



**ΚΕΝΤΡΟ ΕΠΙΜΟΡΦΩΣΗΣ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗΣ**

# **NoSQL Βάσεις Δεδομένων**

**Μάρκος Καραμπάτσης**



# NoSQL Databases



## Τι είναι η NoSQL ΒΔ

Το NoSQL αντιπροσωπεύει το "Not Only SQL", καθώς μπορεί να υπάρχουν και περιπτώσεις όπου η NoSQL ΒΔ χρησιμοποιούνται μαζί με τις SQL ΒΔ.

Η NoSQL βάσεις δεδομένων παρέχουν ένα μηχανισμό αποθηκευσης και ανάκτησης δεδομένων που διαφέρει από τον παραδοσιακό μηχανισμό που έχουν οι σχεσιακές ΒΔ γιατί απλά δεν ακολουθούν τις αρχές των παραδοσιακών σχεσιακών ΒΔ.



# NoSQL Databases

## Βασικά χαρακτηριστικά

- Κύριο χαρακτηριστικό τους είναι ότι δεν τηρούν RDBMS μοντέλο (Relational Database Management System)
- Δεν χρησιμοποιούν δομημένο σύστημα για τα δεδομένα που περιλαμβάνουν όπως για παράδειγμα πίνακες
- Χρησιμοποιούν non-relational τρόπους οργάνωσης και ανάλυσης των δεδομένων
- Είναι σχεδιασμένες για να αποθηκεύουν δεδομένα μεγάλης κλίμακας. Ταυτόχρονα παρέχουν δυνατότητα παράλληλης επεξεργασίας δεδομένων μοιρασμένα σε έναν μεγάλο αριθμό από servers (Clusters)
- Χρησιμοποιούνται συνήθως σε δεδομένα μεγάλου όγκου και εφαρμογές ιστού που εκτελούνται σε πραγματικό χρόνο.



# NoSQL Databases

| Name             | Year | Type         | Developer          |
|------------------|------|--------------|--------------------|
| <b>MongoDB</b>   | 2009 | Document     | 10Gen              |
| <b>CouchDB</b>   | 2005 | Document     | Apache             |
| <b>Cassandra</b> | 2008 | Column Store | Apache             |
| <b>CouchBase</b> | 2011 | Document     | Couchbase          |
| <b>Riak</b>      | 2009 | Key-value    | Basho Technologies |
| <b>SimpleDB</b>  | 2007 | Document     | Amazon             |

## Βασικά χαρακτηριστικά

- Better Scaling
- Significantly Cheaper
- Schema-less (Partially)
- Less Management
- Cheaper Hardware
- Heterogeneous Data



# NoSQL Databases

## Γιατί, που και πότε είναι χρήσιμες

- Τα NoSQL συστήματα δεν εμφανίστηκαν για να αντικαταστήσουν τα σχεσιακά, αλλά για να τα συμπληρώσουν.
- Τα NoSQL συστήματα διαχείρισης βάσεων δεδομένων είναι εξαιρετικά χρήσιμα όταν κάποιος δουλεύει με τεράστιο όγκο δεδομένων, η φύση των οποίων δεν απαιτεί κάποιο σχεσιακό μοντέλο.
- Η κατακευματισμένη τους φύση τα καθιστά ιδανικά για μαζική επεξεργασία δεδομένων (πχ ενοποίηση, φιλτράρισμα, διαλογή, στατιστικές ενέργειες κλπ).



# NoSQL Databases

## Γιατί, που και πότε είναι χρήσιμες

- Είναι επίσης πολύ καλά για ανάκτηση και ανταλλαγή δεδομένων μεταξύ μηχανημάτων (machine-to-machine), καθώς και για την επεξεργασία συναλλαγών μεγάλου όγκου.
- Παρέχουν σχετικά φθηνή, υψηλής επεκτασιμότητας αποθήκευση μεγάλου όγκου. Παραδείγματα δεδομένων: ιστορικά δεδομένα, αρχεία καταγραφής, αρχεία τηλεφωνικών δεδομένων, ενδείξεις αισθητήρων κλπ.
- Έχουν την ικανότητα της κλιμάκωσης
- Τα NoSQL συστήματα έχουν πιο αδύναμα μοντέλα συνέπειας των δεδομένων, μπορούν να «θυσιάσουν» την συνοχή για την αποδοτικότητα

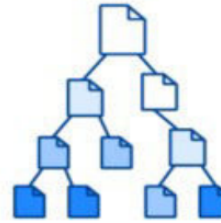


# Τύποι NoSQL ΒΔ

Υπάρχουν οι εξής 4 διαφορετικοί τύποι NoSQL βάσεων δεδομένων:

1. Document databases
2. Key-value stores
3. Wide-column stores
4. Graph databases

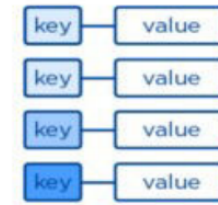
Document



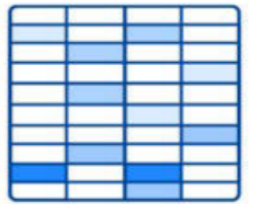
Graph



Key-Value



Wide-column





# Τύποι NoSQL ΒΔ

**Document databases**: Κάθε κλειδί (key) αντιστοιχεί σε μια σύνθετη δομή δεδομένων που ονομάζεται έγγραφο (document). Τα documents μπορούν να περιέχουν πολλαπλά ζεύγη key-value, ή key-array, ή φωλιασμένα έγγραφα.

Document databases:

- MongoDB
- Elastic Search
- CouchDB
- Terrastore
- OrientDB
- RavenDB

```
<Key=CustomerID>
{
  "customerid": "fc986e48ca6" ← Key
  "customer":
  {
    "firstname": "Pramod",
    "lastname": "Sadalage",
    "company": "ThoughtWorks",
    "likes": [ "Biking", "Photography" ]
  }
  "billingaddress":
  {
    "state": "AK",
    "city": "DILLINGHAM",
    "type": "R"
  }
}
```





# Τύποι NoSQL ΒΔ

**Key-value stores:** Η πιο απλή μορφή NoSQL ΣΔΒΔ. Κάθε στοιχείο στη βάση δεδομένων αποθηκεύεται με ένα όνομα κλειδί ('key'), μαζί με την τιμή του (value).

Key-value databases:

- Riak
- Redis
- Memcached
- Berkeley DB
- Amazon DynamoDB (not open-source)
- Project Voldemort and Couchbase

| Key        | Value        |
|------------|--------------|
| Name       | Joe Bloggs   |
| Age        | 42           |
| Occupation | Stunt Double |
| Height     | 175cm        |
| Weight     | 77kg         |



# Τύποι NoSQL ΒΔ

**Wide-column stores:** Έχουν βελτιστοποιηθεί για να απαντούν ερωτήματα σε μεγάλα σύνολα δεδομένων και αποθηκεύουν ενιαίες τις στήλες των δεδομένων (αντί για τις πλειάδες).

Column-store databases:

- Cassandra
- Base
- Hypertable
- Amazon DynamoDB

| ColumnFamily |             |       |       |
|--------------|-------------|-------|-------|
| Row Key      | Column Name |       |       |
|              | Key         | Key   | Key   |
|              | Value       | Value | Value |
|              | Column Name |       |       |
|              | Key         | Key   | Key   |
|              | Value       | Value | Value |
|              |             |       |       |

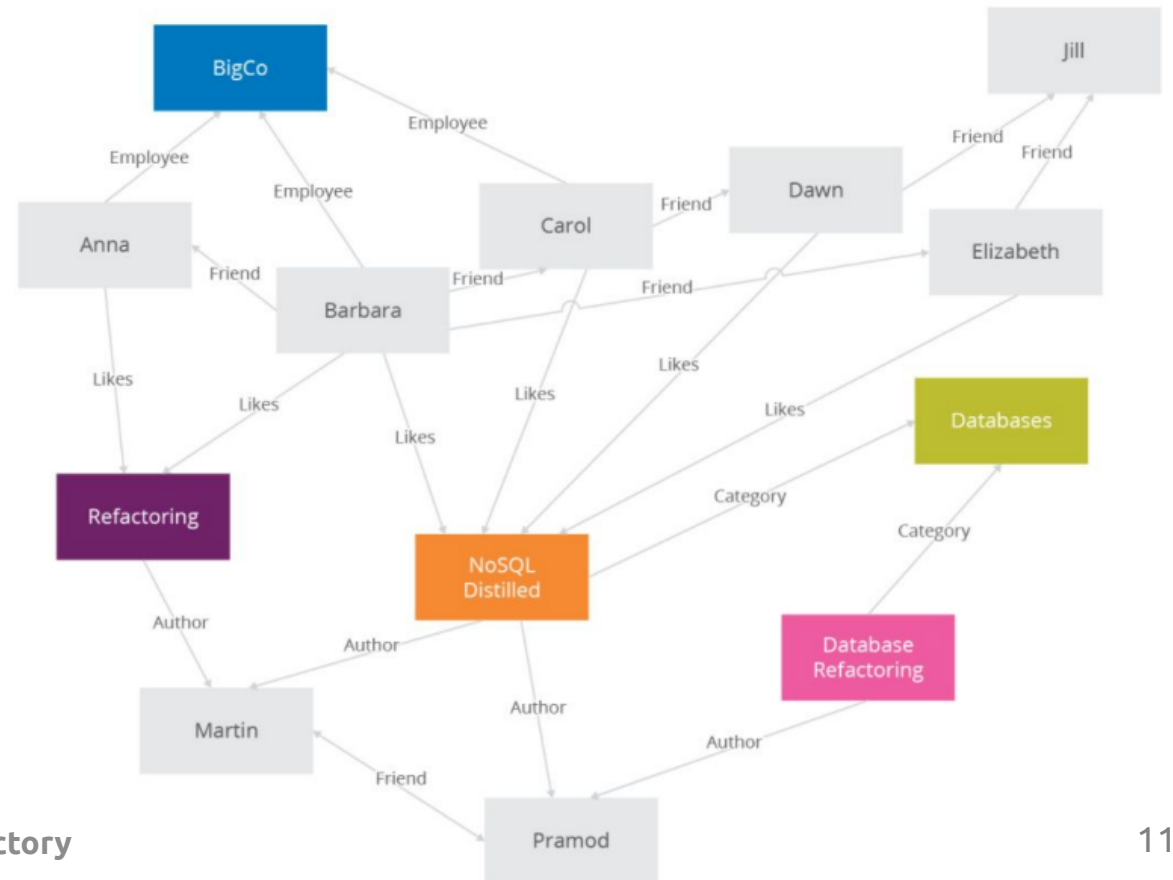


# Τύποι NoSQL ΒΔ

**Graph stores:** Χρησιμοποιούνται για να αποθηκεύσουμε πληροφορία για δίκτυα δεδομένων όπως π.χ. για τις συνδέσεις σε ένα κοινωνικό δίκτυο.

Graph databases:

- Neo4J
- RedisGraph
- OrientDB
- DGraph





# Τύποι NoSQL ΒΔ

## Χαρακτηριστικά ανά κατηγορία NoSQL συστημάτων

| Μοντέλο Δεδομένων | Απόδοση  | Επεκτασιμότητα  | Ευελιξία | Πολυπλοκότητα | Λειτουργικότητα    |
|-------------------|----------|-----------------|----------|---------------|--------------------|
| Key-value         | high     | high            | high     | none          | Variable (none)    |
| Column            | high     | high            | moderate | low           | minimal            |
| Document          | high     | Variable (high) | high     | low           | Variable (low)     |
| Graph             | variable | variable        | high     | high          | graph theory       |
| Relational        | variable | variable        | low      | moderate      | relational algebra |



# RDBMS vs NoSQL: Data Modeling Παράδειγμα

## RDBMS

### Users

| ID | first_name | last_name | cell       | city   |
|----|------------|-----------|------------|--------|
| 1  | Leslie     | Yepp      | 8125552344 | Pawnee |

### Hobbies

| ID | user_id | hobby          |
|----|---------|----------------|
| 10 | 1       | scrapbooking   |
| 11 | 1       | eating waffles |
| 12 | 1       | working        |

## NoSQL

```
{
  "_id": 1,
  "first_name": "Leslie",
  "last_name": "Yepp",
  "cell": "8125552344",
  "city": "Pawnee",
  "hobbies": ["scrapbooking", "eating waffles", "working"]
}
```