

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Pavel Marek
Název práce ALTREP Data Representation in FastR
Rok odevzdání 2020
Studijní program Informatika **Studijní obor** Softwarové systémy

Autor posudku Štěpán Šindelář **Role** Vedoucí
Pracoviště Oracle Labs

Text posudku:

Předkládaná práce se zabývá implementací ALTREP frameworku programovacího jazyka R v jeho alternativní implementaci FastR. Součástí ALTREP frameworku (Alternate Representations for R Objects) je C API pro registraci uživatelských reprezentací R objektů a integrace těchto alternativních reprezentací se zbytkem systému: Zbývajícím C API pro manipulaci s R objekty a se samotným interpretrem jazyka R. ALTREP se stal součástí referenční implementace R ve verzi 3.5.0 vydané v roce 2018.

Autor v první části práce popisuje použité technologie: FastR, GraalVM, Truffle a Sulong (technologie, na kterých je postaveno FastR), referenční implementaci ALTREP a R balíčky, které ALTREP API již využívají. V druhé části autor podrobněji analyzuje kontrakt ALTREP API a některé podstatné implementační detaily FastR tak, aby pak mohl popsat návrh designu implementace a samotnou implementaci. V závěrečné části vyhodnocuje výkonnost jeho implementace a porovnává ji s referenční implementací.

Analýza ALTREP frameworku i některých implementačních detailů FastR byla významnou součástí práce. Její složitost byla vyšší i kvůli neexistenci ucelené vývojářské ani uživatelské dokumentace ALTREP API. Zdrojový kód referenční implementace byl tak téměř jediným zdrojem.

Vlastní implementace je poměrně rozsáhlá a její kvalita je rozumná. Významná je šíře použitých technologií a nutnost dobré znalosti jejich vnitřního fungování – například garbage collection v JVM nebo základní orientace v IR (intermediate representation) překladače za účelem zlepšení kvality generovaného kódu. Dalším faktorem je to, že autor musel svoje dílo navrhnout a implementovat v rámci existujícího, velmi komplexního a více než 7 let vyvíjeného systému (FastR). Autorovi se podařilo vhodně integrovat svůj kód do zbytku systému. Součástí implementace je i sada rozsáhlých a dobře promyšlených testů.

Při vyhodnocení výkonnosti v závěrečné části autor poměrně detailně analyzuje naměřené výsledky. Příloha k práci dokonce obsahuje i Jupyter notebook s naměřenými daty a skripty, které byly použity pro jejich zpracování.

Práce je celkově na velmi dobré úrovni.

Práci doporučuji k obhajobě.

Práci nenavrhují 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