



Úvod do Databázových Systemov

Ján Šturc

Matematicko-fyzikálna
fakulta UK Bratislava, zima
2005

Literatúra:

J.D. Ullman: *Database and knowledge-base systems* computer science press, 1988

J.D. Ullman: *Principles of database systems* computer science press, 1982

C.J. Date: *An introduction to database systems* addison-wesley, 1986

H.F. Korth, a. Silberschatz: *Database system concepts* mcgraw-hill, 1986

Česká a Slovenská literatúra:

J. Pokorný: *Databázové systémy a
jejich použití v informačních
systémech* akademia, 1992

A. Scheber: *Databázové systémy*
alfa, 1988

D.C. Tsichritzis, F.H. Lochovsky:
Database management systems
academic press, 1977
český překlad:
Databázové systémy
snti 1987

Najnovšia literatúra

Jeffrey D. Ullman and Jennifer Widom: *A First Course in Database Systems.*

Prentice Hall 1997

Kniha má svoju web stránku. Táto stránka obsahuje slidy, cvičenia, zadania na projekty a skúšky ku kurzu CS 145.

<http://www-db.stanford.edu/~ullman/fcdb.html>

Dôrazne odporúčané !

Dan A. Simovici and Richard L. Tenney: *Relational Database Systems*

Academic Press 1995

Ešte novšie materialy

Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom:
Database systems The Complete Book

Táto kniha by mala byť dostupná
v knižnici.

<http://www.stanford.edu/class/cs145/>

Na tejto stránke nájdete slidy a študijné
materialy k stanfordskému kurzu z leta
2005

Používané systémy

Sálové počítače a pracovní stanice

- * **Adabas**
- * **Db-2**
- * **Informix**
- * **Ingress**
- * **Magic**
- * **MS SQL**
- * **Oracle**
- * **Progress**
- * **Sybase**

Osobné počítače

- * **Access**
- * **Approach**
- * **Caché**
- * **Dbase**
- * **Enterprisedb 2005**
- * **Foxbase / foxpro**
- * **My SQL**
- * **Paradox**
- * **Postgres**
- * **602 SQL**

Konferencie a semináře

- * *Principles of database systems*
- * *International conference on very large data bases (VLDB)*
- * *International conference on database theory (IDBT / EDBT)*
- * *International conference on the management of data*
- * *International conference on foundations of data organization (FODO)*
- * *Datasem*

Časopisy

- * *ACM transaction on database systems*
- * *Acm sigmod*

JACM, SIAM journal on computing, CACM majú sekcie venované databázovej problematike.

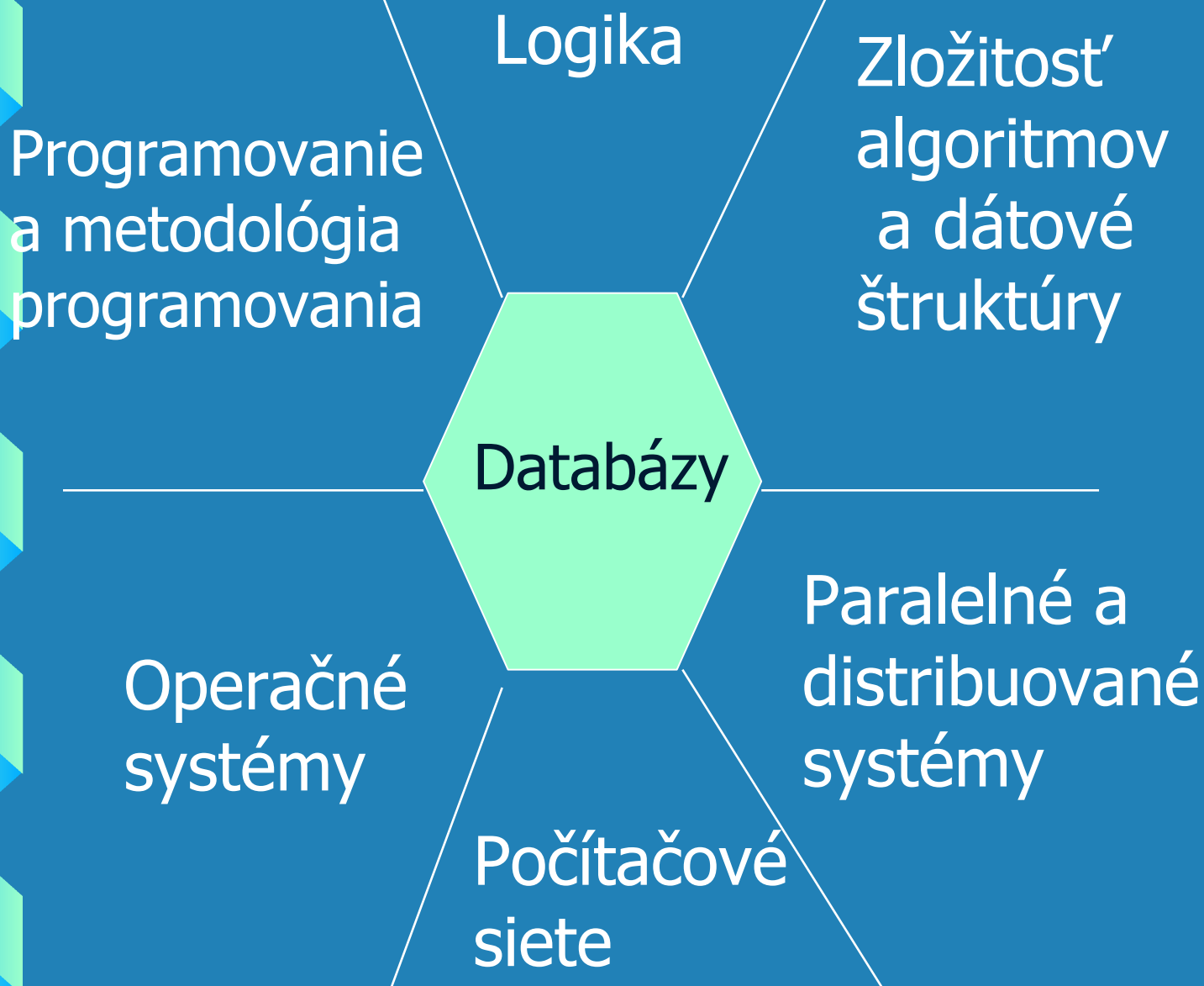
Predmet (účel) databáz

- * **Zobrazenie reality v počítači**
- * **Uchovávanie a udržovanie informácií v konzistentnom stave**
 - **Aktualizácia a modifikácia informácií**
- * **Poskytovanie informácií**
 - **Dotazy**
 - **Periodické správy (reporty, zostavy)**
- * **Ochrana dát pred zničením a neautorizovaným prístupom**

Charakteristika prostredia

- ✦ **Súčasný prístup množstva používateľov k dátam**
- ✦ **Dáta sú rozdelené na viac médií resp. počítačov**
- ✦ **Využíva sa počítačová sieť**
- ✦ **Aplikácie sú rozdelené medzi db-server a klientov**
- ✦ **Zložité dotazy - vysoká priepustnosť systému**
- ✦ **Paralelné a distribuované spracovanie**

Postavenie databáz v informatike



Definícia systému riadenia bázy dát

- * **SRBD poskytuje efektívne “pohodlné”, bezpečné mnohoužívateľské prostredie pre prácu s veľkým objemom perzistentných dát.**
 - **Jednoduché príkazy**
 - **Bezpečnosť**
 - proti chybám systému
 - proti zlým úmyslom užívateľa
 - **Súčasný prístup viacerých užívateľov**
 - transakcie (ACID)

Charakteristika databázového systému (SRBD/DBMS)

- * **Schopnosť pracovať s perzistentnými dátami**
- * **Schopnosť pristupovať k veľkému množstvu dát efektívne**
 - Podporuje aspoň jeden dátový model
 - Podporuje jazyk vyššej úrovne pre definíciu dát a manipuláciu s dátami
 - Podporuje spracovanie transakcií
 - Zabezpečuje ochranu dát a autorizáciu
 - Zabezpečuje zotavenie z chýb systému

Porovnanie databázového a klasického spracovania dát

Klasické spracovanie

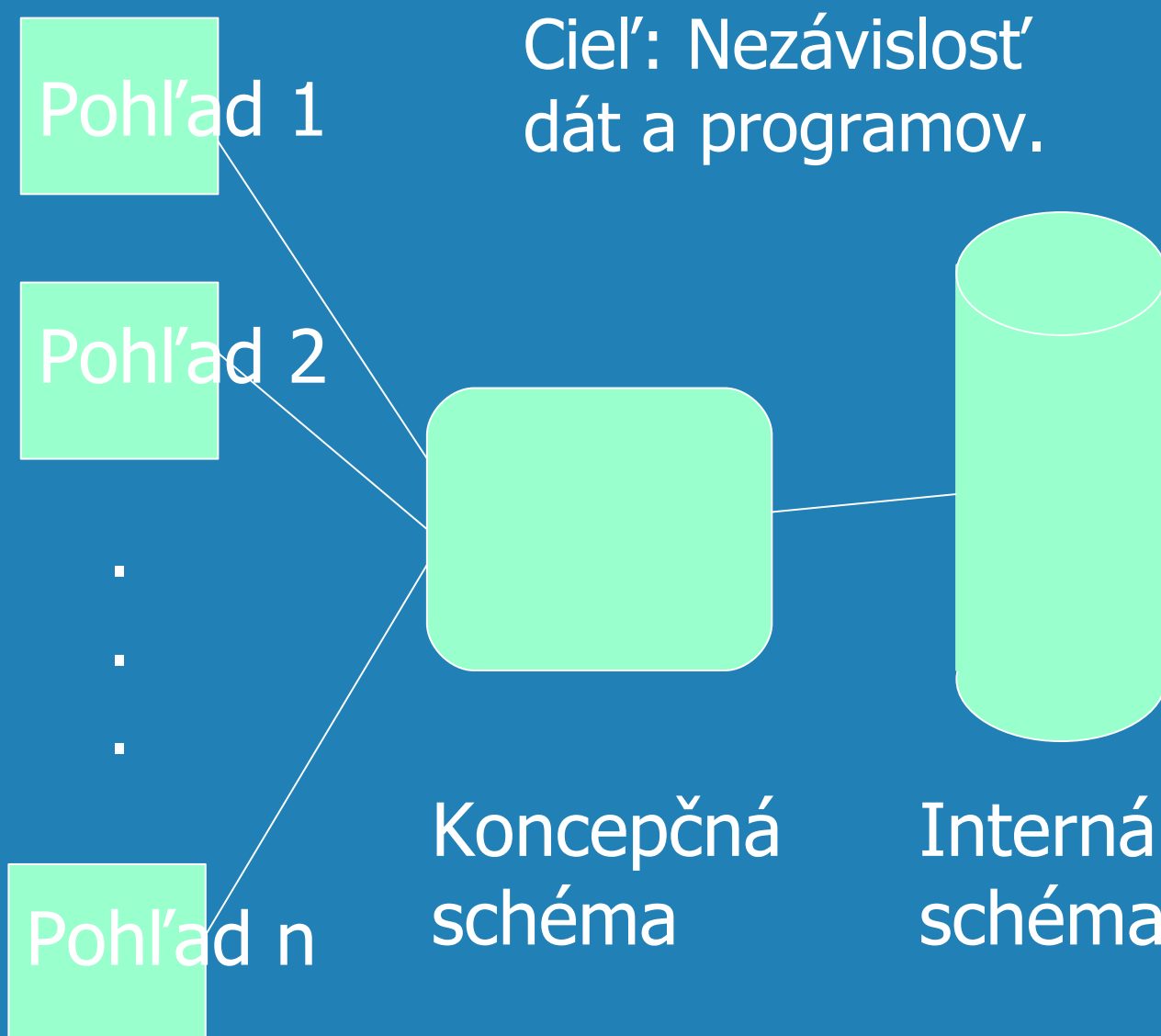
- * Zmena dát implikuje zmenu programu
- * Efektívnosť a ochrana na úrovni OS alebo urob si sám

Databázy

- * Nezávislosť organizácie dát a programov
- * Efektívnosť pri práci s veľkým objemom dát
- * Ochrana dát je transparentná

Trojschémová architektúra

Návrh ANSI Sparc



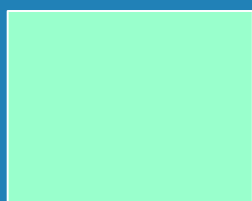
Dátové modely

Dátový model je matematická notácia pre dáta a operácie s dátami. (Algebra)

Poznáme nasledujúce dátové modely:

- entitno – relačný (UML)
- relačný
- navigačný
 - (hierarchický/sieťový)
 - XML
 - Objektový

Základné pojmy entitno-relačného modelu

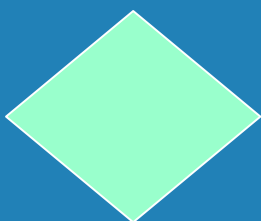


Entita (objekt)



Atribút (vlastnosť)

- identifikačný
- popisujúci



Vzťah (relácia)

arita, horný a dolný pomer

Žiadné operácie

Základné pojmy relačného modelu

Jediný koncept je Relácia (tabuľka)
ako materializácia predikátu

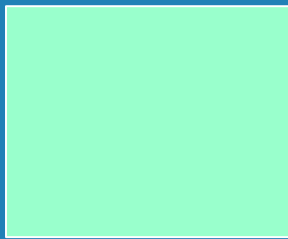
Operácie:

- zjednotenie, prienik, rozdiel
- kartézsky súčin
- selekcia
- projekcia

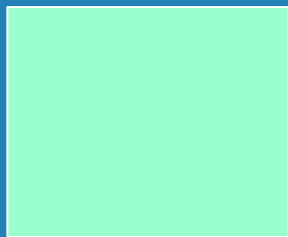
Závislosti:

- funkčné
- multizávislosti
- všeobecné formuly

Základné pojmy siet'ového modelu



Record, record type
(veta, typ vety)



Owner (vlastník)



Set, set
type (typ
setu)



Member
(člen)

Siet'ový model a jeho operácie

Navigačný model je koreňový graf, ktorého vrcholy sú typy viet a hrany typy setov.

Koreňom je systém

Operácie:

- find first (member) within a set
- find next
- find owner within a set

Stav množina smerníkov (currency indicators)

Základné pojmy XML modelu

* XML jazyk na popis dokumentov

- Strom linearizovaný systémom správne konštruovaných zátvoriek (tagov)
- DTD – definícia typu dokumentu
- Otkazy (linky, smerníky)

* Xpath

- Jednoduchý jazyk na výber elementov na základe štruktúry

* Xquery

- Rozšírenie Xpath o testy a príkazy.

Dotazové jazyky

- **SQL** structured query language
- QBE query by example
- relačná algebra
- relačný (predikátový) kalkul
- **Datalóg**
- Xquery
- Hostiteľský jazyk

Úlohy v databázach

- * Navrhovanie (design)
- * Dotazy
- * Aplikácie
- * Formuláre
 - Plnenie databázy
 - Periodické reporty
- * Transakcie

suroviny

suroviny

výrobky

výrobky



nákupná
objednávka

pokyn
k nákupu

pokyn
k výrobe

predajná
objednávka

dáta
dodávateľa

dáta
nákupnej obj.

vyúčtovanie
dlžníka

uloženie
nákupnej
objednáv-
ky

tvorba
vyúčto-
vania
dlžníka

dáta
dod.

dáta nákupnej
objednávky

Zákazník a
účtové dáta

pamäť pre dáta

informačný systém

Plán nasledujúcich prednášok

Plán prednášok je značne približný. Zahrňuje vecný obsah preberanej látky. Časový rozvrh môže sa môže občas o jednu prednášku líšiť.

Doporučujem v predstihu naštudovať materialy k preberanej téme a na prednáškach sa pýtať. Konzultácie v priebehu semestra si možno dohodnúť po prednáške alebo e-mailom. V skúškovom období konzultácie neposkytujem.

Matematické základy relačného modelu (2. prednáška)

- * **Predikátový kalkúl**
- * **Relačná algebra**
- * **Vzťah algebry a kalkulu**
- * **Tabuľky**
- * **Domény**
- * **Rozšírenia**
 - **Aritmetika**
 - **Agregačné funkcie**
 - **Rekurzia (riešenie rovníc v relačnej algebre)**
- * **Závislosti, kľúče**

Relačné jazyky

(3. a 4. prednáška)

- * **Kalkul a algebra**
- * **Datalóg**
- * **SQL**
- * **QBE**
- * **Popis syntaxe a definícia sémantiky jazykov.**
- * **Príklady dotazov**

Precvičeniu dotazov a jazyka SQL budú venované cvičenia.

Navrhovanie báz dát

(5. a 6. prednáška)

* **Poloformálne metódy**

- **Entitno-relačné modelovanie**
- **Objektové návrhy**
- **Binárny model, NIAM**

* **Normalizácia relačných báz dát**

- **Funkčné závislosti a ich vlastnosti**
- **Bezstrátové spojenie**
- **BCNF a 3NF**

* **Všeobecné závislosti**

- **Multizávislosti a 4NF**
- **Inklúzne závislosti, referenčná integrita**

Spracovanie transakcií

(7. a 8. prednáška)

- * **Problémy konkurencie**
 - Uviaznutie, starvation
- * **Vlastnosti transakcií - ACID**
 - Commit, rollback
- * **Sériovateľnosť**
 - Zámky a zamykanie
 - Časové razítka
 - Kópie dát
- * **Žurnálovanie**
 - Undo, redo journal
 - Kontrolné body
- * **Bezpečnosť**

Fyzická úroveň dát

(9. a 10. prednáška)

- * **Halda a sekvenčné súbory**
- * **Index sekvenčné súbory**
- * **Varianty B-stromov**
- * **Hašovanie a dynamizácia hašovania**
 - **Dynamické hašovanie**
 - **Rozšíriteľné hašovanie**
 - **Lineárne hašovanie**
- * **Dotazy na čiastočnú zhodu**
- * **Implementácia relačných operácií**

Kompresia dát

(11. prednáška)

- * **Kompresia medzier**
- * **“Run length coding”**
- * **Diferenčné kódovanie**
- * **Statický slovník**
- * **Huffmanové kódy**
- * **Dynamický slovník (LZW)**
- * **Aritmetické kódovanie**

Navigačné modely

(12. a 13. prednáška ako vyjdú)

- * **XML model (nová prednáška)**
 - **Základy XML**
 - **DTD**
 - **Xpath a Xquery**
- * **Sieťový model (asi už len história)**
 - **Record sety**
 - **Navigácia v setoch**
 - **Hostiteľský jazyk (kedysi COBOL)**