



MATEMATICKO-FYZIKÁLNÍ
FAKULTA
Univerzita Karlova

Zápis o obhajobě disertační práce

Akademický rok: 2023/2024

Jméno a příjmení studenta: Mgr. Tomáš Faltín
Identifikační číslo studenta: 46734522

Typ studijního programu: doktorský
Studijní program: Informatika - Softwarové systémy
ID studia: 559861

Název práce: Distributed Graph Query Engine Improvements for Big Data Graphs
Pracoviště práce: Katedra softwarového inženýrství (204. • 32-KSI)
Jazyk práce: angličtina
Jazyk obhajoby: čeština
Školitel: RNDr. Jakub Yaghob, Ph.D.
Oponent(i): Riccardo Tommasini

Datum obhajoby: 11.12.2023 **Místo obhajoby:** Praha
Termín: řádný

Průběh obhajoby: Obhajoba probíhala hybridním způsobem, tak že část komise byla připojena vzdáleně přes platformu Zoom. Předseda komise stručně představil uchazeče a následně prohlásil, že byly splněny všechny podmínky pro konání obhajoby. Školitel přednesl své vyjádření k osobě uchazeče a jeho výsledkům dosaženým během studia.

V další části obhajoby uchazeč prezentoval výsledky své práce v oblasti efektivního dotazování nad velkými grafy dat ("bug data graphs") v distribuovaném prostředí a jejich zpracování.

Hlavní výsledky jsou tyto:

- různé algoritmy průchodu grafem (varianty průchodu "do hloubky", asynchronní) s důrazem na nízkou spotřebu paměti;
- návrh takzvaných "scouting queries" pro účely rychlého vyhodnocení možných plánů provedení složitějších dotazů;
- návrh algoritmu pro vyhodnocování takzvaných "reachability regular path queries (RPQ)" s důrazem na způsob průchodu grafem a minimální spotřebu paměti;
- experimenty provedené za cílem změřit rychlost zpracování "scouting queries" a "reachability regular path queries".

Výsledky měření ukazují na významné zvýšení rychlosti zpracování "scouting queries" v naprosté většině případů.

Celý projekt byl realizován ve spolupráci s firmou Oracle.

Potom se ke předložené dizertaci vyjádřili oba dva oponenti (jeden ve zastoupení přítomným členem komise), shrnuli posudky a položili několik dotazů.

Oponenti hodnotili práci velmi kladně.

Jejich dotazy mířily především na:

- výběr dalších systémů pro účely experimentálního porovnání a určení vhodné baseline
- vliv databázového schématu na rychlost zpracování grafových dotazů
- měření efektů síťové komunikace (posílání zpráv mezi jednotlivými uzly v distribuovaném prostředí)
- způsob výběru a generování "scouting queries" (jak to přesně funguje)
- možnost zobecnění navržených technik na jiné datové modely
- porovnání "scouting queries" a metody zvané PnP

Uchazeč odpověděl na všechny dotazy oponentů, a ti následně doporučili práci ke obhajobě.

Dotazy členů komise se týkaly:

- spolupráce uchazeče s firmou Oracle a jejího pokračování
- hlavní činnosti uchazeče v rámci celého projektu (návrh, implementace, příprava publikací)
- koncept "data sharding" (principy) a jestli všechny techniky použité v rámci experimentů byly nakonfigurované tak aby použily stejný mechanismus pro "data sharding"

Obhajoba pokračovala neveřejnou částí (tedy uzavřeným jednáním komise).

Proběhlo hlasování metodou "zvednutím rukou" s jednoznačně kladným výsledkem (všechny hlasy kladné), na jehož základě byl udělen titul Ph.D.

Výsledek obhajoby:

prospěl/a (P)

Předseda komise:

doc. RNDr. Petr Hnětynka, Ph.D.

Členové komise:

doc. RNDr. Pavel Parízek, Ph.D.

doc. Ing. Lubomír Bulej, Ph.D.

doc. RNDr. Vlastislav Dohnal, Ph.D.

doc. RNDr. Irena Holubová, Ph.D.

prof. Ing. Michal Krátký, Ph.D.

doc. RNDr. Jakub Lokoč, Ph.D.