Dokumentace projektu

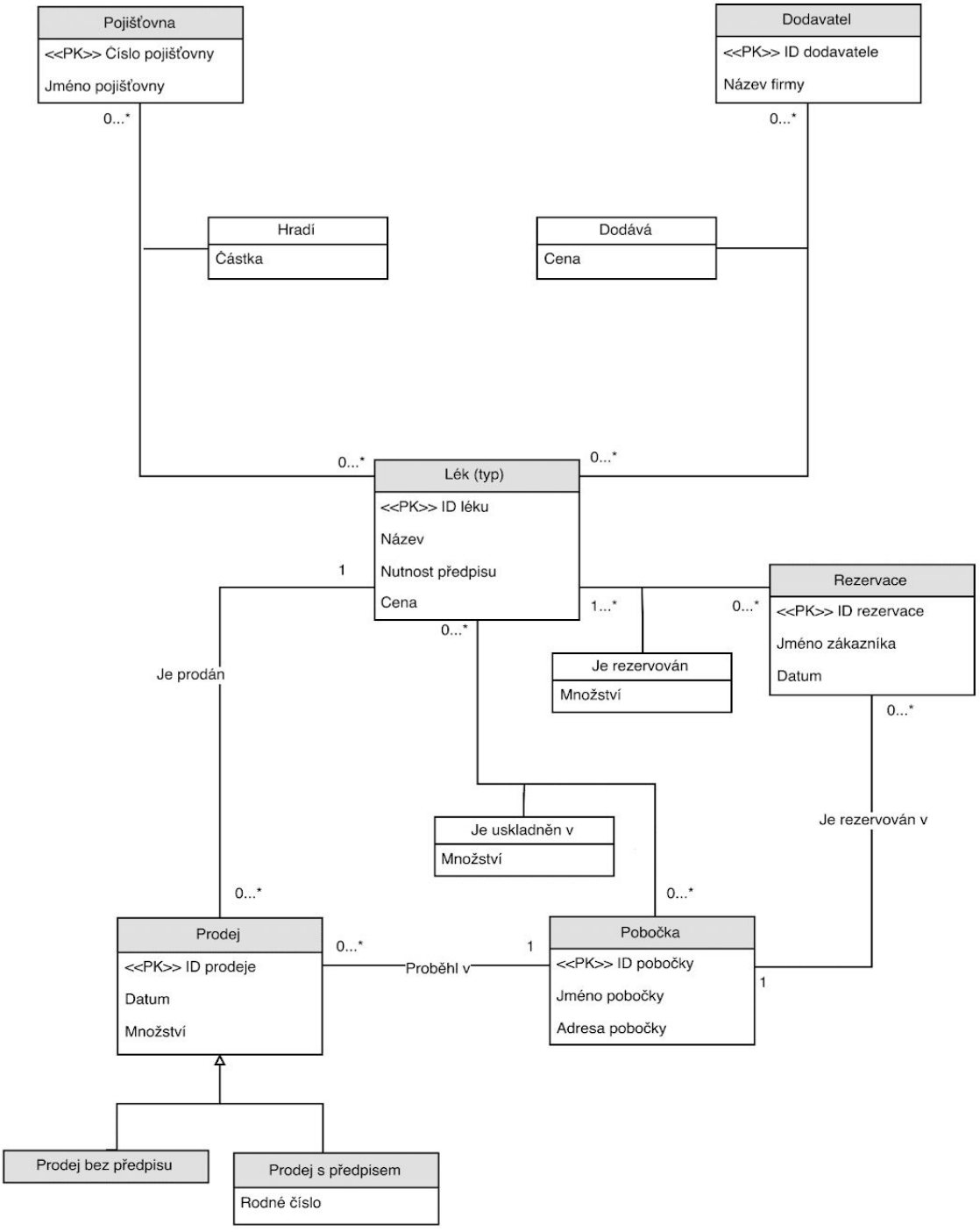
Databázové systémy

Jiří Furda (xfurda00)
Peter Havan (xhavan00)

Fakulta informačních technologií
Vysoké učení technické v Brně

1. května 2018

Návrhový diagram



Popis řešení

V tomto projektu jsme ve dvojici vytvářeli databázový trigger, proceduru, materializovaný pohled a zkoumali plán uskutečnění dotazu.

Popis TRIGGER

V našem řešení se nachází dva databázové triggery. Trigger *incrementid* je využíván na automatické generování a inkrementování identifikačního čísla (PK) v tabulce *dodavatelé*. Druhý trigger zaručuje fakt, že žádná částka v tabluce *hrazeni* nepřesáhne cenu léku, ke kterému se vztahuje

Popis PROCEDURE

Vytvořili jsme dvě procedury. Procedura *percent_predpis* bez parametrů vypíše procentuální zastoupení léku, které vyžadují předpis. Tato procedura pracuje s vyjímkou *zero_divide*, která značí případ kdy počet léku na předpis je 0. V tomto případě se vypíše 0%.
Druhá procedůra s parametrem *jmeno* (léku) vyhledá pomocí tabulky *uskladneni* všechny pobočky, v kterých se daný lék nachází a vypíše je.

Popis MATERIALIZED VIEW

Pro demonstraci jsme v SQL vytvořili dva pohledy - klasický (*example_view*) a materializovaný (*example_view_materialized*). Následně jsme do tabulky, které se pohledy týkají vložili novou položku. Pokud nyní provedeme dotaz *SELECT* na tyto pohledy, zjistíme, že klasický nově přidanou položku obsahuje, ale materializovaný pohled zůstal nezměněn. Tato nekonzistence dat vznikla tím, že materializovaný pohled při svém stvoření vytvoří novou tabulku, kde uloží data, jež obsahuje. Tyto data se v praxi obnovují periodicky, ale v našem případě jsme obnovení nenastavovali.
Naopak klasický pohled žádnou tabulku fyzicky nevytváří, jedná se jen o jakousi zkratku *SELECT* dotazu, který proběhne vždy, když zavoláme dotaz na tento pohled. Proto také obsahuje aktuální data.
Z této ukázky vyplývá, že materializovaný pohled lze využít v případe, kdy nutně nepotřebujeme aktuální data, ale chceme snížit zatížení databáze.

Popis EXPLAIN PLAN

Z důvodu úspory místa následující tabulky neobsahují sloupec *„Time“* obsahující vždy hodnotu *00:00:01.* Dále ve sloupci *„Operation“* byly mezery značící hierarchii operací pro přehlednost nahrazeny pomlčkami.

Plán dotazu bez použití indexu

Id	Operation	Name	Rows	Bytes	Cost (%CPU)
0	SELECT STATEMENT		2	362	7 (15)
1	-FILTER				7 (15)
2	--HASH GROUP BY		2	362	6 (0)
3	---HASH JOIN		2	362	3 (0)
4	----TABLE ACCESS FULL	PRODEJE	2	52	3 (0)
5	----TABLE ACCESS FULL	LEKY	3	465	3 (0)

Plán dotazu s použitím indexu

Id	Operation	Name	Rows	Bytes	Cost (%CPU)
0	SELECT STATEMENT		2	362	6 (17)
1	-FILTER				
2	--HASH GROUP BY		2	362	6 (17)
3	---NESTED LOOPS		2	362	5 (0)
4	----NESTED LOOPS		2	362	5 (0)
5	-----TABLE ACCESS FULL	PRODEJE	2	52	3 (0)
6	-----INDEX UNIQUE SCAN	EXAMPLE_INDEX	1		0 (0)
7	----TABLE ACCESS BY INDEX ROWID	LEKY	1	155	1 (0)

Popis plánu

Oba způsoby začínají operací *SELECT STATEMENT*, která zaštiťuje celý dotaz typu *SELECT*. Ta obsahuje operaci *FILTER*, která odstraní všechny řádky nesplňující podmínku (v tomto případě *SUM(prodej_mnozstvi) > 1*).

Operace *HASH GROUP BY* zase seskupí položky pomocí tabulky s rozptýlenými položkami. Dále se dotazy liší na základě (ne)využití indexů.

Popis plánu bez využití indexu

Operace *HASH JOIN* opět využívá tabulku s rozptýlenými položkami, nejprve do ni vloží všechny záznamy z první tabulky a poté je propojuje s každým záznamem uvnitř druhé tabulky. K tomu je třeba použít dvakrát funkci *TABLE ACCESS FULL*, která kompletně přečte všechny řádky i sloupce dané tabulky.

Popis plánu s využitím indexu

Oproti tomu dotaz s využitím indexu ke spojení použije dvakrát operaci *NESTED LOOP*, která pro každý řádek první tabulky prohledá všechny řádky druhé tabulky.
První *NESTED LOOP* opět přečte celou tabulku pomocí *TABLE ACCESS FULL*, ale u druhé tabulky využije již vytvořené indexy a proto proběhne operace *INDEX UNIQUE SCAN*, která k hledání indexu využije průchod B-stromem.
Druhý *NESTED LOOP* využije záznamy z předešlého *NESTED LOOP* a s využitím operace *TABLE ACCESS BY INDEX ROWID* již bez dvojího prohledávání získá odpovídající řádek.
Z předchozích tabulek lze vyčíst, že tato varianta je levnější co se týče procesorového času.

Zdroje

https://use-the-index-luke.com/sql/explain-plan/oracle/operations
https://docs.oracle.com/en/
https://www.visual-paradigm.com/VPGallery/datamodeling/EntityRelationshipDiagram.html