

Tecnologie e architetture per la gestione dei dati 2023-2024

Docenti	Stefano Iannucci stefano.iannucci@uniroma3.it Paolo Atzeni paolo.atzeni@uniroma3.it Luigi Bellomarini luigi.bellomarini@uniroma3.it
Periodo	01/03 – 14/06/2024
Orario	lunedì, giovedì e venerdì ore 08-10 (8:15-9:45) in aula N14
Ricevimento	concordare per posta elettronica
Sito del corso	raggiungibile dalle home-page dei docenti (attivo dal 4/03)
Iscrizione	sul server Moodle, vedi sito
Teams	vedi sito
Mailing list	sempre attraverso Moodle
Forum	pure su Moodle

Prerequisiti

- Si assume che lo studente abbia una conoscenza di base di modelli, linguaggi e sistemi per basi di dati e sia quindi in grado di utilizzare un DBMS, ad esempio con il linguaggio SQL, e sia in grado di progettare una base di dati di media dimensione
- In sostanza:
 - I contenuti di Basi di dati I

Obiettivi formativi

- Esporre modelli, metodi e sistemi fondamentali per la tecnologia delle basi di dati (strutture fisiche e transazioni) insieme ad alcune recenti direzioni di evoluzione della tecnologia stessa.
- Affrontare le recenti direzioni di evoluzione delle metodologie e tecnologie delle basi di dati, con riferimento alle principali famiglie di problemi di interesse: integrazione di basi di dati eterogenee e autonome; utilizzo di basi di dati per applicazioni di analisi e supporto alle decisioni.
- Superato il corso, lo studente conoscerà le tecnologie fondamentali su cui sono basati i DBMS relazionali e le principali metodologie e tecnologie per l'integrazione e per lo sviluppo di datawarehouse.

Programma

- Tecnologia delle basi di dati attuali:
 - strutture fisiche
 - gestione delle transazioni e architetture distribuite
 - tecnologie per il monitoraggio delle basi di dati e strumenti per la modellazione delle prestazioni
- Seminari (da confermare)

Riferimenti

- Testo
 - P. Atzeni, et al. Basi di dati 5/Ed 2018 o 6/Ed 2023
oppure
 - P. Atzeni, et al. Basi di dati: architetture e linee di evoluzione. McGraw-Hill, 2003 o seconda edizione 2007
 - D. Menasce et al., Performance by Design: Computer Capacity Planning by Example, Prentice Hall 2004
- Materiale disponibile in rete in aggiornamento

Esami

- Modalità:
 - prova scritta (peso 67%)
 - progetto su Capacity Planning di un database (peso 30%)
 - svolto individualmente o in gruppi di due persone
 - valutato con una discussione orale con i docenti
 - ripresentabili in caso di esito negativo dell'esame
 - svolgimento di due o tre piccoli progetti
 - finalizzati alla preparazione allo scritto
 - non specificamente valutati
 - svolti anche in collaborazione con consegna individuale
 - modalità più precise via via
 - da presentare insieme allo scritto (o prima)
- Appelli: (date da definire) giugno o luglio, settembre, febbraio

Esami:

modalità ad "accertamento continuo"

- Valutazione basata su due prove e progetto:
 1. fra il 1 e il 15 aprile (da confermare)
 2. fra il 10 e il 25 maggio (da confermare)
 3. Progetto, da discutere a fine corso (entro pochi giorni)
- Per essere ammessi alla seconda prova è necessario superare la prima
- Per sostenere le prove è necessario consegnare i progetti e le prove di autovalutazione secondo le scadenze che verranno indicate

Esami e prove in itinere

	2022-2023			2021-2022			2020-2021			2019-2020			2018-2019		
	Promossi	Non promossi	Totale	Promossi	Non promossi	Totale	Promossi	Non promossi	Totale	Promossi	Non promossi	Totale	Promossi	Non promossi	Totale
Prove in itinere	43	0	43	28	0	28	26	0	26	28	0	28	45	0	45
Primo appello senza prove in itinere	1	0	1	4	1	5	6	2	8	20	3	23	8	11	19
Secondo appello	0	0	0	1	1	2	5	2	7	6	3	9	11	3	14
Terzo appello	2	1	3	4	0	4	1	1	1	6	0	6	2	2	4

Metodo di studio

- studio individuale, con riflessione sui concetti e riferimento alle esperienze personali
- svolgimento degli homework e dei progetti

Software

- Un DBMS (Postgres) e un DBMS didattico (Simple DB), da utilizzare per
 - verifiche di vari concetti
 - lo sviluppo di esercitazioni e dei progetti
- Strumenti di monitoraggio di database e sistemi, per il salvataggio di serie temporali, per la visualizzazione delle metriche di interesse, per la simulazione, e per la virtualizzazione (e.g., Prometheus, InfluxDB, Grafana, Java Modeling Tools, Docker)
 - per un progetto