CS143: Database Systems

Professor Cho

Thilan Tran

Winter 2021

Contents

CS143: Database Systems

- **DBMS (Database Management System)** is a way to manage and store data:
  - how does a database differ from a spreadsheet software like Excel?
    - expected to *efficiently* scale to a *massive* amount of data, without suffering
    - expected to persist the data
    - expected to be conveniently used by a large number of clients at a time
    - expected to provide secured and safe access to data
  - program data is not assumed to entirely reside in main memory RAM, but instead in disk
    - this leads to utilization of different data structures
- database architecture:
  1. disk for data (sometimes stored in main memory, if data can fit)
  2. OS
  3. DBMS engine
     - database system may access disk through OS, or directly through raw IO
  4. API
     - eg. standard APIs like JDBC (Java), ODBC (Microsoft)
  5. app, or CLI
  - downloading a DBMS software like MySQL installs parts 3, 4, and CLI
- popular DBMS software:
  - relational:
    - open source: MySQL, PostgreSQL
    - closed source: Oracle, Microsoft SQL, IBM DB2
  - non-relational (NoSQL):
    - MongoDB, Spark