



PDB

návrh aplikace/databázového systému

2022/2023

Zadání úlohy

4. decembra 2022

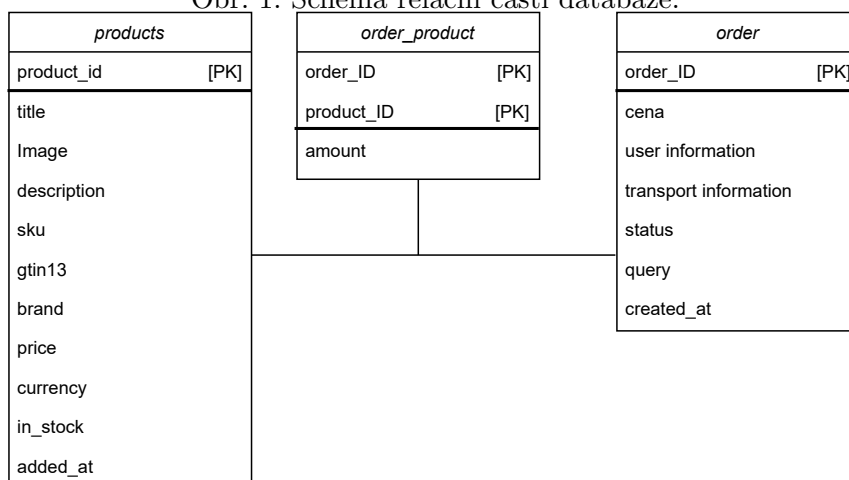
Radoslav Eliáš (xelias18)
Vojtěch Jahoda (xjahod06)

1 Specifikace

Data obchodu Home Depot. Z kterých chceme vytvoriť databázi pro e-shop, kde využijeme výhod relační databáze pro řešení potencionálního konfliktu vykoupení zboží od několika různých uživatelů. Jelikož každé zboží má omezený počet kusů na skladu.

2 Architektura systému

Obr. 1: Schéma relační části databáze.



2.1 CQRS

Systém bude využívať model architektúry CQRS. Oddelenie dotazovacej a príkazovej časti systému nám umožní zaistiť dve veci. Zaprvé, tranzakčné spracovanie údalostí zápisu znamená, že bude existovať jedno absolútne poradie a história príkazov a pri konflikte bude jednoznačné ktorá akcia bola vykonaná prvá. Teda v našom prípade, ak sa dvaja zákazníci pokúsia zakúpiť rovnaký produkt, budeme vedieť rozhodnúť ktorá udalosť prebehla skôr. Podruhé, modul pre dotazovanie dát z databáze s využitím distribuovaného systému zabezpečí rýchle dotazovanie a celkovo vysoký priemerný čas dostupnosti systému.

Čítanie bude prebiehať priamo nad NoSQL databázou. Dáta budú uložené v podobnom formáte v akom sa budú posilať, čo umožní vynechať proces formátovania a zabaľovania, a teda skráti trvanie každého dotazu. Zapisovanie zapuzdrí model, ktorý najprv vykoná validáciu požadovanej zmeny. Overí, či neporušuje existujúce databázové integritné obmedzenia, či nerozbiže konzistenciu databáze a či výsledné dáta budú korektné podľa špecifikácií

dátovej sady. Ak zmena spĺňa všetky pravidlá, modul sa pokúsi vykonať zápis do zdroju pravdy - teda relačnej databázi. V prípade neúspechu vráti chybovú hlášku užívateľovi cez aplikačné rozhranie. V prípade úspechu vykoná zápis aj do noSQL databáze a tým navráti systém do stavu konzistencie.

3 Dotazy

- základní CRUD operace s produktem
- vyhľadávání produktů podle ceny i různé rozmezí cen
- vyhľadávání produktů podle času vložení
- vyhľadávání podle počtu kusů na skladě
- vyhľadávání podle kategorií zakódovaných v SKU
- vyhľadávání podle značek
- vyhľadávání podle popisu produktu (klíčová slova hledané v popisu)

4 Popis Dat

Data pochází z opensource databáze firmy Home depot ¹.

product_id Unikátní ID produktu v obchodu. [integer]

title Název produktu. [string]

images Url adresa pro načtení obrázku produktu. [string]

description Popis produktu. [string]

sku (Skladová jednotka) je jedinečný identifikátor položky prodávané maloobchodníkem. [integer]

gtin13 (Global Trade Item Number), jsou unikátní identifikátory platné kdekoliv na světě. [integer]

brand Označení zda se jedná o značkový nebo neznačkový produkt. [string]

price Cena produktu v dané měně. [integer]

currency Měna daného produktu. [string]

in_stock Počet kusů na skladě. Zde přichází potencionální konflikt, když si daný produkt bude chtít koupit více uživatelů zároveň. [integer]

added_at Čas přidání daného produktu. [timestamp]

5 Technologie

Hlavním implementačním jazykem bude pro náš systém Python3. Tento skriptovací jazyk je často používaný pro práci s velkým množstvím dat. Navíc poskytuje velký počet knihoven, pro komunikaci s jinými aplikacemi, například objektovo relačné mapovanie a rozhranie pre väčšinu bežných databázových systémov. Python využijeme hlavne na tvorbu aplikačného rozhrania. Na to použijeme framework FastAPI².

Relačnú databázu, ktorú použijeme ako "zdroj pravdy", sme vybrali Oracle SQL databázu spravovanú fakultou VUT a poskytnutou študentom v rámci kurzu PDB. Pre čítanie sme vybrali NoSQL databázu MongoDB. Je to dokumentová databáza ktorá ukladá dáta vo formáte podobnom JSON súboru. Podľa webového blogu db-engines³ je to najpoužívanejšia NoSQL databáza.

¹<https://query.data.world/s/7ywhs3irjz3jwexgfc7idx6pmz62kl>

²<https://fastapi.tiangolo.com/>

³<https://db-engines.com/en/>

6 Testování

Testovanie funkčnosti a kvality vytvorenej aplikácie bude prebiehať pomocou systémového testovania. Na aplikačné rozhranie budú posielané korektné aj chybné dotazy. Potom bude skúmané ako sa systém zachoval, musí byť schopný odhaliť chybné akcie a zamietnuť ich s užitočnou chybovou hláškou. Ďalej musí systém vrátiť správne filtrované dáta po dotaze čítania a po zápise alebo zmene hodnôt musí byť interný stav databázy korektne zmenený.