Relational Databases

Bio-Info Club
IJM
09/03/12

Ataaillah BENHADDOU
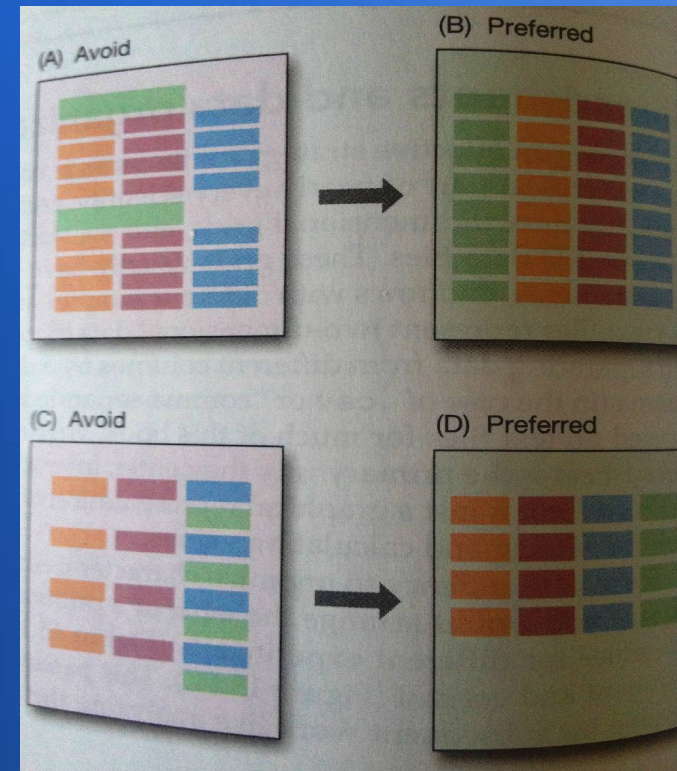
Storing/manipulating information

- Storing in text file:
  - 2D grids :
    - .csv, .txt fils
    - 2D : row and columns

  - XML format
- Manipulating information
  - Spread sheet : a graphical representation of this two dimentional grid with tools for editing and calculating values

The need to relational databases

- For complex and big databases :
  - Ex : molecular sequences of multiple genes
    - Multiple files linked to a main file were to store individual gene sequences.
    - It's is generally faster to find a piece of information in database than to scan a through a large text file.
- RDBMS : relational database management system
  - A server program that runs continuously in the background and manages one or more databases.
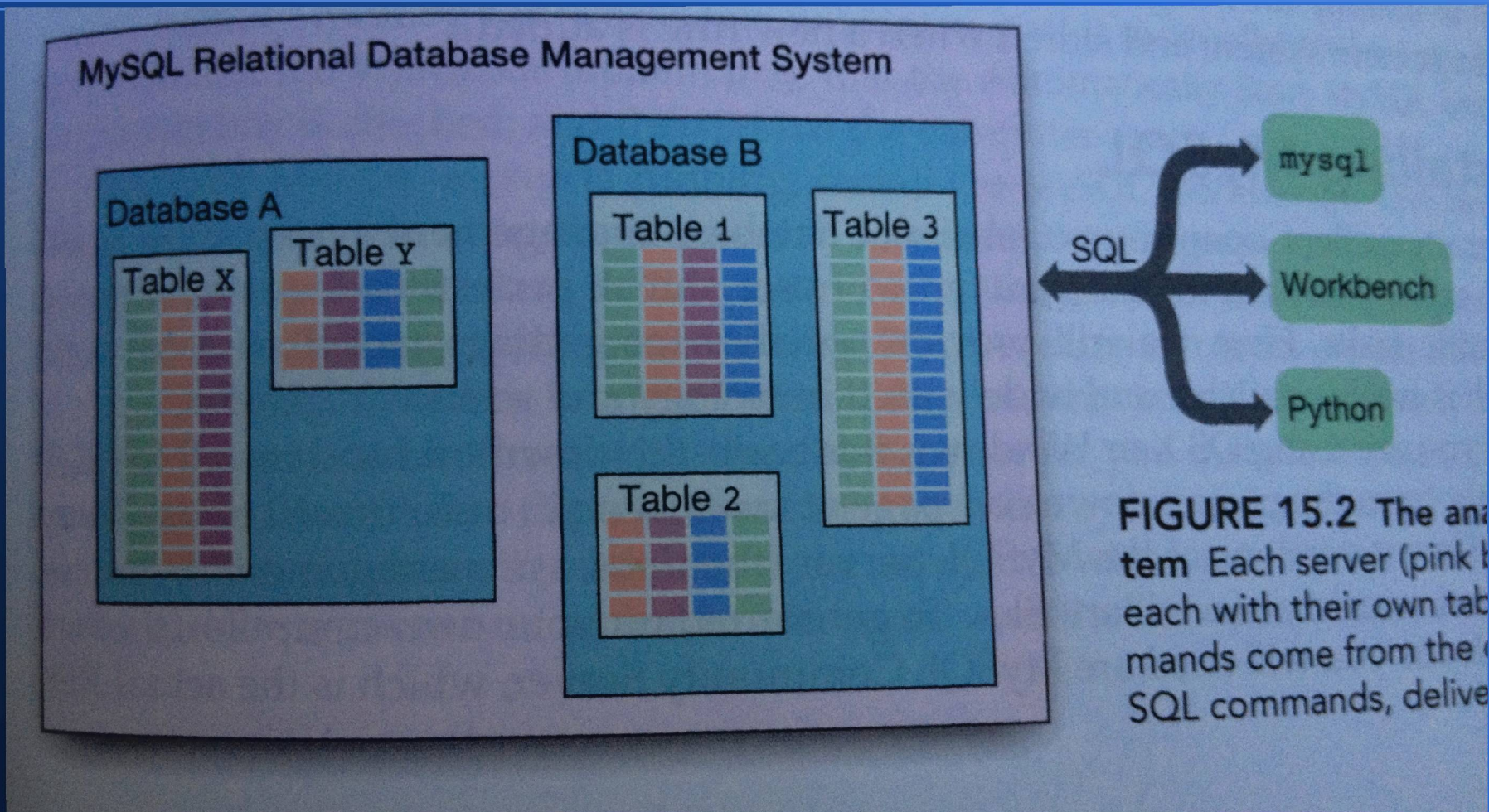
RDBMS

- Creation, organisation and optimization of databases files

- Receive requests to add, edit and look up data.

  - Commercial : FileMaker, Microsoft Office Access, Microsoft SQL Server, the Oracle software suites

  - Open source : MySQL, PostgreSQL, SQLite.

- Uses Structured Query Language (SQL)

  - Directly

    - Commande-line

    - Graphical interface.

  - Indirectly :

    - Software packages : R, Python, MATLAB

    - Web servers

DBMS

- Allows centralized database files :

    - Easly backed up

    - Avoid redundancy(losing time for backing up and space for storing, reduces the chance of inconsistancies)

    - Get several programs talk each other

    - More than one program can access the data at the same time

    - Flexibility, efficient in therm of memory and computations

    - The driving concept : each piece of information is stored only once, and then liked through relation to other pieces rather than copied

Anatomy of a database

FIGURE 15.2 The ana[...]
tem Each server (pink [...]
each with their own tab[...]
mands come from the [...]
SQL commands, delive[...]

Common RDBMS data types

TABLE 15.1 Common RDBMS data types

Data type	Description
INTEGER	An integer ranging in value from -2147483648 to 2147483647; INT can be used as an abbreviation for INTEGER
FLOAT	A floating point number, including scientific notation: 3.14159 or 6.022e+23
DATE	A date in 'YYYY-MM-DD' format
DATETIME	A date and time in 'YYYY-MM-DD HH:MM:SS' format
TEXT	A string containing up to 65535 characters
TINYTEXT	A string containing up to 255 characters
BLOB	A piece of information encoded in binary, including images or other non-text data; there are four sizes of blob data types, with different storage capacities

Hands on session !

- www.mysql.com/download
  - MySQL community server
  - MySQL Workbench