# 10. Databázový procesor - teorie, pojmy

## Databáze

* Uspořádaná množina dat (informací) uložena na paměťovém mediu
* Součástí databáze jsou softwarové prostředky pro manipulaci s daty + přístup k datům

### Rozdělení

* Systémy sálových počítačů (Mainframe)
* dBase
  + Souborově orientované databáze s indexsekvenční metodou přístupu
  + Každá tabulka má samostatný .dbf soubor
  + Software: dBase, FoxPro, Paradox, Access
* Relační databázové systémy (lepší datová integrita, bezpečnost…)
* Objektově orientované databáze (specializované uplatnění, data se ukládají jako objekt s vlastnostmi)

## Databázový procesor

* Nástroj, který slouží pro práci s velkým množstvím dat; MS Access, Firebird, Oracle
* V databázi se data upravují, ukládají, získávají
* Obsahuje jednotlivé akce – moduly:
  + Tabulka
  + Dotazy
  + Formuláře
  + Sestavy

## SŘBD; DBSŘ; DBMS

* Systém řízení báze dat; Databázový systém řízení; Database management system
* Softwarové vybavení, které zajišťuje práci s databází (tvoří rozhraní mezi aplikačními programy a uloženými daty)
* **Databázová aplikace** je program, který umožňuje vybírat, prohlížet a aktualizovat informace uložené prostřednictvím SŘBD
* SŘBD musí být schopen efektivně pracovat s velkým množstvím dat a také musí být schopný řídit data (vkládat, modifikovat, mazat) a definovat strukturu těchto dat

### Služby

* Definice dat (definování a uchovávání datové entity)
* Údržba dat (každému členu entity vyhrazuje záznam skládající se z položek)
* Manipulace s daty (služby umožňující vkládání, aktualizaci, rušení a třídění dat)
* Zobrazování dat (poskytuje metody prezentace dat uživateli)
* Integrita dat (metody pro zajištění správnosti dat nepovolením vložení duplicitního řádku s unikátním klíčem)

## Architektury DB

### Centrální

Tato architektura je typická pro terminálovou síť, kdy se po síti přenáší vstupní údaje z terminálu na centrální počítač do příslušné aplikace, výstupy z této aplikace se přenáší na terminál. Protože aplikační program i vlastní zpracování probíhá na centrálním počítači, který může zpracovávat více úloh, mají odezvy na dotazy určité zpoždění.

* Data i SŘBD jsou v centrálním počítači

### File-Server

Tato metoda souvisí zejména s rozšířením osobních počítačů a sítí LAN.

* SŘBD a databázové aplikace jsou na jednotlivých počítačích
* Data jsou na File-Serveru

##### Komunikace uživatele se systémem:

* Uživatel zadá dotaz
* SŘBD přijme dotaz, zasílá požadavky na data file-serveru
* File-server posílá bloky dat na lokální počítač, kde jsou data zpracovávána podle zadaného dotazu (vyhledávání, setřídění…)
* Výsledek dotazu se zobrazí uživateli

### Klient-Server

V podstatě je založena na lokální síti (LAN), personálních počítačích a databázovém serveru. Na počítačích běží program pro komunikaci se serverem.

* Redukuje množství přenesených dat (v porovnání s File-Serverem)

##### Komunikace:

* Uživatel zadává dotaz (buď přímo v SQL, nebo musí být do tohoto jazyka přeložen)
* Dotaz je odeslán na server
* Server vykoná dotaz
* Výsledek dotazu je poslán zpět na vysílací počítač, kde je zobrazen

### Distribuované

Množina databází, která je uložena na několika počítačích. Uživateli se však jeví jako jedna velká databáze**.**

* **Funkční**
  + Vertikální členění
* **Objektové**
  + Horizontální členění

## Návrh DB

* Určení účelu DB
* Vyhledání a uspořádání požadovaných informací
* Rozdělení informací do tabulek
* Převod jednotlivých informací do sloupců
* Zadání primárních klíčů
* Vytvoření relací mezi tabulkami
* Úprava návrhu
* Použití normalizačních pravidel
  + Sémantické modelování – analyzuje požadavky a zobrazuje tyto požadavky určitými grafickými prostředky
  + Entitně –relační modelování (E/R diagram)

## E/R Diagram

* Entita
  + Subjekt, o němž se bude v databázi uchovávat informace
* Relace
  + Propojení tabulek
* Tabulky
  + Slouží k uložení dat
* Sloupce, atributy
  + Popisuje určitou část dat, kterou má každý záznam
  + Sloupec představuje část tabulky
  + Atribut se vztahuje k reálné entitě
* Domény
  + Popisují typ dat, obor hodnot
* Řádky, záznamy, n-tice
  + Každý řádek v tabulce představuje záznam o jedné entitě

## Klíče

Databázová konstrukce, sloužící ke zrychlení vyhledávacích a dotazovacích procesů v databázi, definování unikátní hodnoty sloupce tabulky

* **Primární klíč**
  + Svou hodnotou jednoznačně identifikuje každý záznam
* **Unikátní klíč**
  + Nemusí být jediný
* **Cizí klíč**
  + Odkaz mezi tabulkami

## Kardinalita

Vyjadřuje, kolik entit jednoho typu může být ve vztahu s kolika entitami z druhého typu entit

* 1:1
  + Používá se, pokud záznamu odpovídá právě jeden záznam v jiné tabulce
* 1:N
  + Přiřazuje jednomu záznamu více záznamů z jiné tabulky
* N: M
  + Umožňuje několika záznamům z jedné tabulky přiřadit několik záznamů z tabulky druhé
  + V praxi se spíše používá 1:N a M:1 pomocí jedné propojovací tabulky

## Relační Algebra

* Základní prostředek pro manipulaci s daty
* Teoretický základ dotazovacích jazyků (SQL,  LINQ, DMX, MDX, Datalog)
* Je dána operátory, které se aplikují na relace a výsledkem jsou opět další relace
* R({A1, A2, … An}); S({B1, B2, … Bm})

### Základní operace

* Sjednocení (Union) R
  + Vytvoření relace obsahující všechny řádky (prvky) obou relací, ale shodné řádky se neopakují
  + Relace S, R musí být kompatibilní
    - Mají stejný počet atributů, v některých případech musí mít atributy stejný název a datový typ
* Průnik (Intersection)
  + Vytvoření relace obsahující společné řádky obou relací, ale společné řádky se neopakují
  + Relace S, R musí být kompatibilní
* Rozdíl (Diference)
  + Vytvoření relace obsahující jen ty řádky první relace, které nejsou obsaženy v druhé relaci
  + Relace S, R musí být kompatibilní
* Kartézský součin (Cartesian product)
  + Vytváří relaci obsahující všechny řádky první relační tabulky zřetězené postupně se všemi řádky druhé relační tabulky

### Speciální operace

* Projekce (Projection)
  + Projekce R[C] na relaci se schématem R({A1, A2, … An }) na množinu C, kde C je menší, nebo rovno množině { A1, A2, … An }
  + Odstraní se i duplicitní řádky
* Selekce, Restrikce (Selection, Restriction)
  + Je relace se schématem R({A1, A2, … An}) podle logické podmínky ϕ (ϕ je jednoznačně true/false)
  + Projekce a selekce jsou operace s jednou relací – unární operace
* Spojení (Join)
  + Slouží pro spojení množin na základě společných prvků zvoleného atributu
  + Natural join
    - Podmínka je určována automaticky, ne často se používá
  + Inner join
    - Kartézský součin
  + Full outer join
    - Stejné jako inner join
  + Left outer join
    - Výsledek uzná, pokud existuje levá část vazby a pravá neexistuje
    - Do hodnot sloupců z připojované části se vloží NULL
  + Right outer join
    - Pokud bude existovat pravá, připojovaná část a nebude k ní levá část, bude stejně ve výpisu zahrnut

## Relační kalkul

* Formální neprocedurální jazyk
* N-ticově a doménově orientovaný

### Zápis:

* Termy
  + Proměnné
  + Jejich komponenty
  + Konstanty
* Predikáty
  + >
  + <
  + >=
  + <=
  + <>
  + =
* Atomické formule
  + Konjunkce – &
  + Disjunkce - ∨
  + Negace - ¬
  + Implikace - ⇒
  + Ekvivalence - ⇔
* Kvantifikátory
  + Univerzální (∀) – „pro každý “
  + Existenční (∃) - „existuje