

Scaletta workshop

1. (10 min) **Introduzione**

- Che cos'è un dato? e un database? ...
- Ciclo di vita del dato: *raccolta* → *archiviazione* → *elaborazione*
- Categorie di database: relazionali vs non relazionali
- Cenni storici sui database relazionali: come siamo arrivati al relazionale?

2. (20 min) **Modello relazionale:**

- Struttura logica: tabelle, chiavi, relazioni, ...
- Vincoli di integrità, paradigma ACID
- vantaggi: consistenza garantita, linguaggio dichiarativo (SQL), integrità
- *Problemi*: scalabilità verticale, rigidità dello schema, performance

3. (20 min) **Modello relazionale, Distribuito**

- Come funziona un database relazionale distribuito (e.g. HA PostgreSQL)
- Limitazioni
- *CAP theorem*
- Considerazioni su i problemi di networking e bandwidth

4. (35 min) **Modelli non relazionali**

- Cambiare paradigma: NoSQL, quando e perché
- Cenni sulla diversità all'interno del mondo NoSQL
- esempi di database NoSQL: MongoDB, QuestDB

- - - break 15 min- - -

5. (25 min) **Small live demo**

6. (45 min) **Dove salvare i dati?**

- Scelta dello storage appropriato:
 - locale vs distribuito
- File system distribuiti: concetti di base:
 - Replica, ridondanza, resilienza
 - Che succede in caso di guasto?
- Esempi: perché NFS non è un file system distribuito. CEPH

7. (10 min) **Domande?**

Totale minuti: 180