ID</u>	Descr
10	Μισθός
20	Επίδομα Τέκνων
30	Φόρος
...	...



<u>EMP ID</u>	Name	Age
110	Kostas	30
120	Mitsos	48
130	Roula	29



DW ≠ Αποθηκευμένες όψεις!



Εννοιολογική εναρμόνιση

**Source 1:
Personnel
(Cobol)**

<u>EMP ID</u>	Name	DoB	Salary	Total Income	DeptID
110	Kostas	1/1/72	1500	1200	132
...	...				...

**Source 2: Accounting
(DB2)**

<u>EMP ID</u>	<u>IL ID</u>	Amount
110	10	1500
110	30	300

**EMP
INCOME**

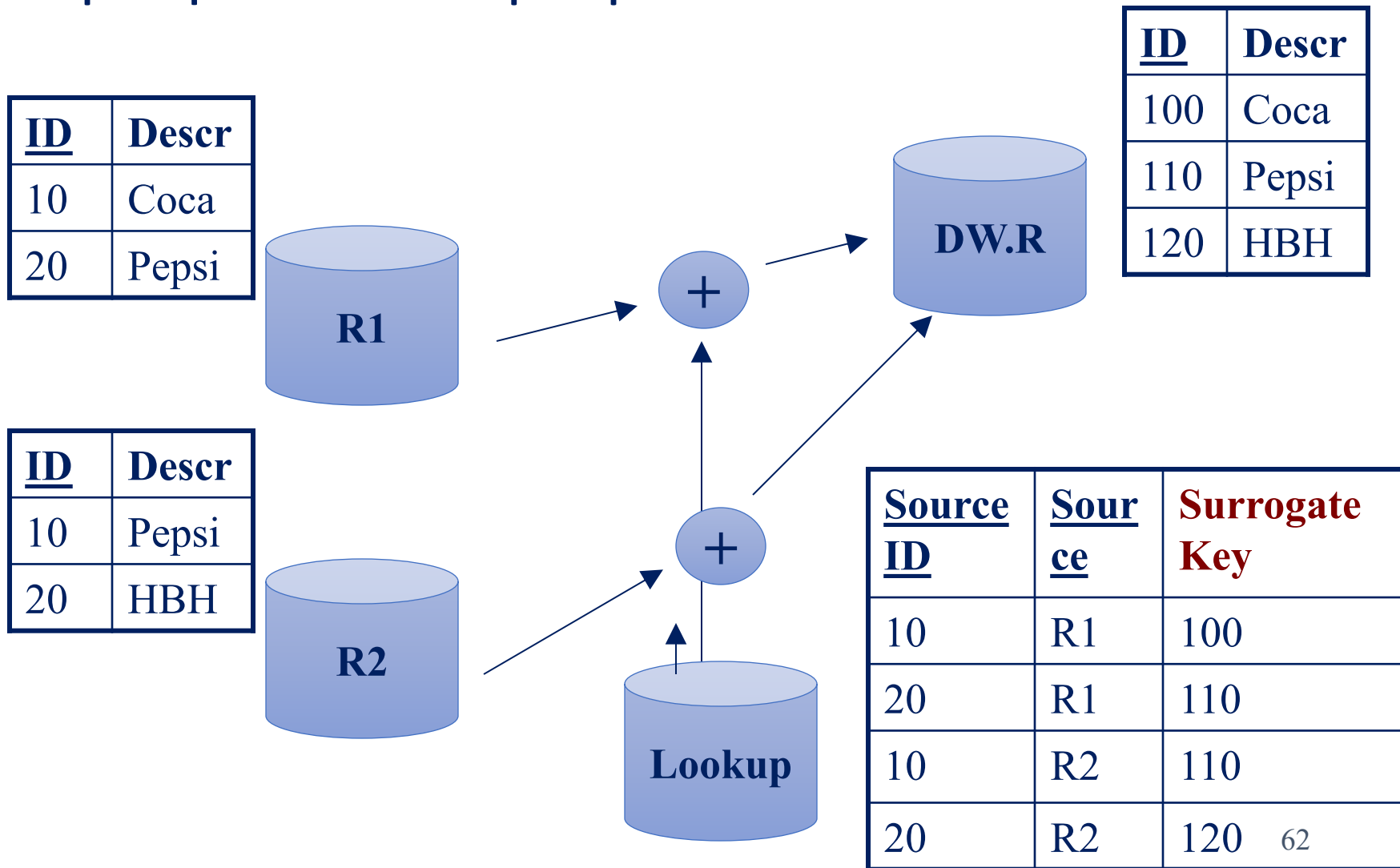
<u>EMP ID</u>	Name	Age
110	Kostas	30
120	Mitsos	48
130	Roula	29

EMP

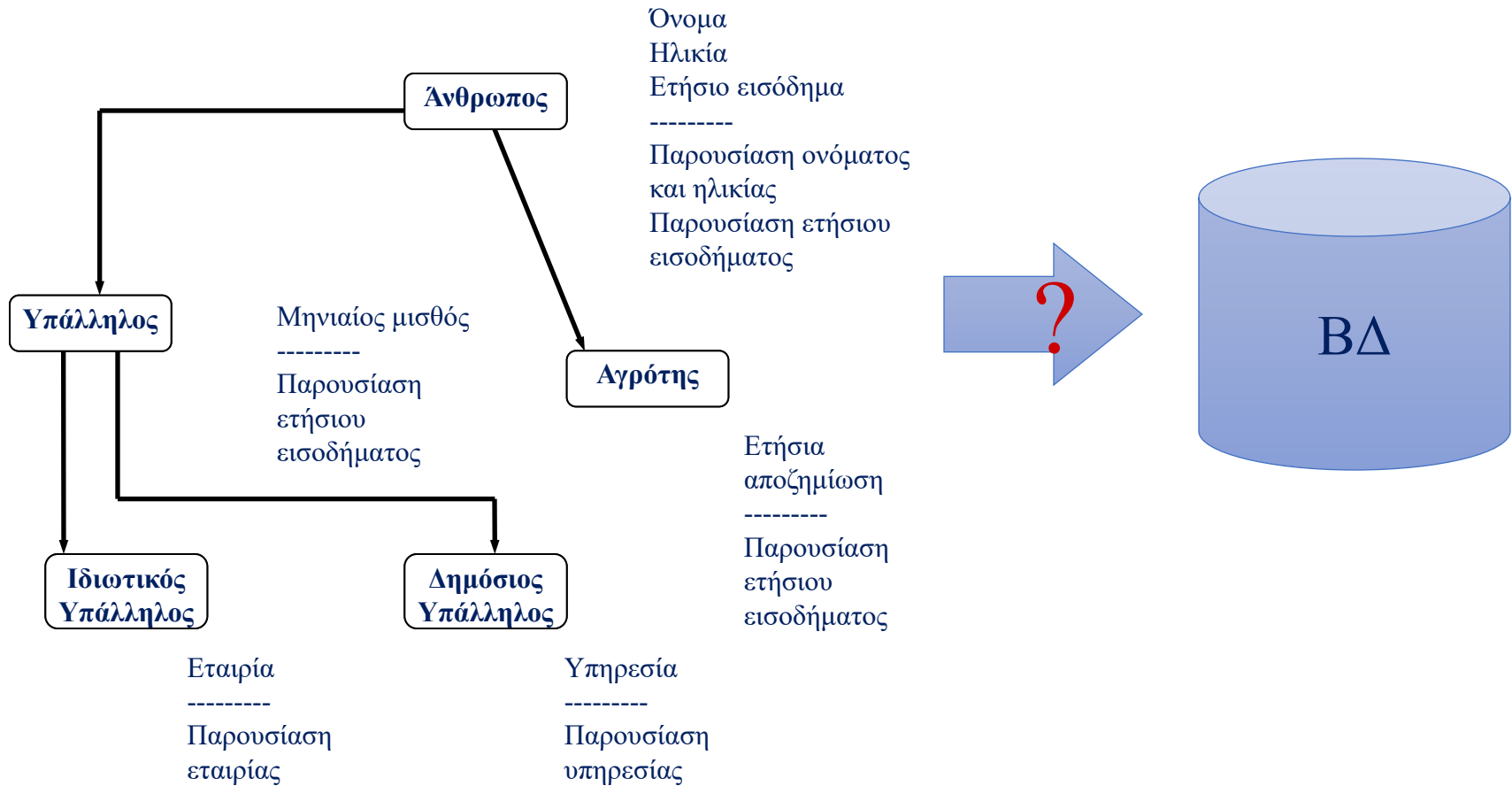
<u>IL ID</u>	Descr
10	Μισθός
20	Επίδομα Τέκνων
30	Φόρος
...	...

**Income
Lookup**

Καθαρισμός των δεδομένων – Ομογενοποίηση κλειδιών



Αντικειμενοστρεφείς ΒΔ



Τεστ σχεδίασης

- Ανάλογα με το πόσο χρόνο θα έχουμε, το μάθημα θα δώσει και μια παράλληλη έμφαση σε **θέματα σχεδίασης**.
- Θα γίνουν χωριστά εργαστήρια στα οποία θα συζητάμε εναλλακτικούς τρόπους σχεδίασης μιας βάσης δεδομένων
 - *ΔΕΝ θα απαιτηθεί ιδιαίτερο διάβασμα για τα θέματα σχεδίασης, αλλά παρακολούθηση των εργαστηρίων!*
- Παλαιά η συμμετοχή ήταν υποχρεωτική στο εργαστήριο, καθώς στο τέλος υπήρχε test. Για φέτος, δεν υπάρχει τέτοια πρόβλεψη.