

Posudek oponenta diplomové práce

Autor práce: **Ondřej Hovjcký**

Název práce: **Uchování a zpracování velkých dat typu JSON**

Dokument práce má velmi dobrou logickou strukturu. Analytická část zahrnuje 2 kapitoly a popisuje možnosti ukládání dat typu *JSON* v relačních a *NoSQL* databázích a také představuje existující benchmarky, které databáze s podporou *JSON* formátu umí otestovat. Na str. 15 se poprvé vyskytuje zkratka *PNUTS*, bohužel v celém dokumentu (ani v seznamu zkratek na závěr) není vysvětleno, co ta zkratka znamená. Text práce je proložen celou řadou obrázků nesoucí mimo jiné spojnicové grafy výsledků. Je velká škoda, že některé tyto grafy jsou pouze černobíle (str. 16, 18, 19) a jiné barevné (str. 30, 31), i když jste autorem všech. Také velikost převážně černobílých grafů je na hranici čitelnosti. V dokumentu postrádám uživatelskou příručku, bez které je nemožné aplikaci uvést do chodu.

Výsledkem práce je webová aplikace, která požadovaným způsobem modifikuje otevřený *YCSB benchmark* pro vybrané databázové systémy (*Apache Cassandra*, *Elasticsearch*, *MongoDB* a *MySQL DS*). Přiložené CD má nevhodnou adresářovou strukturu, ze které nepoznám, kde se nachází zdrojový kód diplomanta a kde zdrojový kód převzatých nástrojů. Zde by mi nepomohl ani textový soubor `readme.txt` v kořenovém adresáři. Také zdrojový kód diplomanta je sporadicky komentován.

Dokument práce má průměrný rozsah 55 stran a je vysázen v *LaTeXu*. Obsahuje nepatrné množství překlepů a velmi dobře se čte. Po typografické stránce nelze dokumentu skoro nic vytknout.

V práci je uvedeno 24 relevantních zdrojů, z toho 7 anglicky psaných odborných knižních publikací, časopisů či sborníků konferencí. Všechny zdroje jsou v textu citovány.

Zadání diplomové práce je splněno bez výhrad.

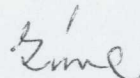
K diplomové práci mám tyto otázky:

1. Podle jakých pravidel byly určovány hodnoty parametrů dotazů uvedených na str. 46 a 47 pro vytváření a spouštění sad dotazů nad generovanými daty.
2. Umožňuje aplikace naplnit testované databáze identickými daty a tato data otestovat identickou sadou dotazů.
3. Co znamená uvedená zkratka *PNUTS* a jak zapadá do řešeného tématu.

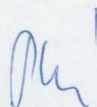
Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.

V Plzni 25. 8. 2017

Ing. Martin Zíma, Ph.D.



**SOUHLASÍ
S ORIGINÁLEM**



Západočeská univerzita v Plzni
Fakulta aplikovaných věd
katedra informatiky a výpočetní techniky