

Geografické informační systémy

Slajdy pro předmět GIS

Martin Hrubý

hrubym @ fit.vutbr.cz

Vysoké učení technické v Brně
Fakulta informačních technologií,
Božetěchova 2, 61266 Brno

—

akademický rok 2004/05

Databázové systémy pro GIS

historie, současnost

Úvod

- Struktura a funkce databázových systémů/uložení dat v GIS.
- Různé varianty uložení. Atributové údaje.
- V předchozí přednášce se probíraly datové modely uložení vektorových a rastrových dat.

Přehled

- souborové systémy
- soubory + databáze
- databáze

Souborové systémy

- Uložení prostorových a atributových dat v souborech.
- GIS obsahuje sadu funkcí pro manipulaci s daty v souborech.
- Problém při víceuživatelském přístupu. Nebezpečí poškození dat (integrita, nezávislost, bezpečnost)
- GRASS: soubory a lokace.
- Řešení je ve využívání databázového přístupu ke zpracování údajů.
- Báze dat + systém řízení báze dat - database management system (DBMS)

Položka, záznam, soubor, databáze, ...

- Entitně-relační přístup - modelování vztahů mezi objekty.
- K ER a objektu se pojí množina atributových údajů.
- Databázový soubor - množina záznamů (datový záznam) stejného druhu
- Víc souborů pro jednu aplikaci (oblast zájmu) - databáze.

Příklad: každá meteorologická stanice je definována svou polohou, včetně nadmořské výšky. Záznam obsahuje: x-souřadnice, y-souřadnice, výška (Sliač: 450000,1220000,320).

Organizace dat (schéma dat)

Další vrstva abstrakce.

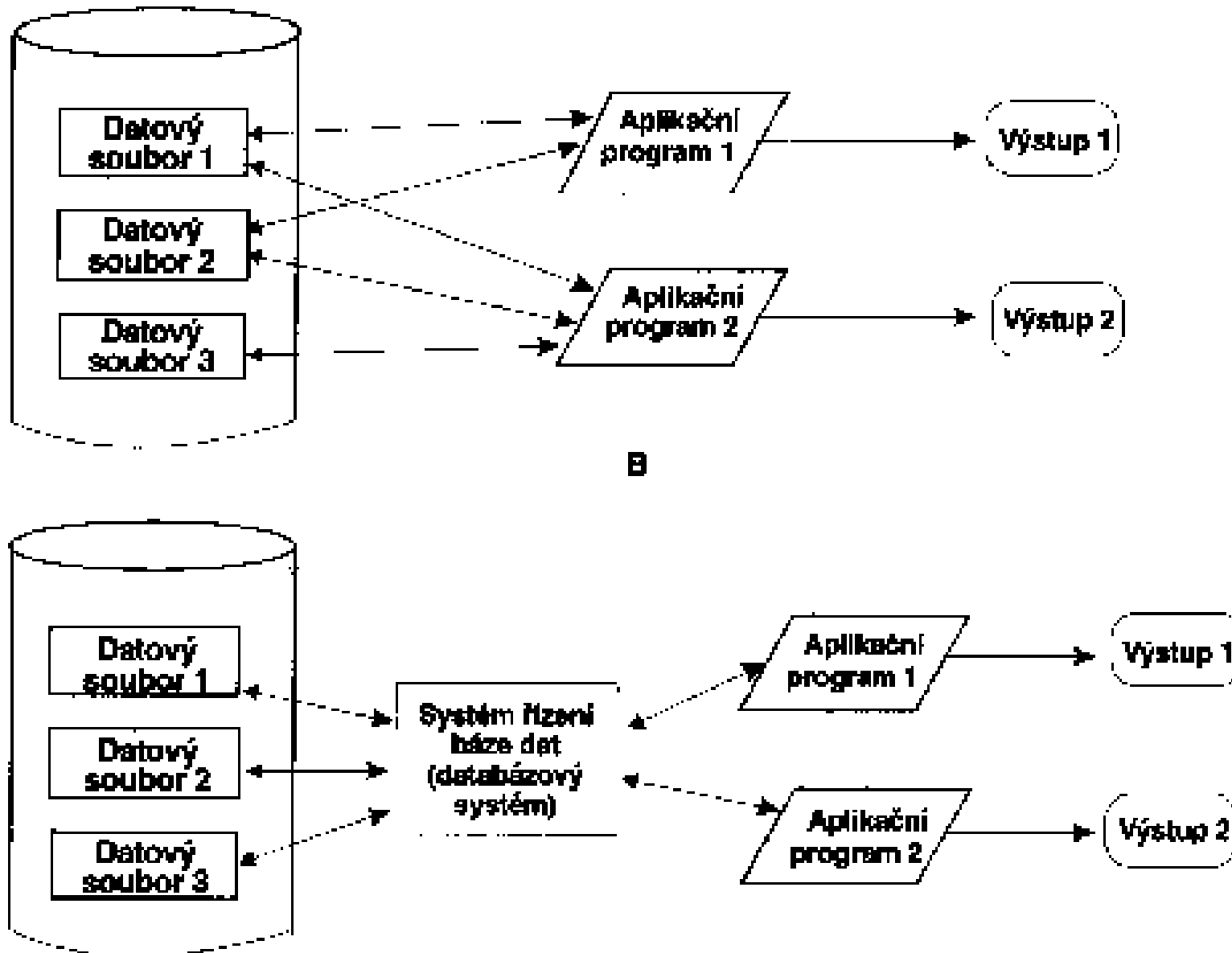
- Interní - konkrétní uložení dat na médiích (disk).
Optimalizace na efektivnost přístupu.
- Externí - uživatelský pohled na data - virtuální pohled.
- Konceptuální - vymezuje celkovou logickou strukturu dat, jejich vlastnosti a vnitřní vazby. Měl by to být nejobecnější pohled na realitu (strukturu dat), který by se neměl měnit příliš rychle.

SŘBD

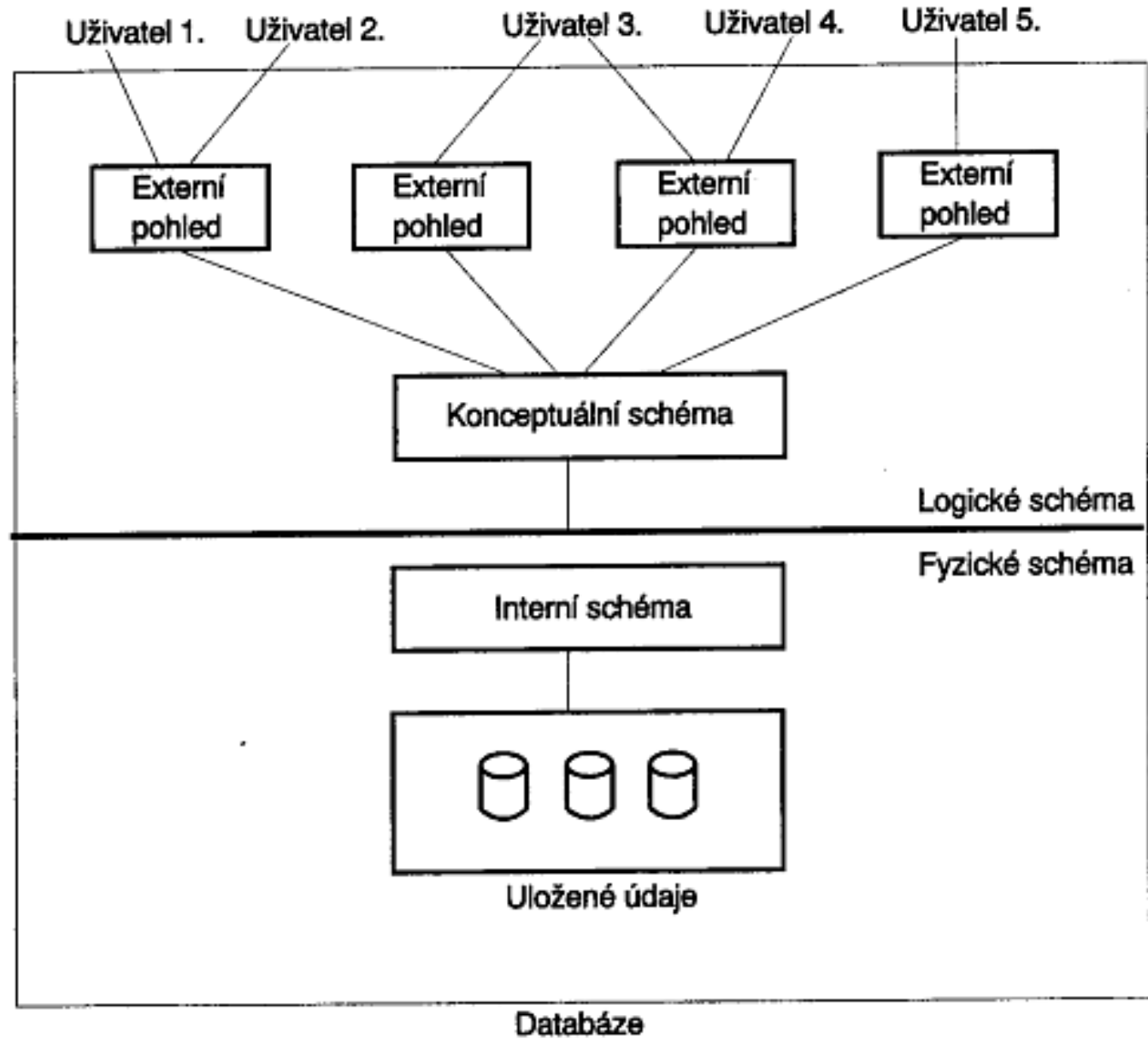
Vlastnosti:

- Přístup k datům pro všechny aplikace bez vícenásobného ukládání (redundance)
- Současný přístup více uživatelů
- Různé vyhledávací metody
- Ochrana dat - neoprávněný přístup, chyby HW a SW
- Prostředky pro centrální správu dat
- Nezávislost aplikací na datech
- Možnost vytvářet i složité datové struktury
- Ukrytí mechanismu struktur a ukládání dat

SŘBD



Vrstvy, schemata



Poznámky

- Jazyk pro manipulaci s daty (např. SQL)
- Architektury klient-server
- Distribuované db systémy
- Role uživatelů - správce dat, aplikační programátoři, znalí/pasivní uživatelé

Databázové modely

- (Hierarchický, síťový) - zastaralé - omezená forma vazeb, prohledávání
- Relační - operace nad relacemi, klíče
- (známé z předchozích předmětů)
- Objektově orientované modely

OO model

Gemstone, Smalltalk, Self.

- Každá entita je modelována jako objekt s vlastní identitou. Tato identita je poskytována OODBMS a je trvalá po celou dobu existence objektu.
- Každý objekt je zapouzdřený, má vlastní strukturu a vlastní chování.
- Objekty komunikují mezi sebou zasíláním zpráv.
- Objekty se stejnou strukturou a chováním jsou popsány jako třída objektů.
- Odvozování tříd

Dotazovací/řídící jazyky

- Databáze má svoji strukturu - uživatel ji zná/nezná.
- Procedurální dotazovací jazyky - jazyky, které vyžadují, aby uživatel poznal hierarchii v DB - váže se na ty "zastaralé" db modely.
- Neprocedurální (SQL), relační db - velká možnost (volnost) ve tvorbě dotazů.
- geoSQL - projekty

GIS architektury

Generace GIS

- komputerizované mapovací systémy s hlavním cílem v produkci map. Soubory, žádné atributové údaje.
- s atributovými údaji, neexistují standardy uložení
- uložení v relačních databázích
- objektově orientované přístupy

Postrelační - rozšířené relační přístupy, OO.

První generace GIS

Systemy bez atributových souborů.

- Čistě rastrový přístup neodděluje polohové a atributové informace. Rastr=jeden soubor.
- Nepřítomnost DB systému v rastrovém přístupu není omezující, naopak - uložení je jednodušší...
- Problém nastává při kombinování více rastrových vrstev - mnoho souborů
- Flat soubory - jednoduché tabulky údajů. Všechny mají stejnou strukturu. Individuální tabulky v relačním modelu se mohou nazývat flat soubory.
 - prostorové údaje se nemusí opakovat
 - soubor s polohami + sada souborů s atributy, ...

Druhá generace GIS, duální systémy

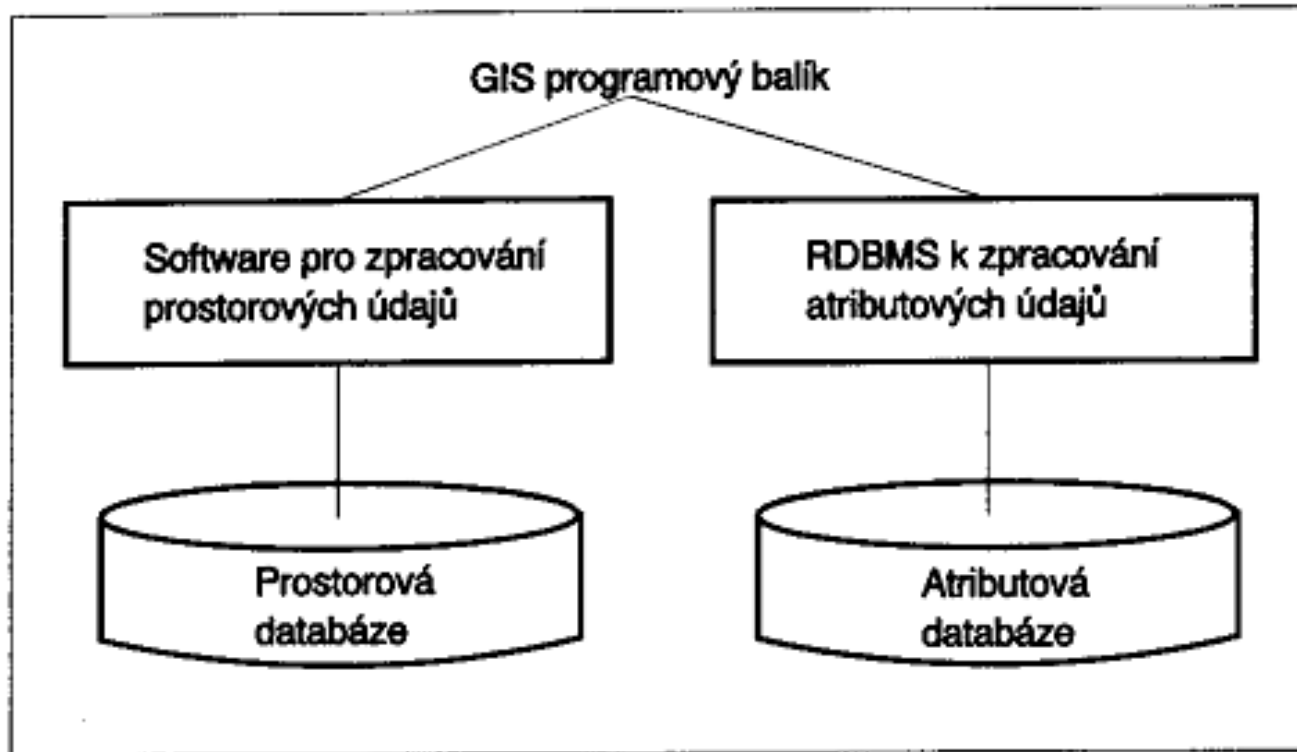
Relační přístup. V podstatě nejrozšířenější přístup (Arc/Info, Systém 9, Geo/SQL).

V tabulkách se ukládají: a) pouze atributové údaje, b) veškerá data

- Arc/Info je typický příklad GIS systému, který ukládá do tabulek pouze atributové údaje (duální architektura).
- GIS systémy implementují rozhraní na různé DBMS nebo mají vlastní (ArcSDE). Případně DB podporu již obsahují.
- Vektorový objekt/buňka rastru je identifikována číslem - klíč do databázové tabulky.
- Pak se nabízí dvojí použití spojení polohová/atributová data:
 - k objektům dodávají popisky (jména řek...)
 - systém zobrazuje objekty na základě DB dotazu (zobraz města s počtem obyvatel...)

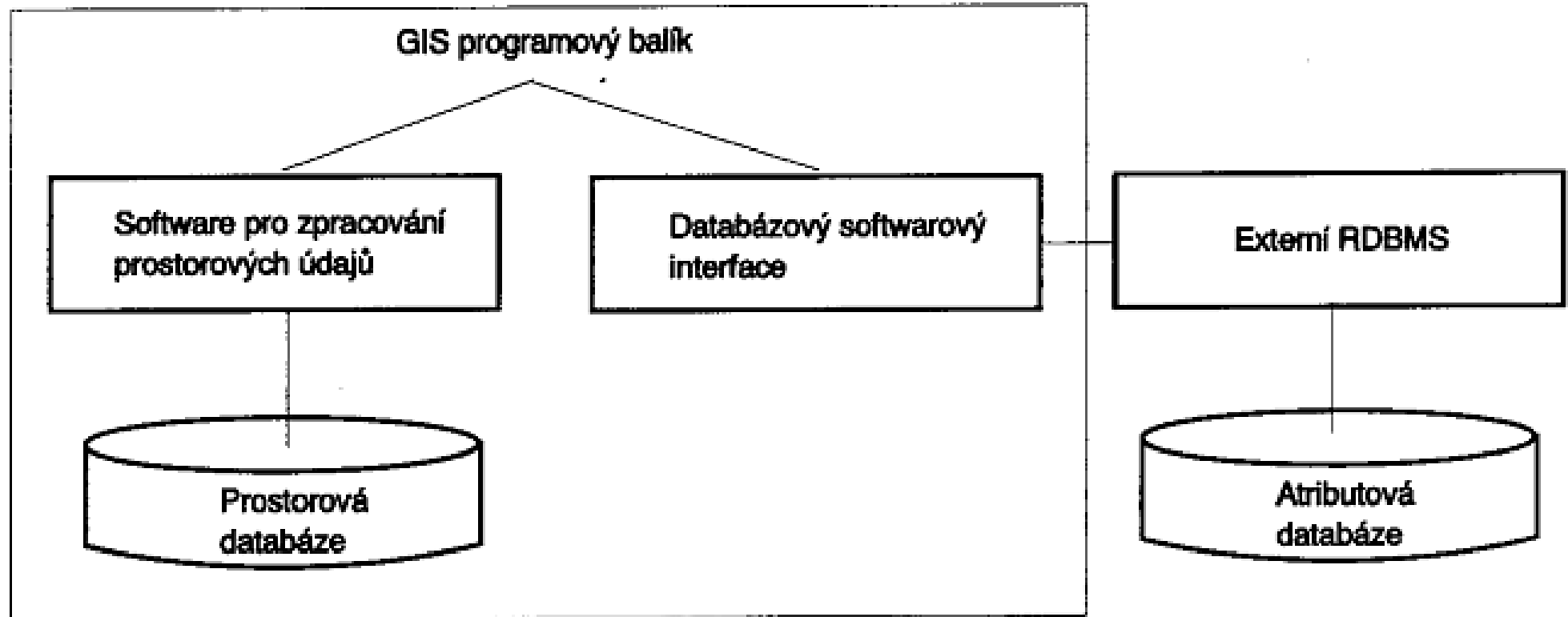
Duální s., DBMS implementovaný do GIS

a) DBMS implementovaný do GIS

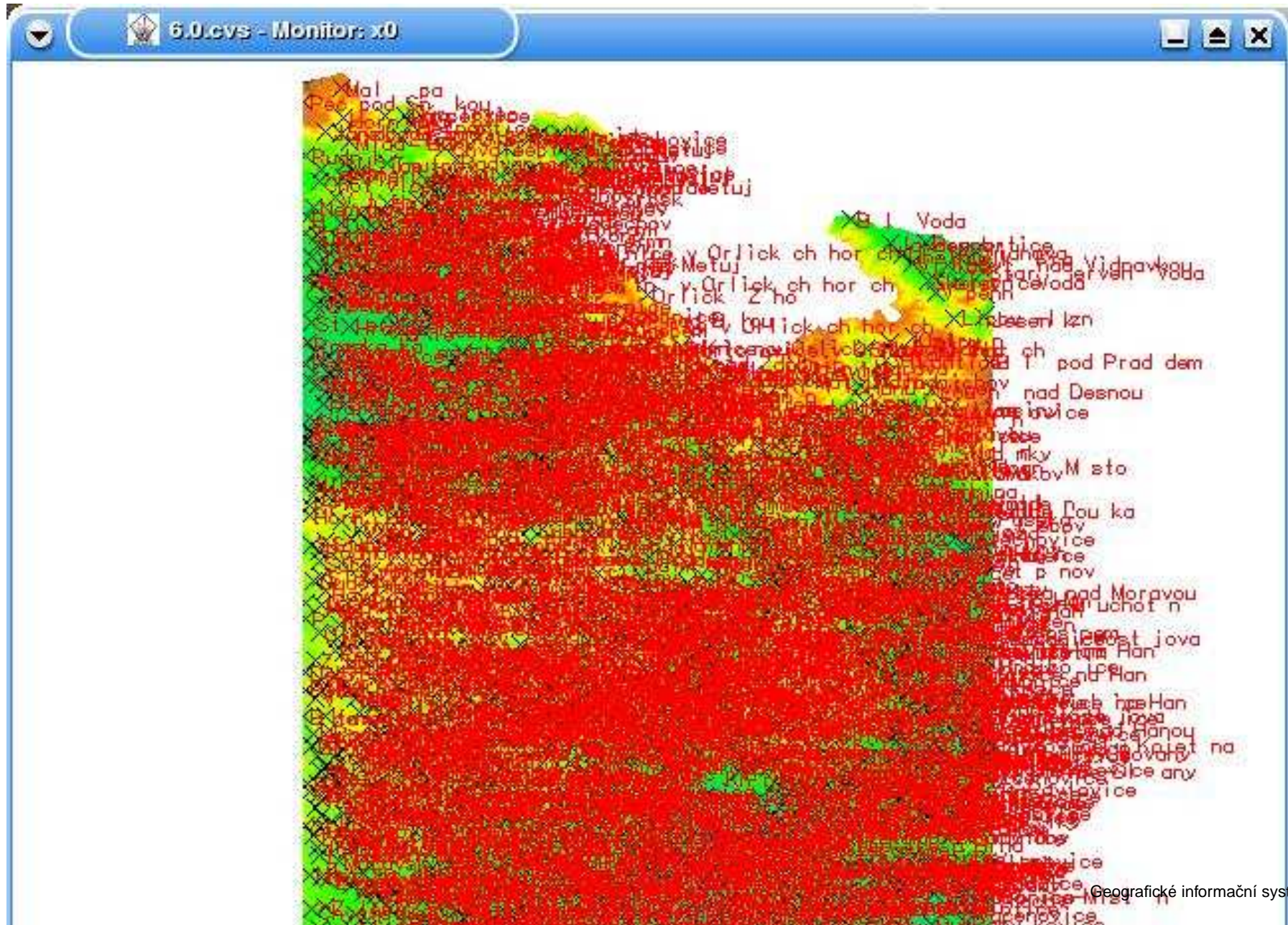


Duální systémy, externí DBMS

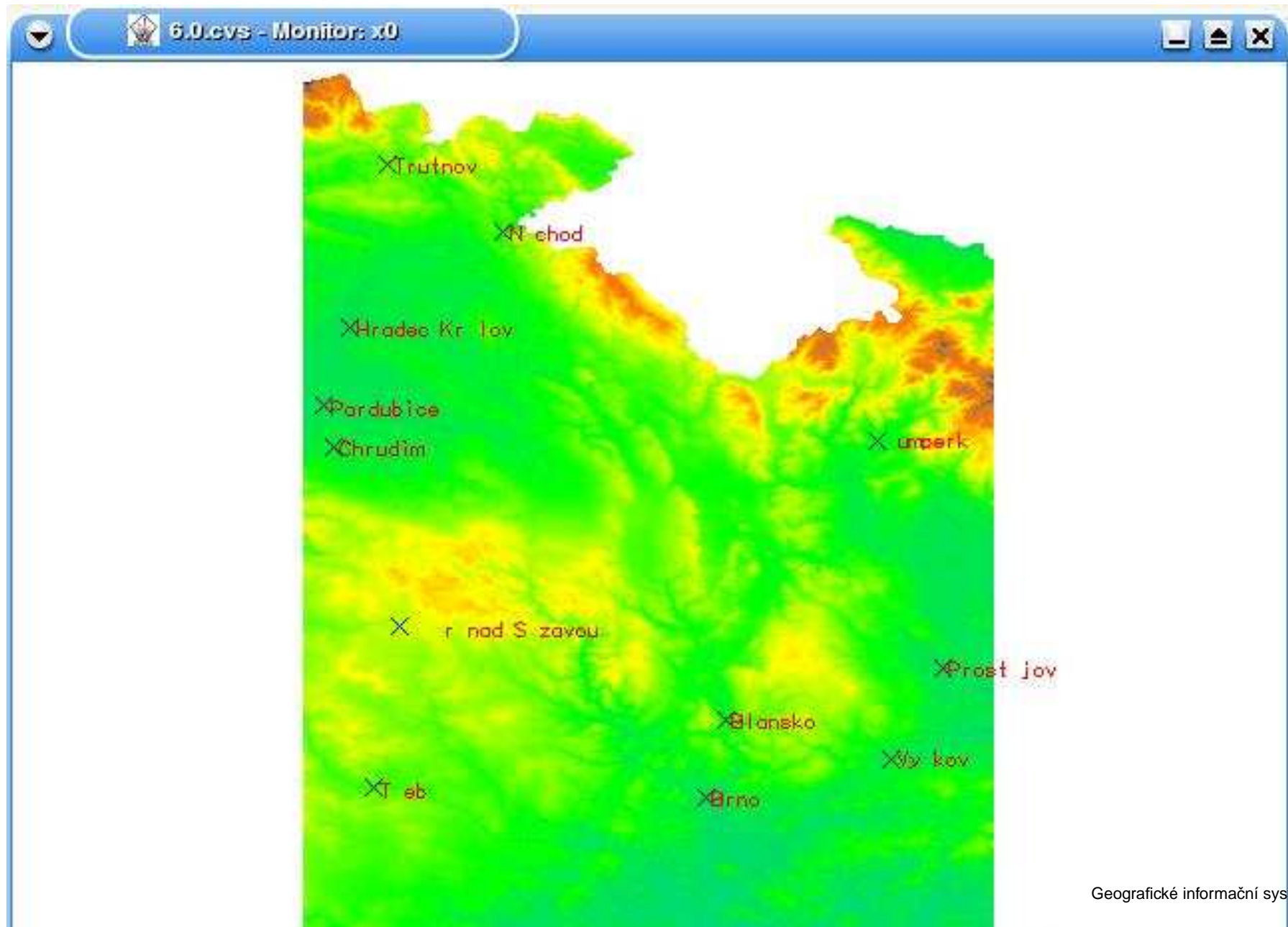
b) Externí DBMS



Vektory+Atributy



Vektory+Atributy



II. generace GIS, integrované systémy

- integrují prostorové a atributové údaje do stejné db struktury (System 9)
- problém - nestejně dlouhé řetězce liniových objektů - rozdělit do více tabulek
- (implementovat uložení vektorového formátu v DB)
- integrita - robustnost profesionálního produktu DBMS
- formování dotazů - problém pouze s formováním dotazů pro prostorové údaje
- (geoSQL)
- ulehčování vytváření distribuovaných databází (GIS systémy beztak směřují k propojování, Internetu, ...)

Třetí generace GIS

Postrelační...

- Relační datový model je poměrně omezený.
- Novinky: (ideální požadavky na rozšíření)
- možnost používat záznamy různé délky
- možnost definovat komplexní objekty
- možnost vyměňovat kompletní GIS databáze mezi různými GIS systémy
- podporuje rastrovou i vektorovou reprezentaci
- prostorové operátory jako například vzdálenost, plocha, obvod
- možnost prostorových indexových souborů

Třetí generace GIS

Objektové.....

- snad východisko...pro moderní nároky na informační systém
- snahou je dosáhnout vyšší úrovně abstrakce
- síla a současně slabost relačního přístupu je v jednoduchosti reprezentace reality
- komplexnost reality se pak rozpustí do množství tabulek. Pak je tu ještě hledisko teorie modelování - v tabulkách modelujeme elementy, které neexistují (identifikační číslo řeky...)
- objekty zahrnují i příslušné operace
- identita objektu je jiná než u tabulek (jedinečnost údajů). Objekt vyjadřuje fyzickou existenci věci.
- OO je pohodlnější pro zavádění nových datových typů

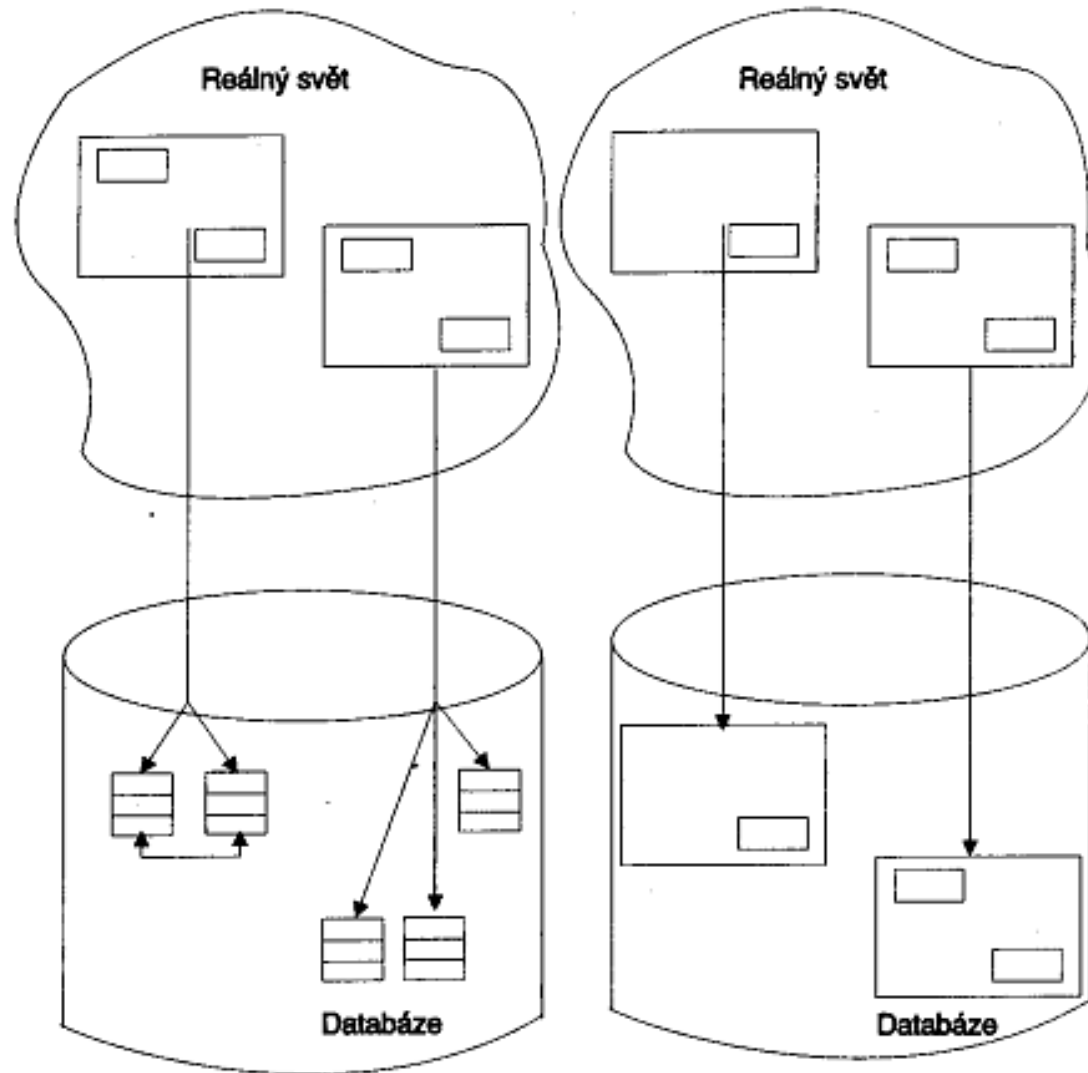
OO přístup, závěr

Důvody atraktivity OO přístupu pro GIS shrneme takto:

- V GIS pracujeme s komplexními objekty. OO přístup je vhodnější než relační.
- Neexistuje rozdíl mezi atributovým a prostorovým údajem. Variabilita délky záznamu (přínos!)
- Vhodnost pro grafické operace.
- Integrovaný příkazový jazyk.
- Alternativní geografická reprezentace. Polytématické vrstvy nejsou v OO problém.

Smallworld, Argus, Intergraph MG Dynamo

Relace, objekty



a) Relační model

Jeden objekt → Mnoho záznamů

b) Objektově orientovaný model

(Jedna (reálná) entita → Jeden (databázový) objekt

Vektory ve Squeak Smalltalku

<http://minnow.cc.gatech.edu/squeak/3317>

