

České vysoké učení technické v Praze
Fakulta elektrotechnická
Katedra počítačů

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: Moravec Jakub

Studijní program: Otevřená informatika
Obor: Softwarové inženýrství

Název tématu: Analýza a návrh abstraktní vícevrstvé architektury pro práci s grafovou databází realizující metadatové úložiště pro data lineage

Pokyny pro vypracování:

Projekt Manta Tools poskytuje sadu nástrojů pro data lineage. Jako úložiště metadat se používá grafová databáze (GDB). Různé GDB poskytují různá rozhraní s různou úrovní abstrakce. Použití GDB jako úložiště metadat pro data lineage lze rovněž formalizovat do podoby specifického aplikačního rozhraní (API). Cíle práce jsou následující: 1. Seznamte se s API pro práci s data lineage v projektu Manta Tools a API vybraných GDB (výběr bude upřesněn po diskuzi s vedoucím práce). 2. Analyzujte požadavky na API pro data lineage a pro práci s GDB. 3. Navrhněte vícevrstvou architekturu a příslušná API, která umožní použití různých GDB při zachování stejného API pro práci z pohledu nástrojů Manta Tools. Problémová doména je velmi komplexní, je třeba uvažovat transakční zpracování a jeho podporu, využití specifických indexů a další optimalizačních přístupů konkrétních GDB. 4. Implementujte prototyp vašeho návrhu a otestujte ho na vhodných datech, která dodá vedoucí práce.

Seznam odborné literatury:

- [1] Neo4j Graph Data Modeling (ISBN 978-1-78439-344-1)
- [2] Practical API Design (ISBN 978-1-4302-0973-7)
- [3] Mining and indexing graph databases (ISBN 978-1-303-78261-9)
- [4] Graph Data Management: Techniques and Applications (ISBN 978- 1613500538)

Vedoucí: Ing. Michal Valenta., Ph.D.

Platnost zadání do konce letního semestru 2018/2019

prof. Dr. Michal Pěchouček, MSc.
vedoucí katedry



prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 27.09.2017