

**Česká zemědělská univerzita v Praze**

**Provozně ekonomická fakulta**

**Katedra Informačního Inženýrství**



**Teze diplomové práce**

**Možnosti optimalizace přístupu k databázově evidovaným datům**

**Bc. Miroslav Smutný**

© 2020 ČZU v Praze

## **Abstrakt**

Diplomová práce se zabývá problematikou optimalizace přístupu k databázově evidovaným datům a jejím možnostem. Teoretická část popisuje metody přístupu k datům, analýzu exekučních plánů, optimalizační techniky a samotný proces optimalizace.

V praktické části je nejprve provedeno zmapování současné situace, po kterém následuje prezentace příkladu optimalizace v nejmenované společnosti. Poté je navržen datový model pro správu životopisů, na kterém jsou identifikovány datové výstupy. Pro definovaný datový výstup je sestaven text dotazu. Následuje zátěžový test, jehož průběh je po dokončení analyzován. Na základě analýzy je navrženo použití optimalizačních technik, které jsou následně analyzovány a porovnány s prvotní analýzou. Pro každou optimalizační techniku je spuštěn samostatný ověřovací zátěžový test. Všechny testy jsou následně porovnány s průběhem prvního testu a je identifikována nejvhodnější optimalizační technika. Na závěr je řešení zhodnoceno a řešení je zobecněno pro další využití v praxi.

**Klíčová slova:** relačně databázová technologie, SQL, indexování, SRBD Oracle

## **Cíl práce a metodika**

Diplomová práce je zaměřena na problematiku optimalizace přístupu k relačně databázově evidovaným datům. Hlavním cílem je navrhnout postup optimalizace přístupu k datům, který povede ke snížení časové a zdrojové náročnosti přístupu na základě požadavků identifikovaných v teoretické části. Využity budou postupy optimalizace jako transkripce dotazů, úprava exekučních plánů, indexace sloupců, tabulkové oddíly, nápovědy optimalizátoru exekučních plánů nebo úprava parametrů. Dílčími cíli je seznámení s principy a kritérii optimalizace náročnosti přístupu k datům. Dalšími dílčími cíli je zmapování momentálního stavu, navržení a ověření řešení na vzorové databázové evidenci, Navržené řešení je zobecněno pro další použití.

Použitá metodika této diplomové práce bude založena na studiu a analýze dostupných informačních zdrojů a existujících řešení v oblasti optimalizace přístupu. Stěžejními metodami této práce budou principy optimalizace, mezi které patří úprava exekučních plánů, indexace sloupců, tabulkové oddíly, nápovědy optimalizátoru exekučních plánů nebo úprava parametrů. Navrhované řešení bude zohledňovat identifikované požadavky a očekávání spojená

s optimalizací přístupu. Na podkladě syntézy teoretických poznatků a dosažených výsledků optimalizace identifikovaných dotazů budou formulovány závěry této diplomové práce a následně zobecněny pro další možná použití.

## **Vlastní práce**

Teoretická část popisuje metody přístupu k datům v kontextu s indexací, oddíly, statistiky datových objektů. Dále jsou popsány exekuční plány, optimalizátor dotazů. Následně jsou popsány jednotlivé optimalizační techniky, konkrétně změna sémantiky dotazu, parametrizace a změna využití fyzických zdrojů. Nakonec je popsán samotný proces optimalizace.

V praktické části byly identifikovány problémy přístupu k datům v sociálně známých společnostech, jež měly mediální dopady. Poté byl prezentován příklad optimalizace z reálného produkčního prostředí nejmenované společnosti.

Pro účely analýzy a ověření optimalizačních technik byl vytvořen datový model na němž byly identifikovány aplikační dotazy. Datový model reprezentuje problematiku správy životopisů. Následoval zátěžový test, na jehož základě se identifikoval problematický dotaz.

Problematický dotaz následně prošel řadou optimalizačních technik, přičemž pro každou techniku bylo nutné analyzovat exekuční plán. Plány byly poté porovnány s původním exekučním plánem v oblastech nákladovosti, době trvání a použitých zdrojích.

Samotné ověření optimalizace dotazů probíhalo individuálně zopakováním zátěžového testu a monitorováním stejných statistik a parametrů jako v případě prvního testu. Výsledky všech testů bylo možné porovnat a ověřit prospěšnost optimalizace.

Navržené řešení je pouze jednou z několika možností, jak procesem optimalizace dosáhnout definovaného cíle. Dalšími řešeními může být úprava zdrojů serverů, úprava parametrů datových uložišť nebo úprava parametrů databázového a operačního systému.

Mírně odlišnou možností optimalizace jsou databáze běžící na více serverech, nazvané clustery. V tomto případě se nejedná o snížení nákladovosti a doby exekuce jednotlivých dotazů, ale spíše o rozložení zátěže celého rozsahu aplikačních dotazů. Díky rozložení zátěže jsou směrovány aplikační přístupy na různé servery, které následně využijí zdrojů na těchto rozdílných serverech.

## Závěr

Přínosem diplomové práce je ověření aplikovatelnosti různých optimalizačních technik a jejich komparace. Aplikovatelnost lze pojmut jako splnění cílů optimalizace pomocí dosažení nízké náročnosti na výpočetní výkon, na diskové uložení a na paměťové struktury v exekučních plánech dotazů při zachování stejných výstupních dat.

Dalším přínosem je také analýza chování dotazu pro rozdílných technikách optimalizace. Každá optimalizační technika má své kladné a záporné stránky, které byly do určité míry prezentovány na stabilitě a době exekuce dotazů.

V této diplomové práci byl navržen postup optimalizace přístupu k datům, který vedl ke snížení časové a zdrojové náročnosti přístupu na základě požadavků identifikovaných v teoretické části. Bylo provedeno seznámení s principy a kritérii optimalizace náročnosti přístupu k datům. Byl zmapován momentální stav a bylo navrženo a ověřeno řešení na vzorové databázové evidenci. Navržené řešení bylo z hlediska procesu optimalizace zobecněno pro další použití.

## Seznam použitých zdrojů

1. **Laurenčík, Marek.** *SQL - Podrobný průvodce uživatele.* Praha : Grada, 2018. 9788027107742.
2. **Microsoft.** SQL Docs. *Microsoft.* [Online] 2019. [Citace: 22. Únor 2020.] <https://docs.microsoft.com/en-us/sql/sql-server>.
3. **Roeser, Mary Beth.** SQL Language Reference. *Oracle Help Center.* [Online] April 2017. [Citace: 29. Červen 2019.] <https://docs.oracle.com/en/database/oracle/oracle-database/12.2/sqlrf/>. E83703-01.
4. **Ashdown, Lance.** SQL Tuning Guide. *Oracle Help Center.* [Online] 2019. [Citace: 17. Únor 2020.] <https://docs.oracle.com/en/database/oracle/oracle-database/12.2/tgsql>. E85762-05.
5. **Ashdown, Lance a Kyte, Tom.** Database Concepts. *Oracle Help Center.* [Online] Březen 2018. [Citace: 25. Únor 2020.] <https://docs.oracle.com/en/database/oracle/oracle-database/12.2/cncpt>. E85769-04.