

# Oponentský posudek tezí disertační práce

**Student:** Mgr. Filip Nálepa

**Název práce:** Stream Processing for Computationally Intensive Applications

**Školitel:** prof. Ing. Pavel Zezula, CSc.

**Oponent:** RNDr. David Novák, Ph.D.

**Pracoviště:** Fakulta Informatiky, MU, Brno

Předložené teze disertační práce se zabývají problematikou efektivního zpracování proudových dat s důrazem na výpočetně náročné aplikace, jako je např. zpracování multimédií. Student v textu formuluje dva základní cíle budoucí disertační práce a to jsou

- vytvoření vhodného modelu pro analýzu efektivity systému pro distribuované zpracování proudových dat
- a optimalizace výkonu technik podobnostního vyhledávání v kontextu proudového zpracování dat.

Práce popisuje předběžné výsledky týkající se prvního cíle a student také předkládá dvě publikace na toto téma.

Práce nejdříve popisuje existující techniky pro danou oblast a stavební kameny, na kterých tyto techniky stojí. Tato část je dobře zpracovaná a strukturovaná a poskytuje základní vhled do problematiky modelování proudového zpracování dat, podobnostního vyhledávání a jejich optimalizace. V další kapitole jsou podrobněji rozebrány cíle disertační práce. Poslední kapitola obsahuje dosavadní výsledky práce, tedy zejména model pro analýzu výkonu systémů včetně popisu experimentálního vyhodnocení předběžných výsledků.

Text je dobře čitelný, má strukturu odpovídající vědecké práci a jazykově je na poměrně vysoké úrovni. V celém textu, a zejména v kapitole s výsledky, mi ale trochu chybí přesnější formalizace problému, zavedení jednoznačné notace a příp. schémata popisovaných modelů. K práci mám následující konkrétní otázky.

- Experiment popsáný v sekci 4.1.5 používá jedno konkrétní schéma mapování výpočetních úloh na danou hardwarovou konfiguraci (deployment model). Z textu není zřejmé, jak tento model vznikl - zda byl zvolen intuitivně a následně otestován nebo jestli je tento model výsledkem nějakých optimalizací v rámci CPN simulací.
- Tabulka 4.1 porovnává maximální a minimální časy zdržení, které vyplynuly ze CPN simulací, s reálně naměřenými hodnotami. Tato rozmezí zdržení jsou velmi velká (u některých ukazatelů až dva řády). Čím je tento velký rozsah způsoben a bude jedním z cílů optimalizace proudového zpracování také omezení této variability?

Celkově konstatuji, že práce má jasně a reálně stanovené cíle a očekávané výsledky považuji za disertabilní. Stanovený časový rámec je také rozumný. Předložené publikace odpovídají standardu v tomto oboru a proto lze práci akceptovat současně jako práci rigorózní. Obhajoby tezí se zúčastním.

V Brně, 28. dubna 2016

RNDr. David Novák, Ph.D.

Masarykova univerzita  
Fakulta informatiky

Datum: 28-04-2016

Č.j./E.č.: MU-15/18943/2016/337180/FI-5

Počet listů dokumentu: .....

Počet příloh a listů/sv.: ...../.....

Počet a druh nelist. příloh: ...../.....