



Posudek oponenta závěrečné práce

Oponent práce:	Ing. Lucie Svitáková
Student:	Bc. Tomáš Polačok
Název práce:	Návrh a implementace paralelního zpracování datových toků v projektu Manta
Obor / specializace:	Znalostní inženýrství
Vytvořeno dne:	31. května 2021

Hodnotící kritéria

1. Splnění zadání

- ▶ [1] zadání splněno
- [2] zadání splněno s menšími výhradami
- [3] zadání splněno s většími výhradami
- [4] zadání nesplněno

Zadání bylo splněno v plném rozsahu. Student se nejprve seznámil s potřebnými částmi projektu Manta, poté identifikoval možnosti paralelizace a další optimalizace, ty navrhl a implementoval s následným otestováním. Vzhledem ke komplexnosti jednotlivých algoritmů Manty byl vybrán jeden, nutno zmínit, že jeden z nejzásadnějších a rovněž nejsložitějších.

2. Písemná část práce

80/100 (B)

Písemný rozsah práce se nachází v doporučeném rozsahu pro diplomové práce na FIT ČVUT. Patří mezi delší práce, avšak všechny uvedené části jsou relevantní, v práci neshledávám žádné zbytečnosti. Nalezené nejasnosti byly se studentem diskutovány a objasněny, nejedná se o hrubé nepřesnosti. Vzhledem k délce práce bych si dovedla představit lepší členění do více kapitol, nicméně současné rozložení je vyrovnané a uvnitř jednotlivých kapitol logicky členěné, žádná část neobsahuje příliš málo ani příliš mnoho textu. Kapitoly jsou správně uvedeny popisem, co v dané kapitole čtenář nalezne.

Práce je psaná v anglickém jazyce. Chyby v jazyce nejsou časté. Co však může zhoršovat čitelnost, jsou občasné chybějící čárky ve větách. Někdy jsou věty natolik dlouhé, že zhoršují pochopení textu, což v kombinaci s občas chybějící čárkou dokáže čtenáře na chvíli zastavit. Vzhledem k náročnosti popisovaného algoritmu mohou tak některé části působit lehce zmátačně. Jinak je však práce čtivá a úroveň jazyka je odpovídající závěrečné práci. Typograficky překvapí umístění posledního odstavce druhé kapitoly, což však přisuzuji použité šabloně.

Práce nevychází z mnoha vědeckých článků, pár jich je zmiňováno v analýze. Převážná většina citovaných zdrojů jsou tak různé dokumentace a jiné webové stránky, což odpovídá praktickému charakteru práce. Nicméně v části analýzy bych uvítala více teorie (např. isomorfismus grafů, rozdělení grafu apod.).

U některých obrázků a tabulek by mohly být více vyčerpávající popisy. Diagram 2.11 je pak pochopitelný, nicméně měl by existovat explicitní stav či přechod, do kterého se po probuzení dostane "Wait for result" stav. Ráda bych však ocenila estetickou hodnotu všech diagramů, které student vytvořil. Ač k jejich tvorbě používal základní volně dostupný software, dokázal si s ním vyhrát natolik, že výsledek vypadá velmi profesionálně.

Použitý software a další díla byla použita v souladu s licenčními podmínkami.

3. Nepísemná část, přílohy 100 /100 (A)

Zdrojový kód je logicky členěn, metody mají rozumnou délku a názvy odpovídají svému účelu. Dokumentace kódu je pro druhou osobu dobře pochopitelná. Kód je zároveň psán tak, aby byl jednoduše rozšiřitelný. Všechny použité technologie odpovídají použité licenci.

4. Hodnocení výsledků, jejich využitelnost 100 /100 (A)

Ačkoli měl být dle zadání zhotoven pouze prototyp, výsledná implementace bude moci být přímo využita v nové architektuře projektu Manta. Co bych ráda velmi zvýraznila, je však samotné zlepšení výkonu, které se studentovi podařilo. Mergovací algoritmus, který může u zákazníka vzhledem k svému charakteru složitosti a vykonané práci běžet hodiny, student pro novou architekturu výrazně zrychlil. Pro specifický náročný typ podoperací je toto zrychlení dokonce až 5-ti násobné. Výstup této práce předčil očekávání zadavatele a je rozhodně velikým přínosem!

Celkové hodnocení 90 /100 (A)

Přestože písemná část obsahuje některé zmiňované nedostatky, jedná se spíše o drobnosti. Nepísemná část práce je na velmi dobré úrovni a zároveň je ihned aplikovatelná v chystané nové architektuře projektu Manta. Výsledná implementace přináší podstatné zrychlení jednoho z klíčových algoritmů, což je pro zadavatele důležité a předčilo jeho očekávání. Na práci je možné i dále navazovat s paralelizací a optimalizací dalších algoritmů. Celkově práci hodnotím jako velmi zdařilou za A.

Otázky k obhajobě

1) Popis mergování vrstev v části 1.5.3.2 není z daného odstavce dobře pochopitelný. Zkuste prosím tuto část vysvětlit více srozumitelně a jasně ukázat, proč není přístupováno do databáze.

Proč je v případě mergování resourců typ vrstvy (layer) ignorován?

2) Odpovídají testovací data velikostí a strukturou tomu, jak běžně vypadají data u zákazníků?

3) Paralelizaci diskutujete v rámci jednoho stroje a zároveň je uvedeno, že se v budoucnu plánuje i distribuované prostředí. Zkuste nastínit, jaké úpravy budou ve vaší implementaci pro takové prostředí potřeba, pokud takové jsou.

Instrukce

Splnění zadání

Posudte, zda předložená ZP dostatečně a v souladu se zadáním obsahově vymezuje cíle, správně je formuluje a v dostatečné kvalitě naplňuje. V komentáři uveďte body zadání, které nebyly splněny, posudte závažnost, dopady a případně i příčiny jednotlivých nedostatků. Pokud zadání svou náročností vybočuje ze standardů pro daný typ práce nebo student případně vypracoval ZP nad rámec zadání, popište, jak se to projevilo na požadované kvalitě splnění zadání a jakým způsobem toto ovlivnilo výsledné hodnocení.

Písemná část práce

Zhodnoťte přiměřenost rozsahu předložené ZP vzhledem k obsahu, tj. zda všechny části ZP jsou informačně bohaté a ZP neobsahuje zbytečné části. Dále posudte, zda předložená ZP je po věcné stránce v pořádku, případně vyskytují-li se v práci věcné chyby nebo nepřesnosti.

Zhodnoťte dále logickou strukturu ZP, návaznosti jednotlivých kapitol a pochopitelnost textu pro čtenáře. Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku ZP, viz Směrnice děkana č. 26/2017, článek 3.

Posudte, zda student využil a správně citoval relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami. Zhodnoťte, zda převzatý software a jiná autorská díla, byly v ZP použity v souladu s licenčními podmínkami.

Nepísemná část, přílohy

Dle charakteru práce se případně vyjádřete k nepísemné části ZP. Například: SW dílo – kvalita vytvořeného programu a vhodnost a přiměřenost technologií, které byly využité od vývoje až po nasazení. HW – funkční vzorek – použité technologie a nástroje, Výzkumná a experimentální práce – opakovatelnost experimentů.

Hodnocení výsledků, jejich využitelnost

Dle charakteru práce zhodnoťte možnosti nasazení výsledků práce v praxi nebo uveďte, zda výsledky ZP rozšiřují již publikované známé výsledky nebo přinášející zcela nové poznatky.

Celkové hodnocení

Shrňte stránky ZP, které nejvíce ovlivnily Vaše celkové hodnocení. Celkové hodnocení nemusí být aritmetickým průměrem či jinou hodnotou vypočtenou z hodnocení v předchozích jednotlivých kritériích. Obecně platí, že bezvadně splněné zadání je hodnoceno klasifikačním stupněm A.