

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Oskar Hýbl
Název práce Data Lineage Analysis of Frameworks with Complex Interaction Patterns
Rok odevzdání 2020
Studijní program Informatika **Studijní obor** Softwarové systémy

Autor posudku Doc. RNDr. Pavel Parížek, Ph.D. **Role** Vedoucí
Pracoviště Katedra distribuovaných a spolehlivých systémů

Text posudku:

Tato diplomová práce měla za cíl rozšířit existující platformu Manta Flow, která vytváří grafy data lineage pro softwarové systémy zahrnující databáze a komponenty ve jazyce Java, o podporu složitějších knihoven na správu a zpracování dat. Na základě úvodních konzultací se zástupci firmy Manta Tools byly ještě před formálním zadáním práce vybrány knihovny Apache Spark a Hibernate, jejichž použití ze klientských Java aplikací se vyznačuje poměrně složitými vzory interakce, zejména ve srovnání se jednoduššími knihovnami jako MyBatis a Spring JdbcTemplate, které už platforma Manta Flow podporuje.

V souladu se zadáním tedy autor provedl důkladnou analýzu relevantních aspektů knihovny Apache Spark, a implementoval modul který dokáže spočítat data lineage informace pro volání této knihovny. Modul podporuje docela velkou podmnožinu funkčnosti Spark, a především všechny běžně používané operace. Součástí implementace je také rozsáhlá množina testů, která potvrzuje správnou funkčnost modulu.

Rozsah textové části a kvalita jejího zpracování jsou přiměřené. Nemám zásadní výhrady ani ke struktuře a obsahu textu, přestože byl dokončován v časové tísní. Text obsahuje také porovnání se existujícími nástroji, které řeší podobnou úlohu, jako třeba se nástrojem Spline.

Původní cíle zahrnovaly také podporu knihovny Hibernate. Nicméně implementace modulu pro Spark se postupně ukázala jako velmi složitá a především časově náročnější oproti plánu, takže pro knihovnu Hibernate autor provedl jen základní analýzu, jejíž výsledky tvoří součást textu diplomové práce.

Celkově považuji tuto diplomovou práci za docela zdařilou.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

Pokud práci navrhuje na zvláštní ocenění (cena děkana apod.), prosím uveďte zde stručné zdůvodnění (vzniklé publikace, významnost tématu, inovativnost práce apod.).

Datum 3.9.2020

Podpis