



# ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΤΜΗΥΠ

## ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ Ι

---

*Β. Μεγαλοικονόμου*

*Δ. Χριστοδουλάκης*

## **Εισαγωγή στις Βάσεις Δεδομένων**



# Εισαγωγή

---

- Γιατί συστήματα Βάσεων Δεδομένων
- Όψη Δεδομένων
- Μοντέλα Δεδομένων
- Γλώσσα Ορισμού Δεδομένων
- Γλώσσα Χειρισμού Δεδομένων
- Διαχείριση Δοσοληψιών
- Διαχείριση Αποθήκευσης
- Διαχειριστής Βάσης Δεδομένων
- Χρήστες Βάσης Δεδομένων
- Συνολική Δομή Συστήματος

# Συστήματα Διαχείρισης Βάσεων Δεδομένων (ΣΔΒΔ)

- ΣΔΒΔ = μια συλλογή από συσχετιζόμενα δεδομένα (βάση δεδομένων) + ένα σύνολο προγραμμάτων για την πρόσβαση στα δεδομένα
- Το ΣΔΒΔ περιέχει πληροφορία για μια συγκεκριμένη εταιρεία
- Το ΣΔΒΔ παρέχει ένα περιβάλλον που είναι *βολικό* και *εύχρηστο*
- Εφαρμογές Βάσεων Δεδομένων:
  - Τραπεζικός τομέας: όλες οι δοσοληψίες
  - Αερογραμμές: κρατήσεις, δρομολόγια πτήσεων
  - Πανεπιστήμια: εγγραφές, βαθμολογίες
  - Πωλήσεις: πελάτες, προϊόντα, αγορές
  - Κατασκευές: παραγωγή, αποθέματα, παραγγελίες, εφοδιαστική αλυσίδα
  - Ανθρώπινο δυναμικό: αρχεία εργαζομένων, μισθοδοσία, φοροαπαλλαγές
- Οι βάσεις δεδομένων είναι παντού, σε όλες τις πτυχές της ζωής μας

# Γιατί Βάσεις Δεδομένων;

- Παλαιότερα, οι εφαρμογές βάσεων δεδομένων χτιζονταν πάνω σε συστήματα αρχείων
- Μειονεκτήματα στη χρήση συστημάτων αρχείων για την αποθήκευση δεδομένων:
  - Πλεονασμός πληροφορίας και έλλειψη συνοχής
    - Πολλαπλοί τύποι αρχείων, επαναλαμβανόμενη πληροφορία σε διαφορετικά αρχεία
  - Δυσκολία πρόσβασης στα δεδομένα
    - Χρειαζόταν νέο πρόγραμμα για την εκτέλεση κάθε νέας εργασίας
  - Απομόνωση δεδομένων — δεδομένα σε πολλαπλά αρχεία και τύπους
  - Πρόβλημα ακεραιότητας
    - Περιορισμοί ακεραιότητας (πχ. Υπόλοιπο λογαριασμού > 25 Euro) γίνεται μέρος του κώδικα του προγράμματος
    - Δύσκολη η προσθήκη νέων περιορισμών και η αλλαγή υπαρχόντων

# Γιατί Βάσεις Δεδομένων; (συνέχεια)

- ... (άλλα) Μειονεκτήματα χρήσης συστημάτων αρχείων
  - Ατομικότητα ενημερώσεων
    - Οι αποτυχίες ενημέρωσης επέφεραν κατάσταση ασυνέπειας στη βάση με μέρος των ενημερώσεων να έχουν εκτελεστεί
    - Πχ. Μεταφορά χρημάτων από έναν λογαριασμό σε άλλον θα έπρεπε είτε να ολοκληρωθεί εντελώς είτε να ξαναγίνει
  - Ταυτόχρονη πρόσβαση πολλών χρηστών
    - Αναγκαία η ταυτόχρονη πρόσβαση για την αποτελεσματικότητα
    - Μη-ελεγχόμενες ταυτόχρονες προσβάσεις οδηγούν σε έλλειψη συνοχής
      - Πχ. Δύο άτομα βλέπουν το υπόλοιπο και το ενημερώνουν ταυτόχρονα
  - Προβλήματα ασφάλειας – πρόσβαση σε ορισμένο μέρος της πληροφορίας
- Τα συστήματα βάσεων δεδομένων προσφέρουν λύσεις σε όλα τα παραπάνω προβλήματα!!!



# Όψη Δεδομένων: Επίπεδα Αφαιρετικότητας

---

- Φυσικό επίπεδο
- Λογικό επίπεδο
- Επίπεδο όψης

# Όψη Δεδομένων: Επίπεδα Αφαιρετικότητας

- Φυσικό επίπεδο: περιγράφει *πώς* ένα αρχείο (πχ. πελάτης) αποθηκεύεται
- Λογικό επίπεδο: περιγράφει *ποια* δεδομένα αποθηκεύονται στη βάση και τις μεταξύ τους συσχετίσεις

**type** πελάτης = **record**

*όνομα* : string;

*οδός* : string;

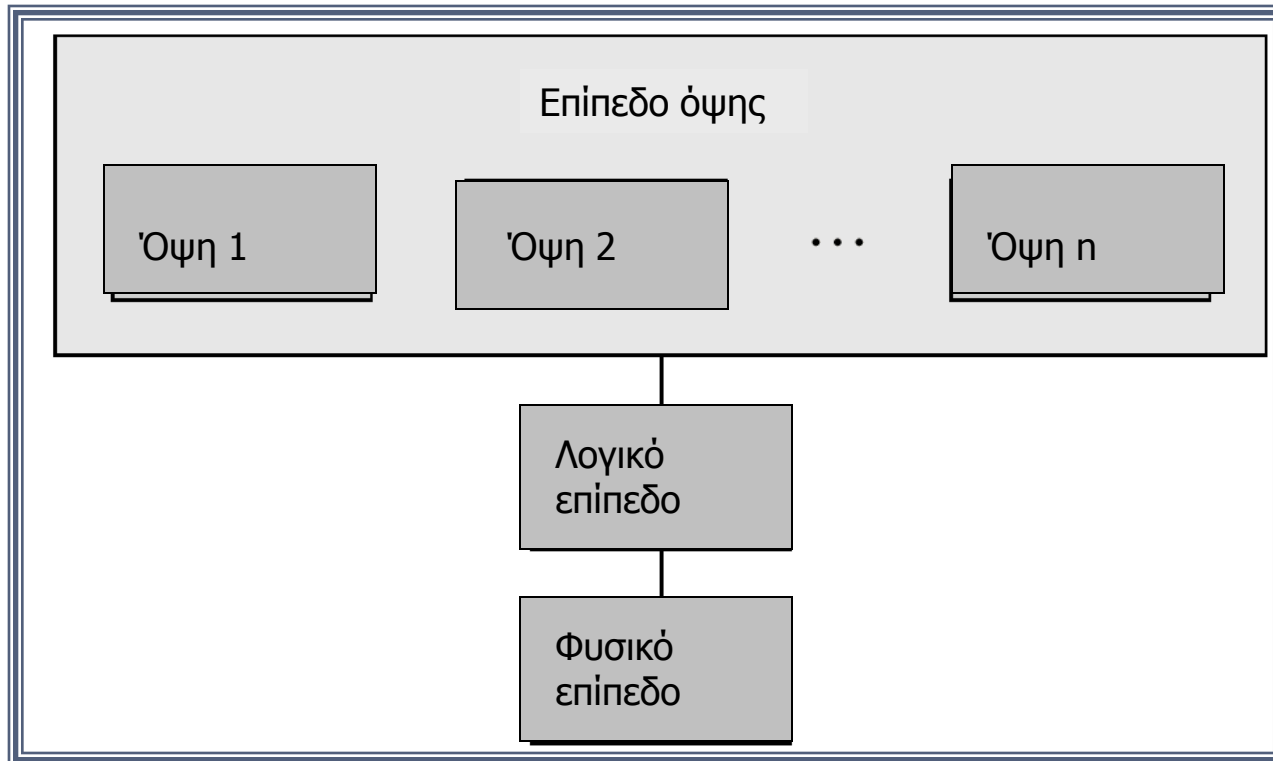
*Πόλη* : integer;

**end;**

- Επίπεδο όψης: προγράμματα εφαρμογών που κρύβουν λεπτομέρειες των τύπων δεδομένων. Οι όψεις μπορούν επίσης να κρύψουν πληροφορία (πχ. μισθός) για λόγους ασφαλείας

# Όψη δεδομένων

Αρχιτεκτονική συστήματος βάσης δεδομένων





# Όψη δεδομένων: Σχήματα και Στιγμιότυπα

- ... παρόμοια με τους τύπους και τις μεταβλητές των γλωσσών προγραμματισμού
- **Σχήμα**– η λογική δομή της βάσης δεδομένων
  - Πχ. Η βάση δεδομένων περιέχει πληροφορία για ένα σύνολο πελατών και λογαριασμών και τις σχέσεις μεταξύ τους
  - Ανάλογη με τον τύπο της πληροφορίας σε μια μεταβλητή ενός προγράμματος
  - **Φυσικό σχήμα**: ο σχεδιασμός της βάσης δεδομένων σε φυσικό επίπεδο
  - **Λογικό σχήμα**: ο σχεδιασμός της βάσης δεδομένων σε λογικό επίπεδο
- **Στιγμιότυπο** –το πραγματικό περιεχόμενο της βάσης σε μια δεδομένη χρονική στιγμή
  - Ανάλογο της τιμής μιας μεταβλητής



# Όψη δεδομένων: Σχήματα και Στιγμιότυπα

---

- **Φυσική Ανεξαρτησία Δεδομένων ;**



# Όψη δεδομένων: Σχήματα και Στιγμιότυπα

---

- **Φυσική Ανεξαρτησία Δεδομένων** – δυνατότητα τροποποίησης του φυσικού σχήματος χωρίς αλλαγή του λογικού σχήματος
  - Οι εφαρμογές εξαρτώνται από το λογικό σχήμα
  - Γενικά, οι διεπαφές μεταξύ των διαφορετικών επιπέδων και των συστατικών θα πρέπει να καθορίζονται έτσι ώστε αλλαγές σε ορισμένα μέρη του συστήματος να μην επηρεάζουν σημαντικά άλλα μέρη



# Μοντέλα Δεδομένων

---

- Μια συλλογή εργαλείων για την περιγραφή
  - Δεδομένων
  - Συσχετίσεων δεδομένων
  - Σημασιολογία δεδομένων
  - Περιορισμών δεδομένων
- Μοντέλο Οντοτήτων-Συσχετίσεων (E-R)
- Σχεσιακό Μοντέλο

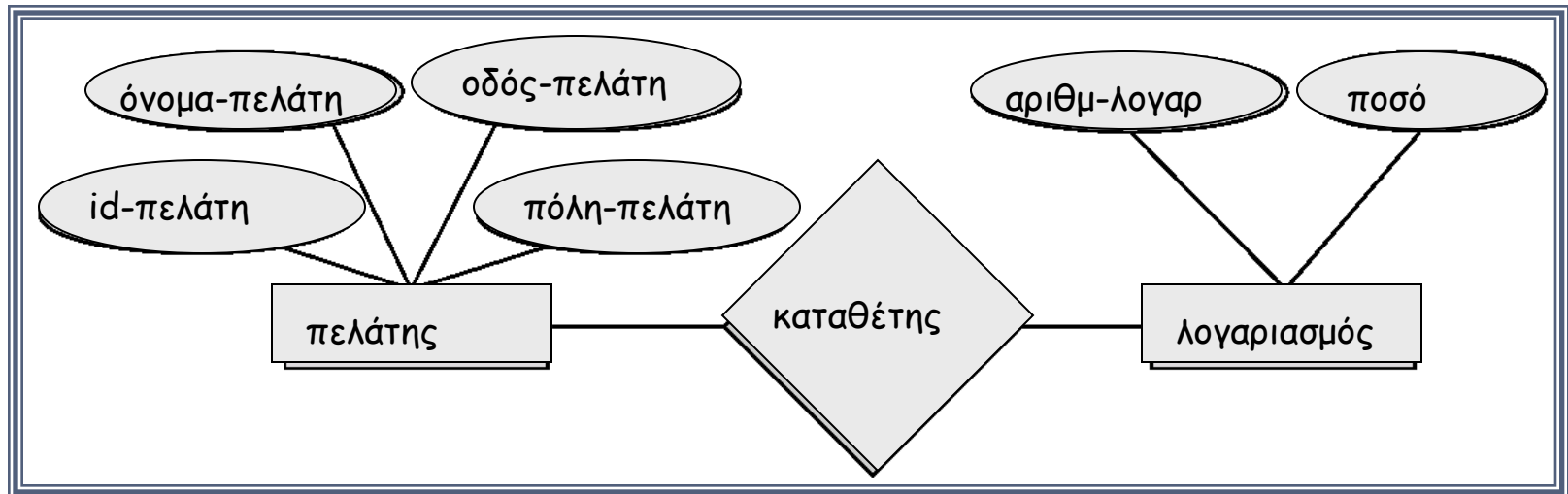


# Μοντέλα Δεδομένων (συνέχεια)

- Άλλα μοντέλα:
  - Αντικειμενοστρεφές μοντέλο (επέκταση του E-R με ενθυλάκωση, μεθόδους και αντικείμενα)
  - Αντικειμενοστρεφές-σχεσιακό μοντέλο
  - Ημι-δομημένα μοντέλα δεδομένων– δεδομένα του ίδιου τύπου μπορεί να έχουν διαφορετικά σύνολα γνωρισμάτων
    - XML
  - Παλαιότερα μοντέλα: δικτυωτό και ιεραρχικό

# Μοντέλο Οντοτήτων-Συσχετίσεων

Παράδειγμα σχήματος σε ένα μοντέλο Οντοτήτων - Συσχετίσεων



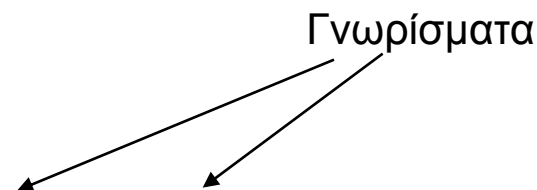
# Μοντέλο Οντοτήτων-Συσχετίσεων (συνέχεια)

- Μοντέλο E-R του πραγματικού κόσμου
  - Οντότητες (αντικείμενα)
    - Πχ. πελάτες, λογαριασμοί, υποκαταστήματα τράπεζας
  - Συσχετίσεις μεταξύ οντοτήτων
    - Πχ. Ο λογαριασμός A-101 διατηρείται από τον πελάτη Φωτίου
    - Ο τύπος συσχέτισης *καταθέτης* συσχετίζει πελάτες και λογαριασμούς
- Ευρέως χρησιμοποιούμενο για το σχεδιασμό ΒΔ
  - Ο σχεδιασμός της ΒΔ στο E-R μοντέλο συνήθως μετατρέπεται σε σχεδιασμό στο σχεσιακό μοντέλο (θα το δούμε στη συνέχεια) το οποίο χρησιμοποιείται για αποθήκευση και επεξεργασία

# Σχεσιακό Μοντέλο

■ Παράδειγμα δεδομένων πίνακα στο σχεσιακό μοντέλο

Γνωρίσματα



<i>Κωδ-πελάτη</i>	<i>Όνομα-πελάτη</i>	<i>Οδός-πελάτη</i>	<i>Πόλη-πελάτη</i>	<i>Αριθμός-λογαριασμού</i>
192-83-7465	Ιωάννου	Αιόλου	Πάτρα	A-101
019-28-3746	Σταύρου	Νότου	Ρόδος	A-215
192-83-7465	Ιωάννου	Αιόλου	Πάτρα	A-201
321-12-3123	Τζίμα	Μηλιάς	Χανιά	A-217
019-28-3746	Σταύρου	Νότου	Ρόδος	A-201



# Παράδειγμα Σχεσιακής Βάσης

Id-πελάτη	Όνομα-πελάτη	Διεύθυνση-πελάτη	Πόλη-πελάτη
192-83-7465	Ιωάννου	Αιόλου 12	Πάτρα
019-28-3746	Σταύρου	Νότου 4	Ρόδος
677-89-9011	Χρήστου	Μάχης 3	Χίος
182-73-6091	Τσάμης	Πατρόκλου 123	Σύρος
321-12-9999	Τζίμα	Μηλιάς 100	Χανιά
336-66-9999	Λυσίου	Παράσχου 175	Πάργα
019-28-3746	Σταύρου	Νότου 72	Ραψάνη

πελάτης

λογαριασμός

αριθμός-λογαριασμού	ποσό
A-101	500
A-215	700
A-102	400
A-305	350
A-201	900
A-217	750
A-222	700

καταθέτης

Id-πελάτη	αριθμός-λογαριασμού
192-83-7465	A-101
019-28-3746	A-215
677-89-9011	A-102
182-73-6091	A-305
321-12-9999	A-201
336-66-9999	A-217
019-28-3746	A-222

# Γλώσσες Βάσεων Δεδομένων: Γλώσσα Ορισμού Δεδομένων (ΓΟΔ)

- Καθορίζει το συμβολισμό για τον ορισμό του σχήματος της ΒΔ
  - Πχ.  
**create table** *λογαριασμός* (  
                  *αριθμός-λογαριασμού*   **char**(10),  
                  *υπόλοιπο*                   **integer**)
- Η μεταγλωττιστής της ΓΟΔ δημιουργεί ένα σύνολο από πίνακες που αποθηκεύονται σε ένα *λεξικό δεδομένων*
- Το λεξικό δεδομένων περιέχει **μετα-δεδομένα** (;)

# Γλώσσες Βάσεων Δεδομένων:

## Γλώσσα Ορισμού Δεδομένων (ΓΟΔ)

- Καθορίζει το συμβολισμό για τον ορισμό του σχήματος της ΒΔ
  - Πχ. **create table** *λογαριασμός*  
(*αριθμός-λογαριασμού* **char**(10),  
*υπόλοιπο* **integer**)
- Ο μεταγλωττιστής της ΓΟΔ δημιουργεί ένα σύνολο από πίνακες που αποθηκεύονται σε ένα *λεξικό δεδομένων*
- Το λεξικό δεδομένων περιέχει **μετα-δεδομένα** (δηλ. δεδομένα για τα δεδομένα)
  - Σχήμα βάσης
  - *Γλώσσα Αποθήκευσης και Ορισμού Δεδομένων*
    - Γλώσσα που χρησιμοποιείται από το σύστημα βάσης δεδομένων για τον προσδιορισμό της δομής αποθήκευσης και των μεθόδων πρόσβασης
    - Συνήθως μια επέκταση της γλώσσας ορισμού δεδομένων
  - Περιορισμοί συνέπειας (Consistency constraints)

# Γλώσσες Βάσεων Δεδομένων: Γλώσσα Χειρισμού Δεδομένων (ΓΧΔ)

- Γλώσσα πρόσβασης και χειρισμού δεδομένων που είναι οργανωμένα από το κατάλληλο μοντέλο δεδομένων
  - Η ΓΧΔ είναι γνωστή και ως γλώσσα ερωτημάτων
- Δύο κλάσεις γλωσσών
  - Διαδικαστική
  - Μη-διαδικαστική

# Γλώσσες Βάσεων Δεδομένων: Γλώσσα Χειρισμού Δεδομένων (ΓΧΔ)

- Γλώσσα πρόσβασης και χειρισμού δεδομένων που είναι οργανωμένα από το κατάλληλο μοντέλο δεδομένων
  - Η ΓΧΔ είναι γνωστή και ως γλώσσα ερωτημάτων
- Δύο κλάσεις γλωσσών
  - Διαδικαστική – ο χρήστης καθορίζει τα δεδομένα που απαιτούνται και πώς θα έχουμε πρόσβαση σε αυτά
  - Μη-διαδικαστική / Δηλωτική – ο χρήστης καθορίζει τα δεδομένα που απαιτούνται χωρίς να προσδιορίζει πώς θα έχουμε πρόσβαση σε αυτά
- Η SQL είναι η πιο ευρέως χρησιμοποιούμενη γλώσσα ερωτημάτων

# SQL

■ SQL: ευρέως χρησιμοποιούμενη μη-διαδικαστική γλώσσα

- Πχ. Βρες το όνομα του πελάτη με κωδικό-πελάτη 192-83-7465

```
select πελάτης.όνομα-πελάτη  
from πελάτης  
where πελάτης.κωδ-πελάτη = '192-83-7465'
```

- Πχ. Βρες τα υπόλοιπα όλων των λογαριασμών που διατηρεί ο πελάτης με κωδικό 192-83-7465

```
select λογαριασμός.υπόλοιπο  
from καταθέτης, λογαριασμός  
where καταθέτης.κωδ-πελάτη= '192-83-7465' and  
καταθέτης.αριθμός-λογαριασμού=  
λογαριασμός.αριθμός-λογαριασμού
```

# SQL (συνέχεια)

- Τα προγράμματα εφαρμογών συνήθως προσπελαίνουν τις βάσεις δεδομένων μέσω μιας εκ των παρακάτω:
  - **Προεκτάσεις Γλωσσών (language extensions)** που επιτρέπουν **την εμφύτευση SQL** (embedded SQL) στο πρόγραμμα της φιλόξενης γλώσσας (χρειάζεται ένας προ-μεταγλωττιστής)
  - **Διεπαφή προγράμματος εφαρμογής (Application programming interface )** (δηλ. ένα σύνολο διεργασιών) (πχ. ODBC/JDBC) που επιτρέπουν την αποστολή SQL ερωτημάτων σε μια βάση δεδομένων
    - ODBC: Open DataBase Connectivity (χρησιμοποιείται στη C)
    - JDBC: Java DataBase Connectivity (χρησιμοποιείται στη Java)

# Χρήστες Βάσεων Δεδομένων

- Οι χρήστες διαφοροποιούνται ανάλογα με τον τρόπο που αναμένεται να αλληλεπιδράσουν με το σύστημα
  - **Προγραμματιστές εφαρμογών** – αλληλεπιδρούν με το σύστημα μέσω κλήσεων της ΓΧΔ
  - **Τεχνολογικά καταρτισμένοι χρήστες** – διατυπώνουν ερωτήματα προς τη βάση σε μια γλώσσα επερωτήσεων
  - **Εξειδικευμένοι χρήστες** – γράφουν εξειδικευμένες εφαρμογές βάσεων δεδομένων που δεν εμπίπτουν στο παραδοσιακό πλαίσιο διαχείρισης δεδομένων (πχ. Γνωσιακές βάσεις και έμπειρα συστήματα, CAD, κα)
  - **Απλοί χρήστες** – ενεργοποιούν ένα από τα μόνιμα προγράμματα εφαρμογών που έχουν προηγουμένως γραφτεί
    - Πχ. Άτομα που προσπελαίνουν μια βάση δεδομένων μέσω διαδικτύου, ταμείες τραπεζών, διοικητικό προσωπικό



# Διαχειριστής Βάσης Δεδομένων

- Συντονίζει τις δραστηριότητες ενός συστήματος βάσης δεδομένων (κατανοεί πλήρως τους πόρους που παρέχουν πληροφορία και τις ανάγκες)
- Καθήκοντα του διαχειριστή της βάσης δεδομένων:
  - Ορισμός σχήματος
  - Ορισμός δομών αποθήκευσης και μεθόδων πρόσβασης
  - Τροποποίηση σχήματος και φυσικής οργάνωσης
  - Εξουσιοδότηση χρηστών για πρόσβαση στη ΒΔ
  - Καθορισμός των περιορισμών ακεραιότητας
  - Συνδετικός κρίκος με τους χρήστες
  - Συντήρηση Ρουτίνας: παρακολούθηση απόδοσης, ανταπόκριση σε απαιτήσεις αλλαγών, τήρηση αντιγράφων, κα.

# Διαχείριση Δοσοληψιών

- Μια *δοσοληψία* είναι μια συλλογή από πράξεις που εκτελεί μια λογική συνάρτηση σε μια εφαρμογή βάσης δεδομένων
- Το συστατικό της διαχείρισης δοσοληψιών διασφαλίζει ότι η βάση δεδομένων παραμένει σε συνεπή (σωστή) κατάσταση ανεξάρτητα από τις αποτυχίες του συστήματος (πχ. απώλεια ισχύος και κατάρρευση λειτουργικού συστήματος) και τις αποτυχίες των δοσοληψιών
- Ο διαχειριστής ελέγχου ταυτόχρονης προσπάθειας (*concurrency-control manager*) ελέγχει την αλληλεπίδραση μεταξύ των ταυτόχρονων δοσοληψιών για να διασφαλίσει τη συνέπεια (και συνοχή) της βάσης δεδομένων

# Διαχείριση Αποθηκευτικού Χώρου

## ■ Διαχειριστής αποθηκευτικού χώρου (Storage manager)

- Ένα πρόγραμμα που παρέχει τη διεπαφή μεταξύ των αποθηκευμένων δεδομένων χαμηλού επιπέδου της βάσης και των προγραμμάτων εφαρμογής και των επερωτήσεων που υποβάλλονται στο σύστημα

## ■ Καθήκοντα διαχειριστή αποθηκευτικού χώρου:

- Αλληλεπίδραση με τον διαχειριστή αρχείων (file manager)
- Αποδοτική αποθήκευση, ανάκτηση και ενημέρωση δεδομένων

## ■ Συστατικά του διαχειριστή αποθηκευτικού χώρου:

- Διαχειριστής εξουσιοδότησης και ακεραιότητας (ελέγχει την ακεραιότητα, κτλ)
- Διαχειριστής δοσοληψιών (διασφαλίζει τη συνέπεια)
- Διαχειριστής αρχείου (διαχειρίζεται την ανάθεση χώρου)
- Διαχειριστής ενδιάμεσης μνήμης (buffer) (φέρνει δεδομένα από το δίσκο)

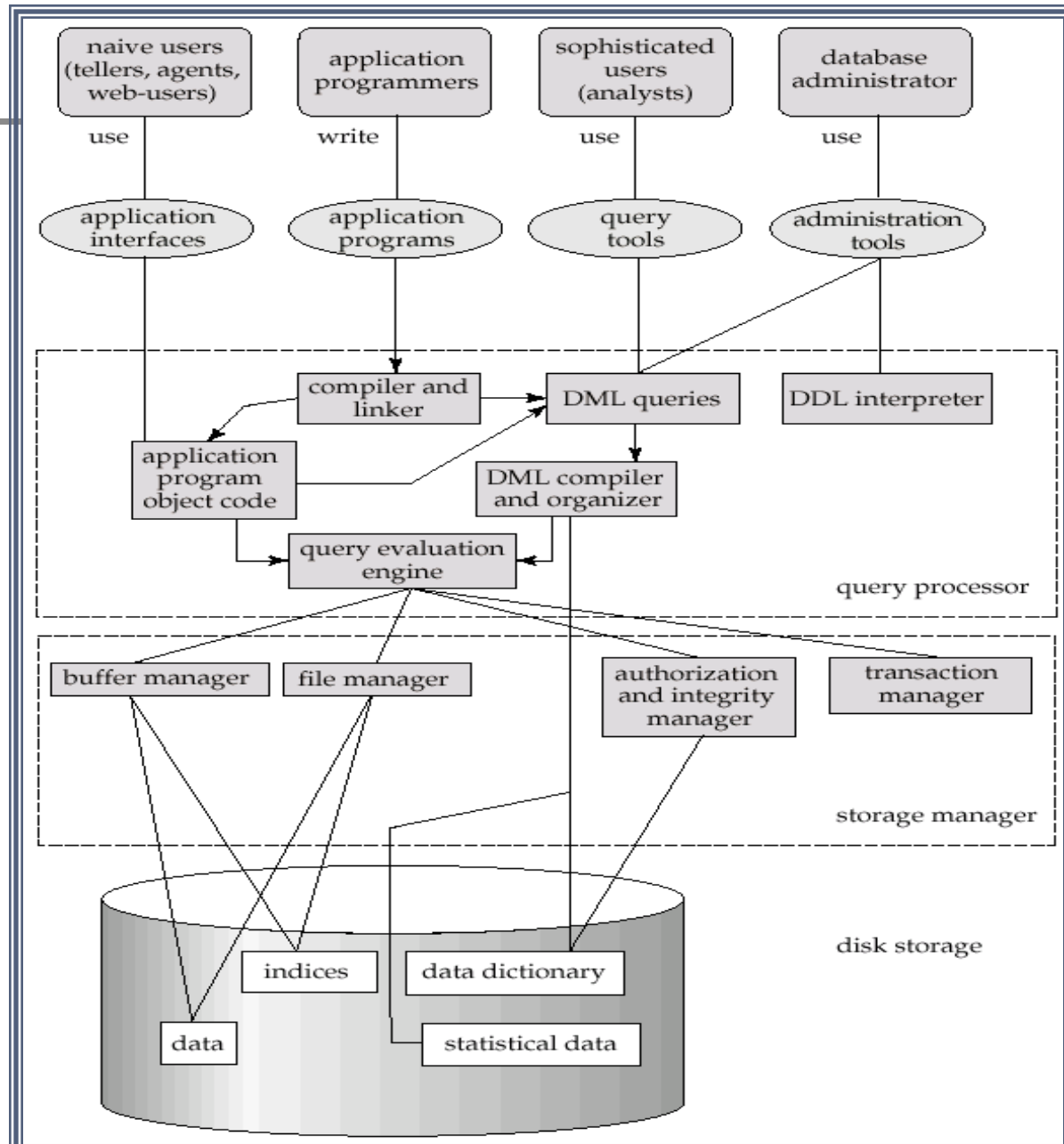


# Επεξεργασία Επερωτήσεων

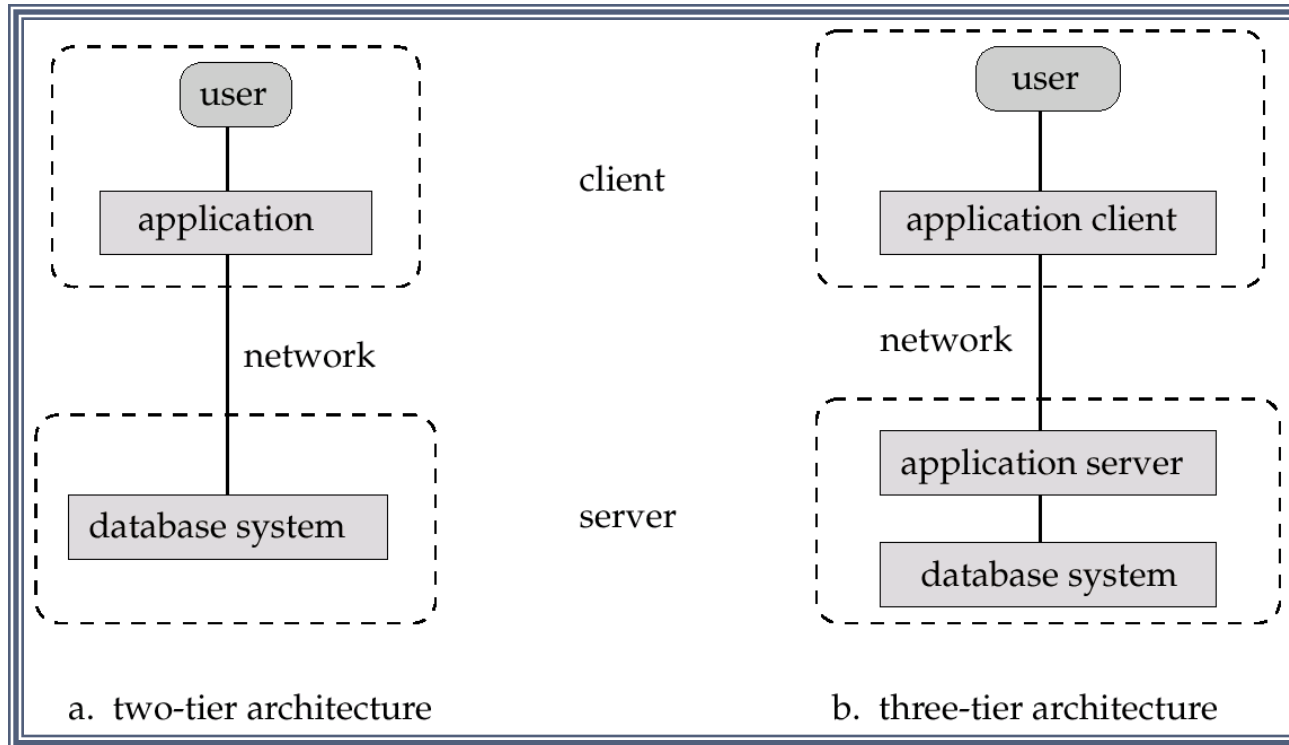
---

- Διερμηνέας ΓΟΔ (DDL interpreter )
  - Καταγράφει τους ορισμούς στο λεξικό δεδομένων
- Μεταγλωττιστής ΓΧΔ (DML compiler)
  - Μεταφράζει τη ΓΧΔ σε ένα πλάνο αξιολόγησης
- Μηχανή αξιολόγησης επερωτήσεων (Query evaluation engine)
  - Εκτελεί instructions χαμηλού επιπέδου που δημιουργούνται από τον μεταγλωττιστή της ΓΧΔ

# Γενική Αρχιτεκτονική Συστήματος



# Αρχιτεκτονικές Εφαρμογών



- **Αρχιτεκτονική δύο επιπέδων:** Π.χ. Προγράμματα εξυπηρετητών χρησιμοποιούν ODBC/JDBC για να επικοινωνήσουν με τη ΒΔ
- **Αρχιτεκτονική τριών επιπέδων:** Π.χ. Δικτυακές εφαρμογές και εφαρμογές που έχουν αναπτυχθεί με τη χρήση “middleware”