

PA152

Implementace
databázových systémů

RAID level 1

- zrcadlení disku
- výpočet MTTF
 - 2 stejné disky, MTTF 3 roky
 - výměna vadného – 3,5 dne
 - výpadek – oba disky během 3,5 dne
 - ▷ $p(\text{výpadku disku za rok}) = 1/6$
 - ▷ $p(\text{výp. d. za 3,5 dne}) = 1/6 * 3,5/365 = 1/600$
 - ▷ $p(\text{výp. celku za rok}) = 2 * 1/6 * 1/600 = 1/1800$
 - ▷ $\text{MTTF} = 900 \text{ let}$

RAID level 4

- datové disky + paritní disk
- kontrolní součet
- MMTF
 - ▷ $p(\text{výp. 2 disků}) * (\text{kombinace disků})$
 - ▷ 3+1 disk
 - ▷ $1/3600 * (4*3)$
 - ▷ $\text{MTTF} = 150 \text{ let}$

RAID level 6 – příklad

disk	datové			redundantní			
	1	2	3	4	5	6	7

1	1	1	0	1	0	0
---	---	---	---	---	---	---

1	1	0	1	0	1	0
---	---	---	---	---	---	---

1	0	1	1	0	0	1
---	---	---	---	---	---	---

disk	obsah
------	-------

1	11110000
---	----------

2	10101010
---	----------

3	00111000
---	----------

4	01000001
---	----------

5	01100010
---	----------

6	00011011
---	----------

7	10001001
---	----------

Formy uložení dat

Co chceme ukládat?

- položky
- záznamy
- bloky
- soubory

Kam ukládat?

- disk
- (operační paměť)

Uložení dat -- terminologie

Položka Záznam Kolekce

Files field record file

C field struct array, file

SQL attribute tuple relation

Uložení datových elementů

Co chceme ukládat?

- jméno
- plat
- datum
- obrázek

Kam ukládat?

- posloupnost bajtů

Typy datových elementů

- celé číslo
 - 2/4 bajty
- reálné číslo
 - plovoucí čárka
 - mantisa + exponent
 - pevná čárka
- znaky
 - kódová stránky ASCII/UTF

Typy datových elementů

datum

- počet dní od "počátku"
- řetězec YYYYMMDD

čas

- počet sekund od půlnoci
- řetězec HHMMSSFF

výčtové typy

- očíslování hodnot
- red→1, green→2, blue→3, yellow→4, ...

Typy datových elementů

- pravdivostní hodnota

- True 11111111

- false 00000000

- Použití méně než 1 bajtu?

- ▷ s rozmyslem

- řetězec znaků

- pevná délka

- ukončení nulou

- uložená délka

Uložení datových elementů

- většinou pevná délka
 - každý typ má svoji bitovou reprezentaci

- proměnlivá délka
 - velikost na začátku

- každý element "má" svůj typ
 - interpretace bitů
 - velikost
 - speciální hodnota "neobsazeno" (NULL)

Uložení záznamů

- záznam
 - kolekce položek
 - je popsán ve schématu
- schéma
 - typ záznamu
 - počet položek
 - typ každé položky
 - pořadí položek v záznamu
 - význam každé položky

Uložení záznamů

- pevný formát
 - schéma je uloženo mimo záznamy
 - každý záznam na stejný počet bajtů

- proměnlivý formát
 - každý záznam obsahuje svoje schéma

Proměnlivý formát záznamu

- použití
 - "řídké" záznamy
 - ▷ většina položek s hodnotou NULL
 - opakování položek stejného typu
 - vyvíjející se formát
 - ▷ změny schématu během života databáze
- může docházet k plýtvání místem
- varianty kombinující pevný a proměnlivý formát

Uložení záznamů do bloků

záznamy

- pevné délky
- proměnné délky

bloky pevné velikosti

Uložení záznamů do bloků

Možnosti

- oddělení záznamů
- souvislé bloky
- bloky se záznamy různých typů
- rozdělení záznamu
- uspořádání záznamů
- odkazy na záznamy

Oddělení záznamů

- identifikace začátku a konce záznamu
 - značky
 - adresy
- záznamy pevné délky
 - začátek pole záznamů
 - počet obsazených prvků pole

Souvislé/nesouvislé bloky

nesouvislé

- každý záznam součástí jednoho bloku
- jednodušší, ale může plýtvat místem

souvislé

- záznam může začínat na konci jednoho bloku a pokračovat na začátku dalšího bloku
- nutné pokud velikost záznamu je větší než velikost bloku

Míchání záznamů

- blok obsahuje záznamy různých typů

- sdružování
- záznamy, ke kterým často přistupujeme dohromady, jsou uloženy ve stejném bloku
 - Př.: student + studium

Míchání záznamů

□ hlavní nevýhoda:

- složitější struktura

□ kompromis:

- bez míchání, ale bloky souvisejícími záznamy jsou na stejném stopě

Rozdělení záznamů

- kombinace položek s pevnou a proměnlivou délkou
- rozdělení do různých bloků
 - blok s pevnou částí
 - blok s proměnlivou částí

Uspořádání záznamů

- záznamy jsou v souboru (a bloku) setříděny podle hodnot klíče

- efektivní čtení záznamů v daném pořadí
 - např. pro merge-join

Uspořádání záznamů

způsob uspořádání

- záznamy v daném pořadí
- seznam propojený ukazateli
- oblast přetečení

Odkazy na záznamy

- Jak ukládat odkazy na záznamy?

- fyzická adresa
 - ID bloku
 - ▷ (č. zařízení, cylindru, stopy, bloku)
 - offset v bloku

- nepřímo
 - převod: ID záznamu → fyz. adresa
 - ID záznamu
 - ▷ libovolná posloupnost bitů

Odkazy na záznamy

výhody

- fyzická adresa
 - cena přístupu
- nepřímou
 - jednoduchost použití

Modifikace záznamů

vkládání

mazání

- správa volného místa

aktualizace

- stejná velikost – na místě
- jiná velikost – smazání + vložení

Vkládání záznamů

bez uspořádání

- vkládáme na konec nebo do volného místa po smazaném záznamu
- problémy alokace místa pro záznamy s proměnlivou velikostí

Vkládání záznamů

uspořádané záznamy

- pokud je "blízko" volné místo, přeuspořádáme
- jinak -- oblast přetečení

Mazání a odkazy

mazání záznamů

- visící ukazatele

- ukazují na neplatné místo

- náhrobky

- ponecháme značku v převodní tabulce nebo na staré fyz. adrese

- ID záznamu v záznamu

- při přechodu přes ukazatel testujeme shodu ukazatele a ID

Správa vyrovnávací paměti

- obecné strategie uvolňování vyrovnávací paměti
 - nejdéle nepoužitý

- přípíchnuté bloky
 - trvale umístěny ve vyrovnávací paměti

Prohazování odkazů

- blok ve vyr. paměti
 - odkazy ukazují do paměti místo na disk

- automatické
 - při načtení do paměti se všechny ukazatele prohodí
- na žádost
 - prohození při prvním použití odkazu
- žádné