

Oponentský posudek diplomové práce

Název diplomové práce: Network Repository for Performance Evaluation Results
Autor diplomové práce: Pavel Ondroušek
Vedoucí diplomové práce: Doc. Ing. Petr Tůma, Dr.

Předmětem práce je vytvoření úložiště výsledků testů výkonnosti, které dokáže importovat výsledky ve formátech typických pro testy výkonnosti, generovat typické grafy, a bude dostupné jako webová služba. Součástí zadání byl seznam známých testů výkonnosti, určených pro demonstraci reálné použitelnosti práce.

Úvodní část práce se zabývá zavedením pojmů a analýzou požadavků. Úvod bohužel nevychází z konkrétních testů výkonnosti, jejich architektury, účelu, ani formy výsledků, ale soustřeďuje se podrobně na analýzu textu zadání. Důsledkem je neintuitivní zavedení pojmů, nejasná aplikace pojmů na zadané testy a rozebírání triviálních problémů (např. nezávislost kódu generujícího grafy na testech) místo problémů reálných (velký objem dat, doplňování výsledků o vypočtené statistiky).

Zadání si student vykládá tím nejjednodušším možným způsobem, např. formulaci “no databases” interpretuje jako zákaz použití nejen relačních databází pro uložení samotných výsledků měření, ale použití jakýchkoliv typů databází pro výsledky i metadata. Požadavek, že systém musí umět vrátit data též v původním formátu, si student interpretuje tak, že data musí též v původním formátu uchovávat.

Srozumitelnost úvodní části značně komplikuje systematicky špatné užívání určitých členů v anglickém jazyce, které navozuje (nesprávný) dojem, že používané pojmy byly dříve vysvětleny.

Podpora pro více úrovní důvěryhodnosti výsledků, požadovaná v zadání, je realizována nejtriviálnějším možným způsobem – uživatel si může přidat atribut.

Analýza ignoruje problémy které jsou uvedeny přímo v doporučené literatuře v zadání – například potřebu opakovaných spouštění testů a potřebu ukládat výsledky statistik k výsledkům.

Přehled existujících projektů nerozebírá důležité projekty, které jsou navíc zmíněny přímo v doporučené literatuře – zejména NIST Automated Benchmarking Toolset, ATL Benchmarking Tools a Open CORBA Benchmarking. Popis projektu RUBiS je nepřesný – výsledky, které generuje, lze těžko považovat za úložiště výsledků.

Text práce obsahuje popis konkrétního řešení a jeho implementace. V této části se student zaměřil zejména na parsování vstupních souborů, navrhl a implementoval konfigurovatelný parser pro XML a ASCII soubory, s jakými se setkal u některých zadaných testů. Bohužel chybí konkrétní příklady, jak parser a úložiště aplikovat na zadané testy (příklady jsou jenom v kódu, navíc jenom pro některé ze zadaných testů). Otázkou je také výkonnost zvoleného řešení, autor řešení aplikoval jen na triviálních příkladech a malých objemech dat, přestože na katedře jsou k dispozici desítky gigabajtů reálných dat.

Student naimplementoval jednoduchý systém pro reprezentaci a uložení metadat – metadata mají stromovou strukturu, která se serializuje na disk ve formě Java objektů. Využití databáze pro metadata by mohlo být škálovatelnějším a spolehlivějším řešením – zejména ale chybí příklad použití na reálné testy, včetně příkladů reálných dotazů a změření výkonnosti při vyhodnocování takových dotazů.

Implementace funguje jako webová služba, ale nemá požadovaný interface pro prohlížení.

Podpora pro generování grafů je omezená – implementace přímo neumožňuje grafy generovat, pouze generuje data v proprietárním XML formátu, ze kterých si uživatel může grafy sám vytvořit. Takto podporované grafy nejsou v práci popsány ani nejsou uvedeny odkazy na jejich popis v externím zdrojích. Grafy navíc nejsou implementovány dostatečně robustně – např. u histogramu se počet sloupců zadává v

kódu, graf pojmenovaný "density" není grafem odhadu hustoty pravděpodobnosti, jak by název nasvědčoval, ale pouze variantou grafu pojmenovaného "simple". Nejsou uvedeny žádné příklady grafů, není ověřeno, že poskytovaná data stačí k nakreslení vhodných grafů pro konkrétní testy.

Implementace funguje jako webová služba a demonstruje technické schopnosti autora naimplementovat webovou aplikaci s využitím jazyka Java a dostupných technologií.

Celkově práce působí nedotaženým dojmem, splňuje jenom část zadání, a to nejjednodušším možným způsobem, analýza není dostatečná a chybí případové studie pro zadané testy výkonnosti. V současné formě práci nedoporučuji k uznání za práci diplomovou, ale věřím, že vhodným rozšířením implementace a přepracováním textové části lze dosáhnout zajímavých výsledků.

V Praze, 29. ledna 2007

Tomáš Kalibera

