



Basi di dati Avanzate

- presentazione del corso -

Prof. Matteo Golfarelli

Alma Mater Studiorum - Università di Bologna

Caratteristiche

□ Docenti

- Prof. Matteo Golfarelli (Modulo I - Teoria)
 - matteo.golfarelli@unibo.it
 - Orario di ricevimento: Ven 11-13
- Prof. Alessandra Lumini (Modulo II - Laboratorio)
 - alessandra.lumini@unibo.it
 - Orario di ricevimento: Lun 11-13

□ *Propedeutico* per il profilo Knowledge & Data Engineering

One size does not fit all!

- ❑ Per oltre 40 anni l'unico modello dati disponibile era quello relazionale
 - ✓ Applicazioni con necessità/caratteristiche diverse venivano implementate sullo stesso modello
 - ✓ Le performance erano limitate



- ❑ La necessità di supportare applicazioni con vincoli di performance stringenti ed enormi moli di dati ha portato alla nascita dei DBMS NoSQL
 - ✓ Ogni modello ha caratteristiche diverse, specifiche per carichi di lavoro
- ❑ Il progettista deve conoscere le caratteristiche e i principi di modellazione dei diversi modelli



One size does not fit all!

- ❑ Per oltre 40 anni l'unico modello dati disponibile era quello relazionale
 - ✓ Applicazioni con necessità/caratteristiche diverse venivano implementate sullo stesso modello
 - ✓ Le performance erano limitate

Modello	Descrizione	Casi d'uso	Applicazioni
Key-value	Associates any kind of value to a string	Dictionary, lookup table, cache, file and images storage	Web session profile, shopping cart, user preferences
Document	Stores hierarchical data in a tree-like structure	Documents, anything that fits into a hierarchical structure	Event log, CMS, blogging platform
Wide column	Stores sparse matrixes where a cell is identified by the row and column keys	Crawling, high-variability systems, sparse matrixes	Event log, CMS, blogging platform, GIS
Graph	Stores vertices and arches	Social network queries, inference, pattern matching	Social network, routing application, fraud detection

Studia
l'applicazione

Scegli il
modello

Modella
l'applicazione

Obiettivi

L'utilizzo del modello relazionale non è più sufficiente a rispondere ai requisiti delle applicazioni in cloud e più in generale delle applicazioni data intensive. Per affrontare qualsiasi applicazione data intensive è necessario:

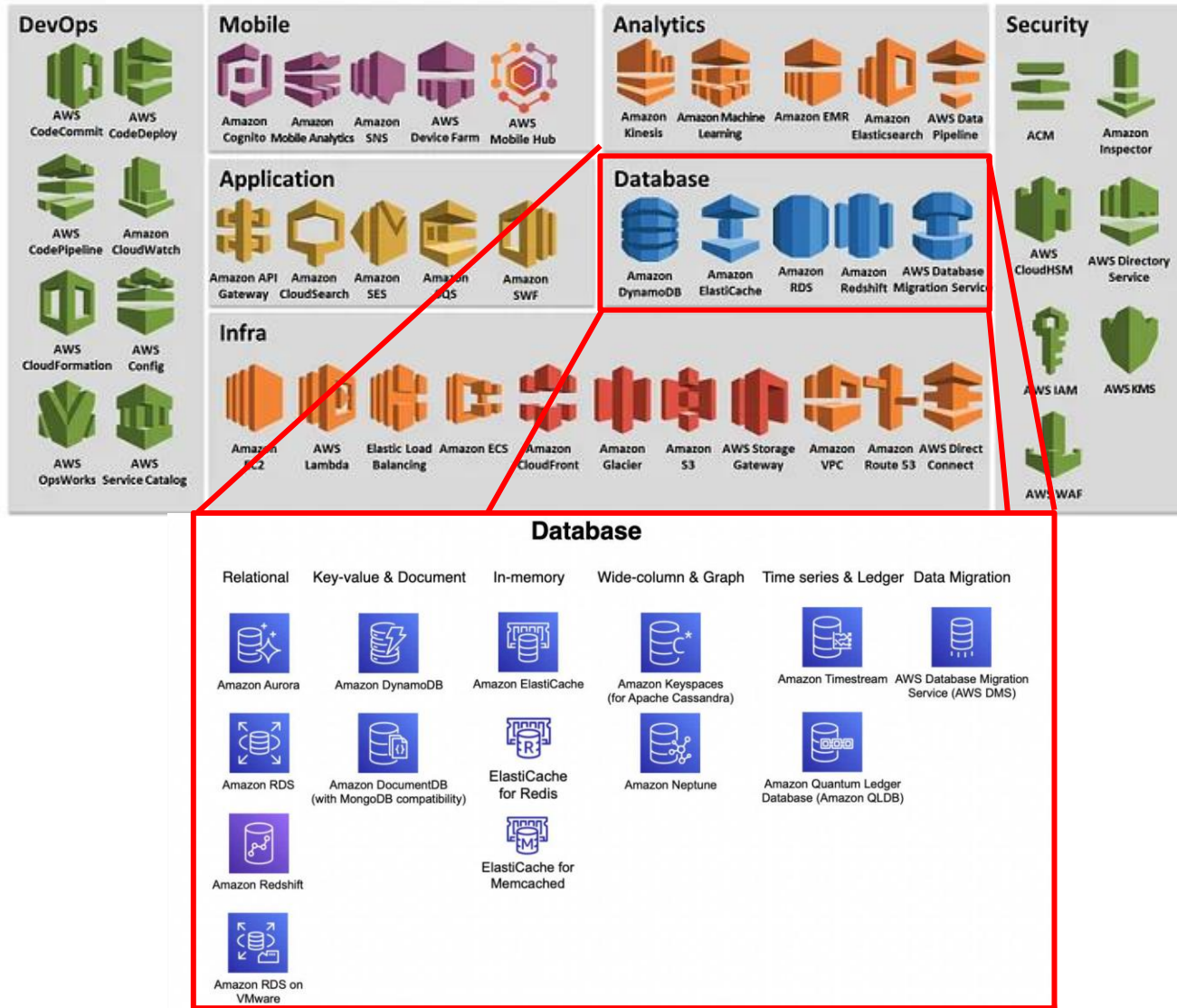
- ❑ Padroneggiare tutti i **modelli dati** offerti sul mercato.
 - Conoscere le caratteristiche dei diversi modelli dati
 - Saper progettare correttamente in modelli dati diversi
- ❑ Avere conoscenze approfondite sulla **gestione** e **realizzazione** delle basi di dati
 - Nel corso di basi di dati si sono affrontati i temi della
 - Progettazione: modello ER e modello relazionale
 - Interrogazione: linguaggio SQL
 - In questo corso si eliminerà il livello di astrazione del modello relazionale per curare gli aspetti di:
 - Amministrazione
 - Programmazione
 - Ottimizzazione in laboratorio

Programma

- ❑ Amministrazione di basi di dati architettura logica e fisica dei DBMS
- ❑ Programmazione avanzata di basi di dati Costrutti avanzati di SQL
- ❑ Ottimizzazione delle prestazioni Il processo di ottimizzazione delle query SQL
 - Esercitazioni sul DBMS relazionale Oracle
- ❑ Modelli dati NOSQL
 - Caratteristiche dei modelli key-value e Document-based, Graph-based
 - Consistenza nei DBMS NoSql
 - Progettazione con il modello documentale
 - DBMS NoSQL: MongoDB – sharding dei dati e cluster
- ❑ L'enfasi sarà posta sulla **componente implementativa** al fine di fornire allo studente **competenze pratiche** sugli strumenti
 - il 50% delle lezioni si svolgeranno in laboratorio

I sistemi NoSQL

- Rappresentano il presente e il futuro delle Data Platform



Modalità di esame

- ❑ L'esame si compone di una prova teorico-pratica
- ❑ **Prova pratica** (da svolgersi sul calcolatore 100 minuti con tutti gli appunti a disposizione)
 - Realizzazione di stored procedure e trigger
 - Svolgimento di attività di gestione di un DBMS
 - Analisi di piani di esecuzione
 - Stima del costo di esecuzione di query SQL
 - Interrogazioni con MongoDB
- ❑ **Prova teorica** (scritta 30 minuti)
 - Domande aperte su tutti gli argomenti trattati
- ❑ Le due prove si svolgeranno nella stessa giornata
- ❑ Il voto è determinato dalla media pesata delle due prove

Informazioni pratiche

- ❑ Sito del corso: **Virtuale**
 - Dispense e materiale didattico
 - Orari
 - Risultati degli esami

- ❑ 5 ore di lezione settimanale
 - L'alternanza dei docenti potrà cambiare durante il corso

Domande?

