



Бази от данни

Въведение в курса

доц. д-р Димитър Димитров

Организационни въпроси

- Да се запознаем
 - Преподавателски екип
 - Студенти – изкарани курсове, интереси
- Курс в <http://learn.fmi.uni-sofia.bg>
 - Учебни материали
 - Съобщения за контролни и изпити
- Междучасия в 9:00 и в 10:10

Оценяване

- Начин на оценяване
 - Контролно през семестъра – задачи
 - Контролно през семестъра – теория
 - Изпит през сесията
 - Писмена част – задачи, минимална оценка: 3.00
 - Теоретична част – тест, минимална оценка: 3.00
- Държавен изпит след 4-и курс

Учебни материали (лекции и упражнения)

- Презентации в Moodle
- Hector Garcia-Molina et.al.
Database Systems: The Complete Book
- Р. Христова, Д. Димитров.
Ръководство по бази от данни.
Университетско издателство "Св.
Климент Охридски"



База от данни



- Една от многото дефиниции:
 - БД е структурирана колекция от данни, организирана за ефективно съхранение, извличане и манипулиране
- Ключови аспекти от дефиницията

СУБД

- СУБД: Система за управление на бази от данни
- Софтуерна система, позволяваща създаването, достъпа, модификацията и администрирането на БД
- Интерфейс между потребители (напр. софтуерни приложения) и БД

Защо?

- Защо не използваме файловата система?
- Данните на една БД така или иначе се съхраняват във файлове

Защото...

- Нека имаме таблица с 1 млн. реда в .txt или .xlsx файл
- Искаме например:
 - Да намерим клетка с текст, отговарящ на даден шаблон
 - Да добавим колона
 - И т.н.
 - И няколко потребители да извършват подобни операции едновременно

Модели на данни

- Абстрактен модел
- Определя структурата на данните
- Популярни модели на данни:
 - Релационен
 - Йерархичен
 - Същност-връзки (Entity-Relationship)
 - Обектно-ориентиран
 - Графов

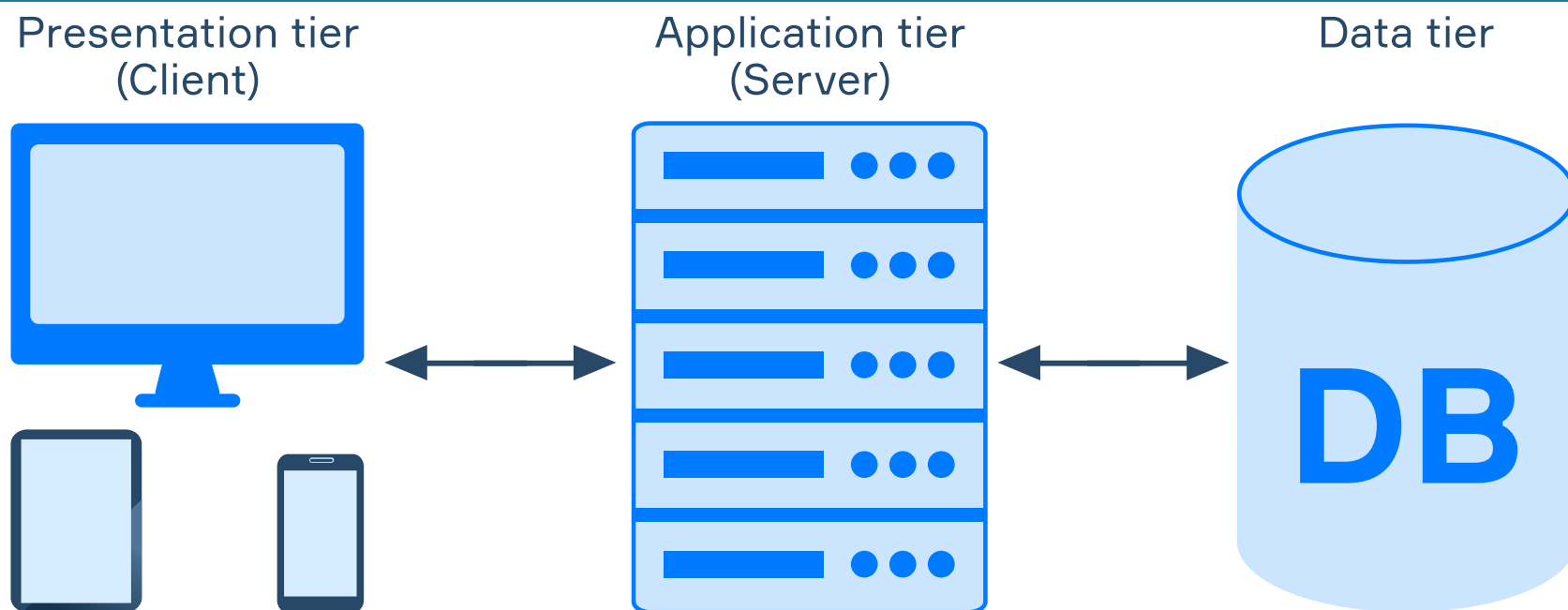
Видове БД

- Конкретни реализации на даден модел на данни
- Популярни видове БД:
 - Релационни
 - Йерархични
 - Обектно-ориентирани
 - И др.

Релационен модел

- Таблици/релации
- Връзки между таблиците – чрез ключове
- Интегритет на данните – чрез ограничения
- Нормализация – за намаляване на излишествата и подобряване на интегритета
- Математическа основа
- Мощен език за заявки - SQL

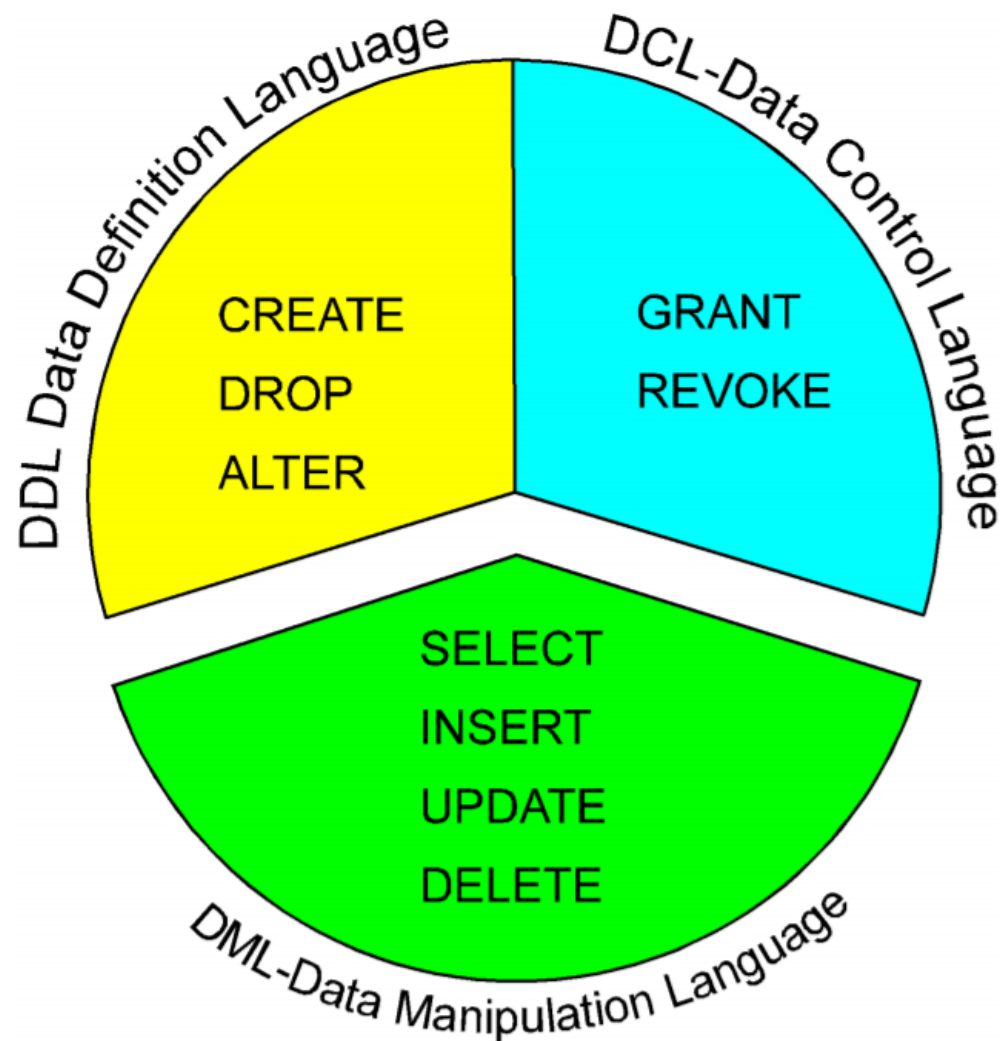
Примерно приложение



- Какви таблици бихме използвали за съхранение на данните на уеб-базирана поща?
- Как ще реализираме основни операции (log-in и т.н.)?

SQL

- Какви усилия биха били нужни, ако трябваше сами да реализираме тези операции?



CAP теорема

- Всяка разпределена система за съхранение на данни може да удовлетвори най-много две от следните свойства:
- **Consistency** (консистентност, съгласуване на данните)
 - Всяко прочитане на данни връща последно записаните данни (или грешка)
- **Availability** (наличност, достъпност)
 - Всяка заявка получава отговор (не и грешка), без да е гарантирано, че съдържа данните от последното записване
- **Partition tolerance** (устойчивост на разпределение)
 - Разделянето на системата на няколко изолирани секции (nodes) (или загубата на съобщения между секциите) не води до некоректни отговори
- При проблем в комуникацията между секциите:
 - Или операцията се отменя => консистентност, но не и наличност
 - Или се продължава => наличност, но не и консистентност

ACID

- Свойства на транзакции в (релационни) БД
- Гарантират надеждността на транзакциите
- **Atomicity**
 - Всичко или нищо
- **Consistency**
 - Различно от CAP: транзакция води БД от едно консистентно състояние в друго
- **Isolation**
 - Конкурентно изпълнение на транзакции води до същите резултати, както и последователното
- **Durability**
 - След като е утвърдена (committed), транзакцията остава дори и след срив

Компоненти на СУБД

- Компилятор на заявки (Query compiler)
- Машина за изпълнение (Execution engine)
- Управление на индекси, файлове и записи (Index/file/record manager)
- Управление на буфера (Buffer manager)
- Управление на паметта (Storage manager)
- Управление на транзакции (Transaction manager)
- Контрол на конкурентостта (Concurrency control)
- Компилятор на DDL (DDL compiler)
- Управление на логове и възстановяване (Logging and recovery)

Популярни релационни СУБД

- Microsoft SQL Server
- PostgreSQL
- MySQL
- SQLite
- MariaDB (клон на MySQL)
- Oracle
- IBM DB2

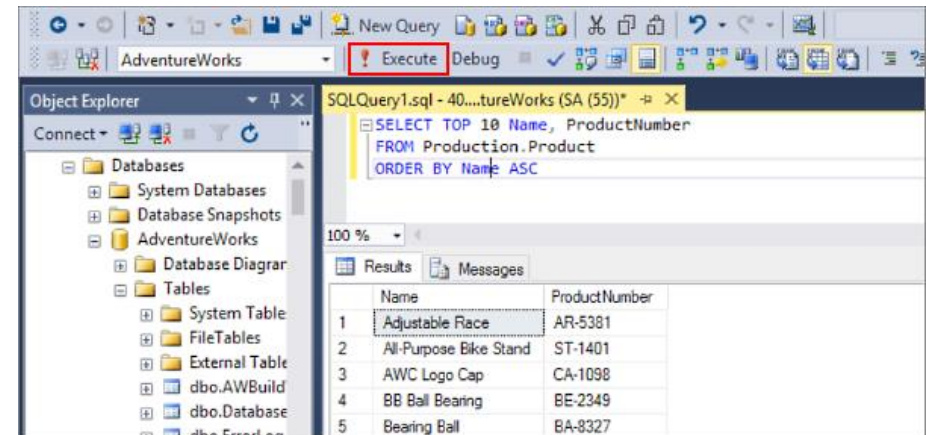


SQL – диалекти

- Transact-SQL (T-SQL)
- PL/SQL
- ...
- В “Ръководство по БД” са разгледани някои разлики

Популярни интегрирани среди

- SQL Server Management Studio
(за MS SQL Server)
- Azure Data Studio
- DBeaver
- DataGrip
- MySQL Workbench
- Различни command-line
tools
- И много други



```
jawadhasan80@hexquote-ub:~$ sudo -i -u postgres
postgres@hexquote-ub:~$ psql
psql (9.5.24)
Type "help" for help.

postgres=# \l

```

Name	Owner	Encoding	Collate	Ctype	Access privileges
postgres	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	
template0	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	=c/p/postgres
template1	postgres	UTF8	en_US.UTF-8	en_US.UTF-8	=c/p/postgres

```
(3 rows)

postgres=#
```

Популярни нерелационни БД

- Ключ-стойност
(Redis, Amazon DynamoDB)
- Документно-ориентирани
(MongoDB, CouchDB)
- Графови
(Neo4j, OrientDB)
- И други

Какво следва?

- Основни теми:

Модел “Същност-връзки” (Entity-Relationship Model)



Релационен модел



Функционални зависимости



Релационна алгебра



Нормализация

- Допълнителни теми

Въпроси?