

Posudek oponenta diplomové práce

Orchestration tool for the modular computing subsystem of the ANALYZA platform

Bc. Richard Kyčerka

Vedoucí: RNDr. Tomáš Rebok, Ph.D.

Oponent: RNDr. Michal Batko, Ph.D.

Cílem práce bylo prozkoumat nástroje pro orchestraci modulárních výpočtů a zpracování dat s preferencí platformy Kubernetes, vybrat vhodné řešení a adaptovat ho pro potřeby orchestrace výpočetních operací projektu ANALYZA.

Práce je strukturovaná do třinácti kapitol. V úvodní kapitole je představen kontext a cíle práce, ale text je zde relativně stručný a postrádám především podrobnější vysvětlení problematiky orchestrace. Druhá kapitola je nazvaná "Big Data analysis", ale fenoménu BigData je věnován prakticky jen první odstavec navíc pouze z pohledu objemu, což je nedostatečné. Spíše se zde jedná o úvod k rešerši orchestračních nástrojů, kterým se věnuje dalších pět kapitol. Rešerše je zpracovaná dobře, kladně hodnotím, že student uvádí informaci i o kvalitě dokumentace a pro každou techniku je k dispozici shrnutí. Ocenil bych ale i celkové shrnutí například tabulkou vlastností a jejich plnění jednotlivými nástroji. Vytknul bych převzetí celého obrázku 4.1, kde jsou obsaženy části, které nejsou relevantní k textu, což je matoucí. Následuje kapitola, kde je stručně představen orchestrační nástroj projektu ANALYZA, jsou zde vyjmenovány vysokoúrovňové požadavky, kde dle mého názoru trochu chybí detailnější popis požadavků na funkcionalitu, a je zde diskutován výběr řešení pro vlastní implementaci. Osmá kapitola se věnuje detailnímu popisu vybraného nástroje Argo a v další kapitole jsou řešeny požadavky pro spuštění orchestrace a stručně i vytvořené API, mapující původní funkcionalitu orchestrátoru projektu ANALYZA na nový nástroj Argo. Zde bych rovněž očekával více podrobností o vlastní implementaci a hlavně detailně rozebrané limity implementace. Další kapitola se věnuje možnému rozvoji orchestrace Argo, kterou by asi bylo vhodnější umístit do závěru. Jedenáctá kapitola se věnuje testování orchestrace, očekával bych ale, že testování bude provedeno nad moduly projektu ANALYZA. Poslední kapitola obsahuje stručné shrnutí práce a možné další rozšíření.

Elektronická příloha obsahuje zdrojové kódy vytvořeného API a šablony pro Argo workflow, které lze sestavit a spustit, ocenil bych však i postup pro nasazení aplikace pro projekt ANALYZA.

Po technické stránce je text na průměrné úrovni, obsahuje menší množství překlepů a gramatických chyb, angličtině by prospěla jazyková korektura. Bibliografické odkazy jsou v pořádku, i když autor rovněž hodně využívá uvádění odkazů formou poznámky pod čarou.

Otázky:

- Jaký přesně je vztah mezi typem úlohy "service" v kontextu původního orchestrátoru projektu ANALYZA a použitou Argo Service? Jak je to v případě typu "job"?
- Je možné podrobněji rozebrat, jak byla modifikována funkcionalita původního API projektu ANALYZA? V sekci 9.2.1 je zmíněna jen jakási modifikace řízení operací na úrovni modul/krok.

Konstatuji, že dle mého názoru práce splňuje zadání a lze ji přijmout jako diplomovou práci. Práci navrhuji hodnotit stupněm **velmi dobře (B)**.

RNDr. Michal Batko, Ph.D.