

# KMA/PDB

# Prostorové databáze

Karel Janečka

*Tvorba materiálů byla podpořena z prostředků projektu FRVŠ č. F0584/2011/F1d*



# Sylabus předmětu KMA/PDB

- **Úvodní přednáška**
  - Základní terminologie
  - Motivace – rozdíl klasické DB x prostorové DB
  - Rastrová data & databáze
  - Historie ukládání (vektorových) prostorových dat v relačních databázích
- **Prostorové objekty a SQL**
  - Geometrický model OGC
  - SQL:1999 a Abstraktní datové typy
  - Formáty WKB a WKT



# Sylabus předmětu KMA/PDB

- **Prostorové datové struktury pro indexaci prostorových objektů**

- Rozdělení prostorových datových struktur podle pojetí výchozího prostoru
- Linearizace prostoru
- Indexace bodů a obecných prostorových objektů

- **Algoritmy prostorových databází**

- Algoritmy pro vyhodnocení prostorového spojení
- Topologická konzistence prostorových dat
- Víceuživatelská editace prostorových dat



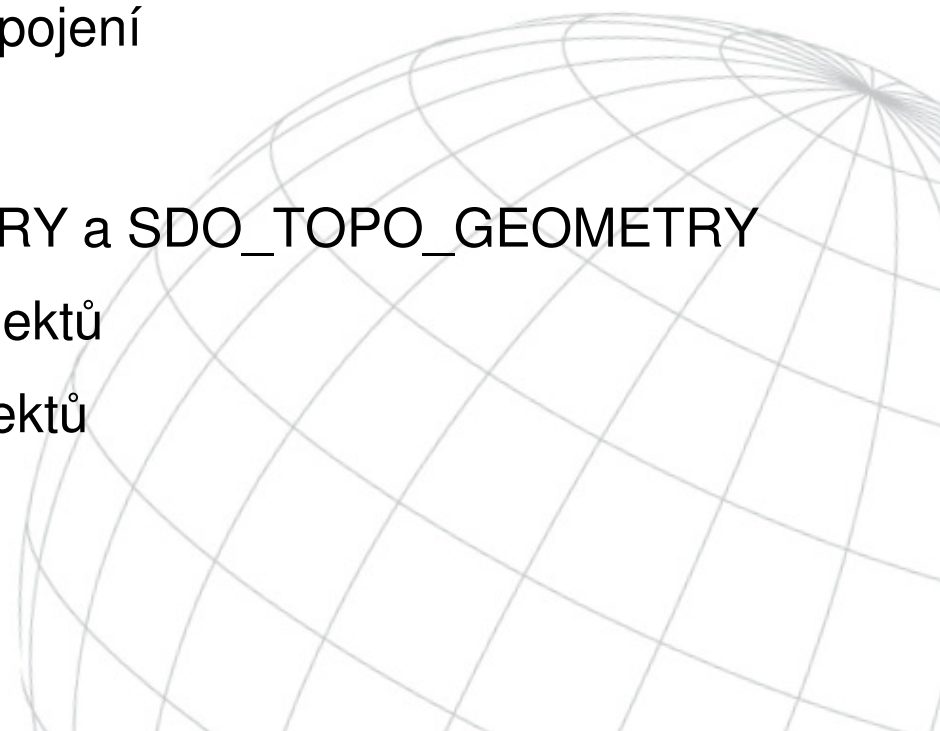
# Sylabus předmětu KMA/PDB

- **Prostorové spojení**

- Prostorové spojení pomocí vnořených cyklů
- Prostorové spojení pomocí dynamických prostorových datových struktur
- Optimalizace prostorového spojení

- **Oracle Spatial**

- Datové typy SDO\_GEOMETRY a SDO\_TOPO\_GEOMETRY
- Modelování prostorových objektů
- Indexování prostorových objektů
- Prostorové dotazy
- Praktická ukázka



# Sylabus předmětu KMA/PDB

- **PostGIS**

- Základní informace o PostgreSQL a PostGIS
- Prostorové datové typy a funkce PostGIS
- Praktická ukázka:
  - import dat do databáze
  - připojení databáze do klientských aplikací
  - prostorové dotazy a indexace
  - publikování dat z databáze v prostředí internetu



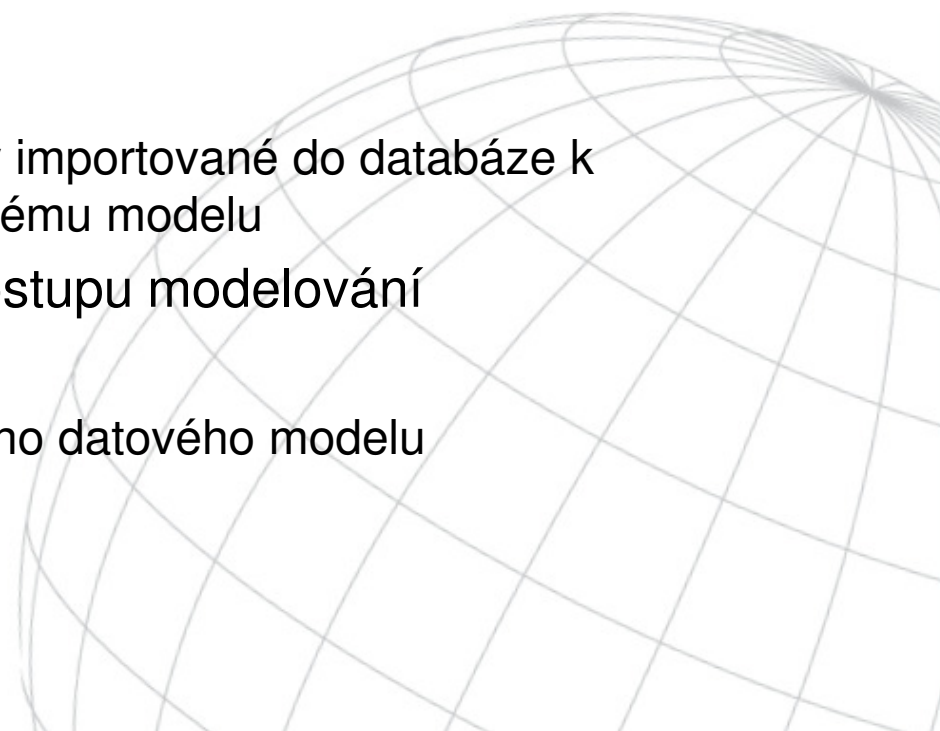
# Sylabus předmětu KMA/PDB

- **ESRI Geodatabase**
  - ArcSDE Binary
  - Struktura (datové sady a třídy)
  - Typy (personal, file, ArcSDE)
  - Cvičení:
    - založení geodatabáze
    - import dat
    - práce se strukturou databáze



# Sylabus předmětu KMA/PDB

- **Modelování v prostorových databázích**
  - Obecný návrh (informačního) systému
  - UML
  - Datové modelování
  - Názvové konvence
  - Cvičení:
    - Přejít od ploché struktury importované do databáze k hierarchicky hlubšímu datovému modelu
  - Nástroje pro dokumentaci postupu modelování
  - Cvičení:
    - Návrh a dokumentace nového datového modelu



# Základní terminologie





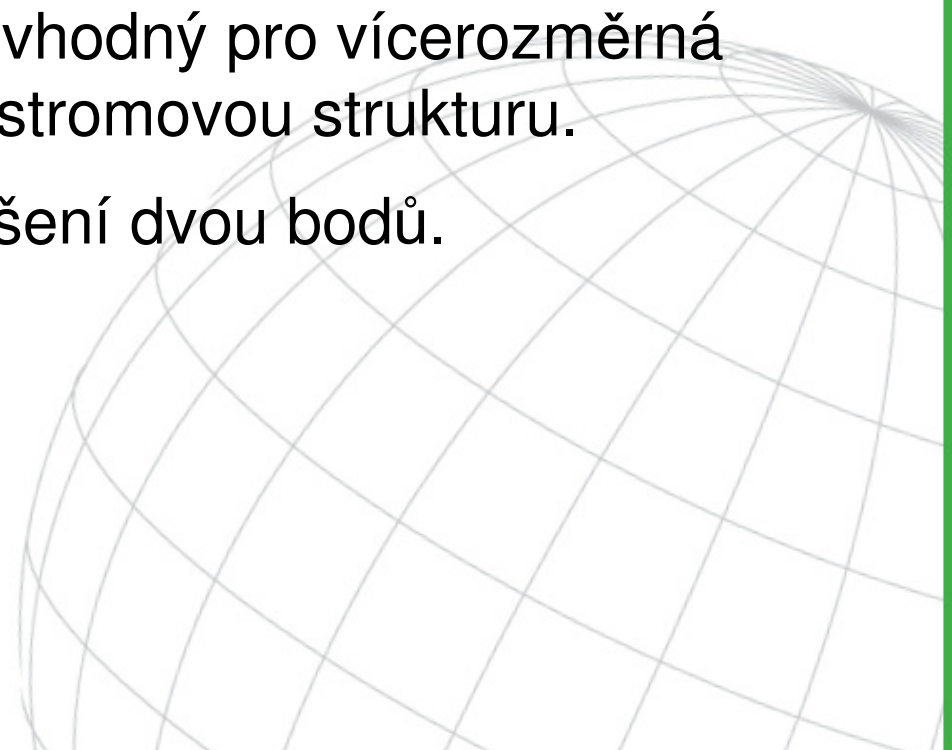
# Vybrané pojmy z oblasti prostorových databází

- **Prostorový datový typ** – datový typ vhodný pro uložení vícerozměrných dat; založeno na mechanismu abstraktních datových typů.
- **Prostorový atribut** – atribut obsahující geometrická data.
- **(Binární) Prostorový predikát** – vztah mezi dvěma prostorovými atributy. Nabývá booleovské hodnoty TRUE/FALSE.



# Vybrané pojmy z oblasti prostorových databází

- **Prostorová relace** – relace s prostorovým atributem. Prostorové relace mají jiné vlastnosti než konvenční relace, neboť odkazují na topologické a metrické vztahy příslušných objektů.
- **Prostorový index** – index vhodný pro vícerozměrná data; nejčastěji se jedná o stromovou strukturu.
- **Tolerance** – přesnost rozlišení dvou bodů.



# Klasické DB x Prostorové DB



# Požadavky na prostorovou databázi

- Aby bylo možné ukládat prostorová data a provádět nad nimi požadované operace, musí databáze podporovat:
  - **prostorové datové typy** (např. bod, lomená úsečka, region apod.)
  - **operace**, kde takové datové typy vstupují jako parametry, případně vystupují jako výsledek. Tyto operace se dělí na:
    - **prostorové predikáty** (např. je\_uvnitř())
    - **prostorové relace** (např. průnik())
    - **ostatní operace** (např. plocha())
  - **prostorové indexování** a efektivní algoritmy pro **prostorové spojení**
  - Dále musí databázový systém zajistit funkci již existujících operací (např. projekce, selekce).

# Rastrová data & databáze



- Systémy pro správu rastrů dříve pracovaly s rastry uloženými v nějakém proprietárním formátu, případně s rastry uloženými v proprietární rastrové “databázi” (např. Mr.SID – *Multi-resolution Seamless Image Database*)
- V současné době lze rastrová data uchovávat i v SŘBD
- Příklad:
  - Oracle Spatial
  - ESRI Geodatabase

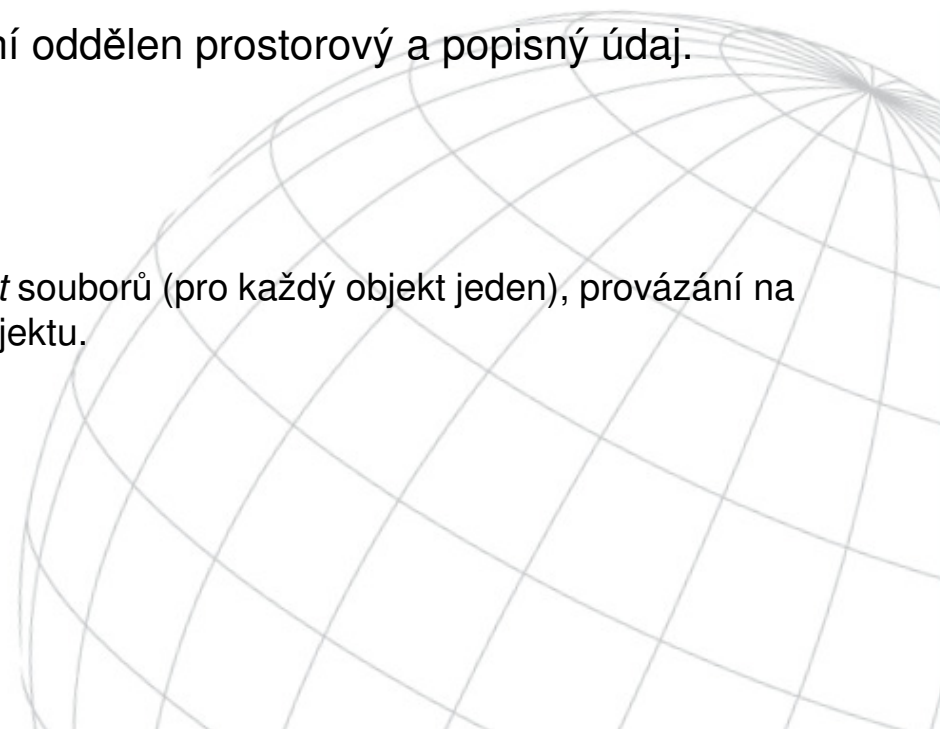


# Historie ukládání prostorových dat v relačních databázích



# Vývoj GIS podle přístupu k propojení atributových a prostorových dat – 1. generace

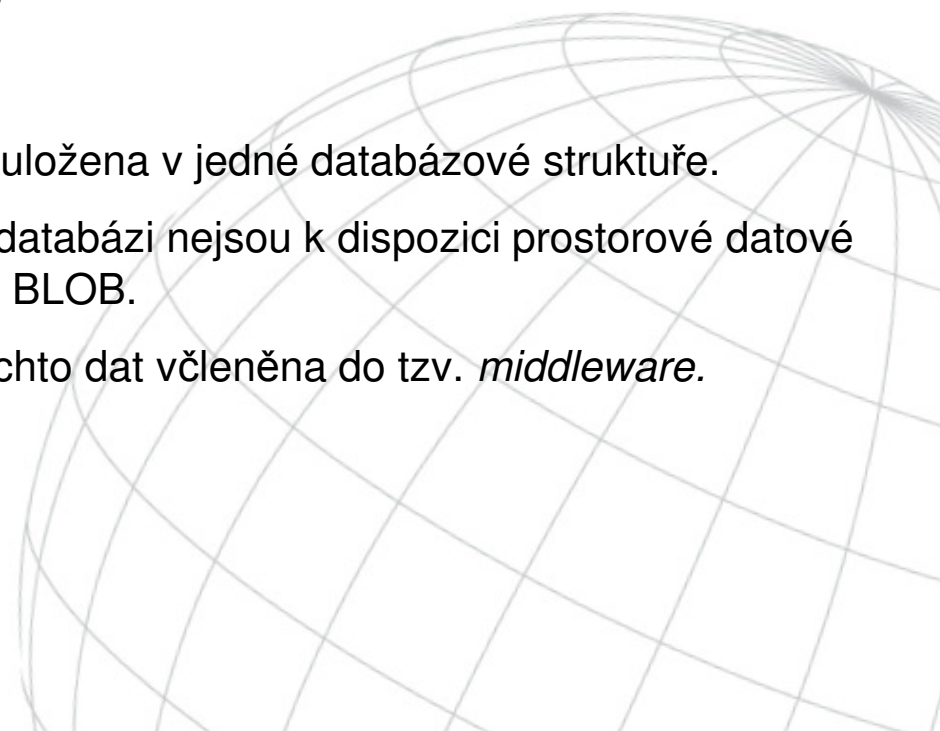
- V první generaci jsou prostorové objekty a jejich atributy uloženy v souborech.
- Dvě skupiny systémů:
  - **Systém bez atributových souborů**
    - používán pro rastrová data, není oddělen prostorový a popisný údaj.
    - 1 rastr = 1 soubor = 1 *atribut*
  - **Systém s *flat* soubory**
    - Atributová data byla ukládána do *flat* souborů (pro každý objekt jeden), provázání na prostorová data přes identifikátor objektu.





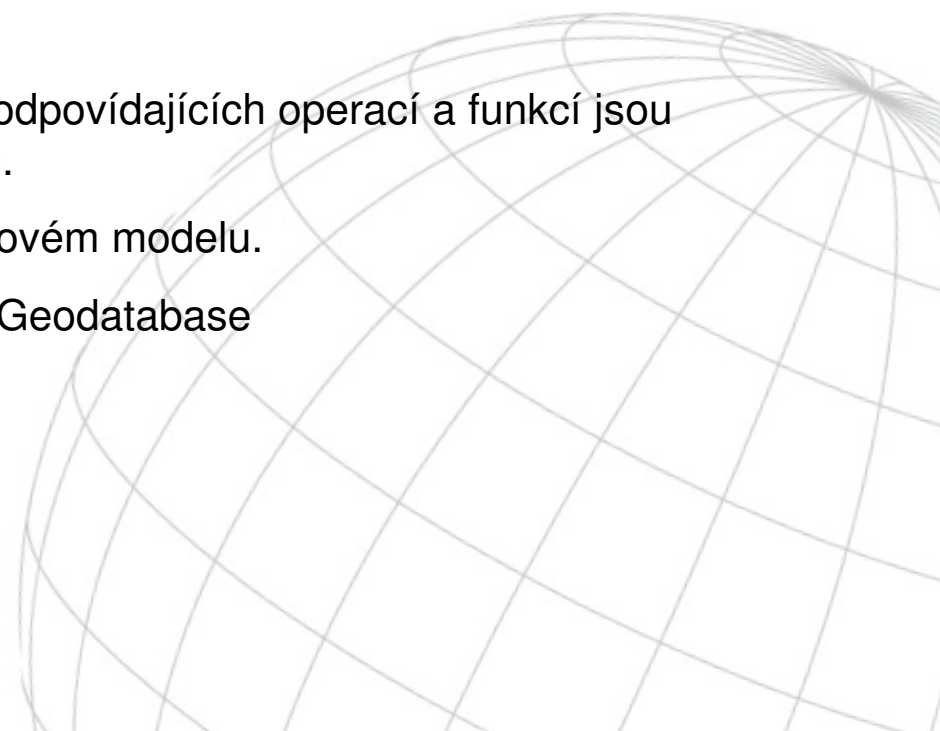
# Vývoj GIS podle přístupu k propojení atributových a prostorových dat – 2. generace

- V druhé generaci můžeme rozlišit dva typy přístupů:
  - **Duální/hybridní**
    - Neprostorová data jsou uložena v relační databázi, prostorové objekty jsou uloženy v souborovém systému.
    - Např. ARC/INFO (dnes ArcInfo)
  - **Integrovaný**
    - Prostorová i neprostorová data uložena v jedné databázové struktuře.
    - Pro uložení prostorových dat v databázi nejsou k dispozici prostorové datové typy, data jsou tak uložena jako BLOB.
    - Funkcionalita pro manipulaci těchto dat včleněna do tzv. *middleware*.
    - Např. ArcSDE



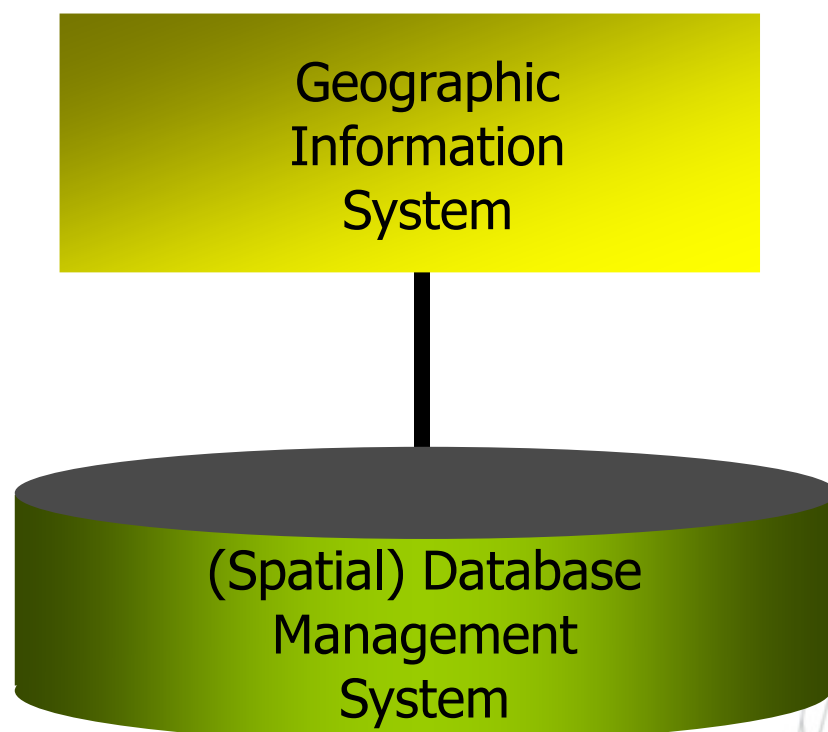
# Vývoj GIS podle přístupu k propojení atributových a prostorových dat – 3. generace

- Opět můžeme rozlišit dva typy přístupů:
  - **Objektový**
    - Jedné reálné entitě odpovídá jeden databázový objekt.
    - Data (prostorová + neprostorová) uložena spolu s metodami objektu.
  - **Objektově-relační**
    - Prostorové datové typy včetně odpovídajících operací a funkcí jsou integrovány do relačního SŘBD.
    - Podpora objektů v relačním datovém modelu.
    - Např. Oracle Spatial, PostGIS, Geodatabase



# Vývoj GIS podle přístupu k propojení atributových a prostorových dat – 3. generace

## Technologie



## Úlohy

- Pořizování dat
  - Vizualizace (Mapy)
  - Pokročilá analýza
- 
- Úložení a správa
  - Indexování
  - Dotazování
  - Verzování
  - Bezpečnost, správa uživatelů
  - Škálování výkonu
  - Vysoká dostupnost

# Přístupy pro uložení prostorových dat v relační databázi

- Prostorová data jsou uložena v tabulkách v 1 NF
- Prostorová data jsou uložena jako nestrukturované rozsáhlé binární objekty (BLOB) či WKB
- Prostorová data jsou uložena v relační databázi rozšířené o typy prostorových dat
- Prostorová data jsou implementovaná pomocí uživatelsky definovaných typů



# Prostorová data uložena v tabulkách v 1. NF

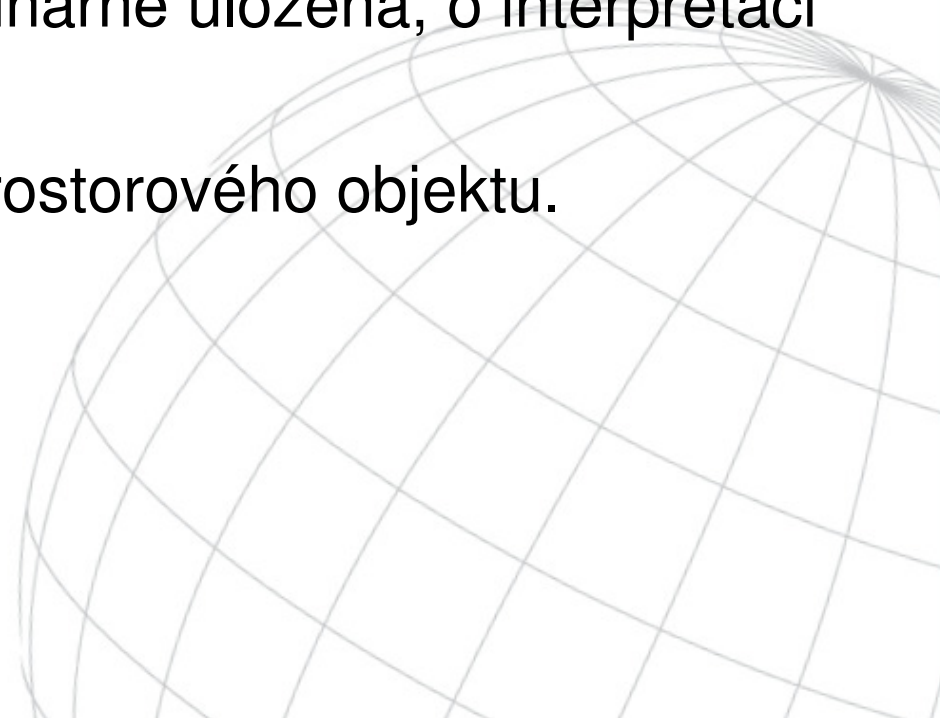
- *(připomenutí)* Relace je v první normální formě, pokud každý její atribut obsahuje jen atomické hodnoty. Tedy hodnoty z pohledu databáze již dále nedělitelné.
- **Je nutné prostorové objekty rozložit na primitivní objekty a ty ukládat pomocí souřadnic a čísel.**
- Příklad: Uložení čtyřstěnu v relační databázi
- Čtyřstěn je dán pomocí rovin, hran a bodů.
- Odpovídající data se uloží ve čtyřech relacích: ČTYŘSTĚNY, ROVINY, HRANY a BODY.

# Prostorová data uložena v tabulkách v 1. NF

- Relace ČTYŘSTĚNY obsahuje dvojice  $(t, f)$ , kde  $t$  je identifikátor čtyřstěnu,  $f$  je identifikátor jedné z jeho rovin.
- Relace ROVINY obsahuje dvojice  $(f, e)$ , kde  $f$  je identifikátor roviny a  $e$  je jedna z hran v té rovině.
- Relace HRANY obsahuje trojice  $(e, p, q)$ , kde  $e$  je identifikátor hrany,  $p$  a  $q$  jsou koncové body hrany.
- Relace BODY obsahuje čtveřice  $(p, x, y, z)$ , kde  $p$  je identifikátor bodu a  $x, y, z$  jeho souřadnice.
- Je zřejmé, že tato reprezentace geometrie nepodporuje příliš zpracování prostorových dotazů.

# Prostorová data uložena jako BLOB

- (připomenutí) BLOB (*Binary Large Object*) je označení pro datový typ blíže nespecifikovaných binárních dat v databázi
- Obsah BLOB položky obdrží klientská aplikace přesně tak, jak byla tato položka binárně uložena, o interpretaci se musí postarat sama.
- Nelze zacházet s částmi prostorového objektu.



# Prostorová data uložena v relační databázi rozšířené o typy prostorových dat

- Rozšíření relační databáze o prostorové typy dat znamená přidat nové struktury (implementované pomocí tabulek) a odpovídající operace
- Podpora indexových struktur





# Prostorová data uložena pomocí UDT

- Dříve (SQL3) se používal pojem **Abstraktní datový typ** (ADT), který zapouzdřuje data a funkce ve stejném smyslu jako u objektově orientovaného přístupu.
- SQL:1999 specifikuje SQL:1999 objektový model, který do SQL přináší **Uživatелеm definované typy** (UDT)
- Funkce (Uživatелеm definované funkce – UDF) modelující chování objektů jsou vnořeny do SŘBD, což umožňuje lepší optimalizaci dotazů.
- Mechanismus vytváření UDT a UDF je součástí konstruktorů jazyka (v daném případě SQL).
- **ADT  $\approx$  UDT + UDF**

# Zdroje

- JANEČKA, K.: [Zajištění konzistence prostorových dat v Informačním systému katastru nemovitostí](#). In: Proceedings of GIS Ostrava 2008. Tanger. Ostrava, 2008. s. 1-8. ISBN 978-80-254-1340-1.
- KOLINGEROVÁ, I.: [Přednášky k předmětu Vybrané algoritmické metody](#). FAV ZČU v Plzni.
- MURRAY, Ch.: [Oracle Spatial Developer's Guide, 11g Release 1 \(11.1\)](#). Oracle. 2009.
- POKORNÝ, J.: [Prostorové datové struktury a jejich použití pro indexaci prostorových objektů](#). In: Proceeding of GIS Ostrava 2000. Ostrava, 2000.
- POKORNÝ, J.: [Prostorové objekty a SQL](#). In. Proceedings of GIS Ostrava 2001. Ostrava, 2001. ISSN: 1213-239X.
- POKORNÝ, J.; ŽEMLIČKA, M.: [Základy implementace souborů a databází](#). Karolinum. Praha, 2004. 211 s. ISBN: 80-246-0837-5.
- ŽEMLIČKA, M.: [Přednášky k předmětu Organizace a zpracování dat II](#). MFF UK v Praze.

