



# Hodnocení vedoucího závěrečné práce

**Vedoucí práce:** Ing. Tomáš Oberhuber, Ph.D.  
**Student:** Tat Dat Duong  
**Název práce:** Implementace B-stromů na GPU  
**Obor / specializace:** Teoretická informatika  
**Vytvořeno dne:** 2. července 2021

## Hodnotící kritéria

### 1. Splnění zadání

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

Všechny body zadání byly splněny. Myslím si, že implementace B-stromů na GPU patří mezi jedny z náročnějších algoritmů pro GPU. Autor pracoval na zadání velmi intenzivně, ale i tak se nevyhnul tomu, že strávil hodně času na náročném hledání chyby ve své implementaci. Hledání chyb v kernelech pro GPU je velmi náročné, protože není možné používat běžné nástroje jako debugger nebo výpis zásobníku. Navíc jde o ladění kódu, zpracovávaného stovkami vláken. Díky tomu a celkové vyšší náročnosti tématu musel autor žádat o odložení termínu odevzdání a na sepisování vlastního textu měl méně času.

### 2. Písemná část práce

90/100 (A)

Práce je psána anglicky a to angličtinou na velice dobré úrovni. Text je dobře strukturovaný, snad bych jen prohodil první a druhou kapitolu, což je ale drobnost. Jak již jsem zmínil, autor měl na samotné sepisování méně času, než by bylo potřeba. Vůbec tím nemyslím, že by psaní nechal na poslední chvíli, ale jde o náročné téma a popsat jej pořádně je také náročnější. Dle mého názoru se to projevuje na čtvrté kapitole popisující implementaci, která by si zasloužila rozepsání do více detailů, a také část týkající se testů by mohla obsahovat testy založené na porovnání se strukturou typu `std::set` a ne pouze `std::map`. Tyto testy by pak mohly podložit nebo vyvrátit tvrzení, že slabší výkonost autorovy implementace je způsobena ne zcela optimálními přístupy do paměti.

### 3. Nepísemná část, přílohy

100/100 (A)

Autor je velice dobrý programátor schopný psát dobře čitelný kód. Zároveň ukázal velikou snahu o vytvoření kvalitního kódu, jak z pohledu funkcionality tak i výkonosti. Výsledný kód by se v budoucnu měl stát součástí knihovny TNL, a proto je důležité, aby se jednalo opravdu o robustní a spolehlivou implementaci. Autor tuto snahu projevoval sám od sebe, což bylo velmi potěšující. Jak již jsem zmínil, v určitý moment autor narazil na chybu ve své implementaci, kterou bylo velice náročné najít. Hledání chyb v paralelních programech patří mezi jedny z nejtěžších. Autor si sám napsal systém pro ukládání stavu stromu a následné krokování ve webovém prohlížeči. Díky tomu se mu chybu skutečně podařilo najít. Tento postup považuji za jasný důkaz toho, že autor má širší vzdělání v IT a umí ho správně využít za účelem vytvoření kvalitního kódu.

### 4. Hodnocení výsledků, jejich využitelnost

100/100 (A)

Výsledný kód bude brzy součástí knihovny TNL. Mohlo by se tak jednat o jednu z nejnázve dostupných implementací B-stromů pro GPU a zároveň také jednu z nejlépe otestovaných. Jak je vidět v samotné práci, implementaci B-stromů na GPU je zatím velmi málo a z toho některé navíc nefungují správně. B-stromy na GPU jsou tématem posledních let a mohou velice napomoci při portování algoritmů, které se doposud mohly zdát být nevhodné pro běh na GPU. Autor navíc ukazuje, že velmi dobrého výkonu lze dosáhnout i bez pomocných linků uvnitř stromu, na kterých je postavena v současnosti asi nejpokročilejší implementace B-stromu na GPU. Osobně považuji tento výsledek za velice zajímavý. Pokud by se provedlo ještě podrobnější otestování a případně doladění výkonu, mohlo by možná jít o publikovatelný výsledek.

### 5. Aktivita studenta

- ▶ [1] **výborná aktivita**
- [2] velmi dobrá aktivita
- [3] průměrná aktivita
- [4] slabší, ale ještě dostatečná aktivita
- [5] nedostatečná aktivita

Student se věnoval tématu velmi aktivně a samostatně. Celou problematiku nastudoval sám. Dělal jsme pravidelné konzultace každý týden, kde jsem studentovi spíš udával směr, kterým se dále ubírat. A pochopitelně jsem mu pomáhal s psaním textu, jak je u bakalářských prací vždy potřeba.

### 6. Samostatnost studenta

- ▶ [1] **výborná samostatnost**
- [2] velmi dobrá samostatnost
- [3] průměrná samostatnost
- [4] slabší, ale ještě dostatečná samostatnost
- [5] nedostatečná samostatnost

Autor pracoval velmi samostatně. V jeden moment jsem se rozhodl, že mu asi budu muset nějak pomoci najít již zmiňovanou chybu v jeho kódu, ale nakonec i to zvládl sám.

## Celkové hodnocení

95 /100 (A)

Celkově lze říci, že autor si zvolil velmi náročné téma. Sám jsem občas pochyboval, jestli jsem toto téma neměl vypsát spíš jako diplomovou práci, a ani jsem neočekával, že by se autorovi podařilo vše dotáhnout tak daleko, jak se nakonec stalo. Sice se nepodařilo dosáhnout stejné efektivity jako u implementace, ze které jsme vycházeli, ale autor věnoval opravdu hodně úsilí, aby vlastní kód optimalizoval co nejlépe - dokonce se ponořil až do assembleru pro GPU. Nutno ale podotknout, že implementace, ze které jsme vycházeli, nefunguje správně pro větší objemy dat a není napsána pro obecné datové typy. Trojnásobný rozdíl výkonu může být na GPU často otázkou jednoduché optimalizace, na kterou jsme třeba jen prozatím nepřišli. Autor navíc ukazuje, že dobrého výkonu lze dosáhnout i bez pomocných interních linků, což považuji za velmi hodnotné zjištění pro výzkum podobných datových struktur na GPU. Autorova implementace se stane součástí knihovny TNL a tím by mohla napomoci rozšířit třídu algoritmů, které lze efektivně provozovat na GPU. Jediné, kde lze uvažovat o slabším hodnocení, je text samotné práce, který je na dobré úrovni, ale mohl by být dotažen ještě dále, a chybějící testy se strukturou `std::set`. To bych zároveň autorovy položil jako dotaz do diskuze k obhajobě:

Jakého urychlení je možné dosáhnout při ukládání pouze klíčů a žádných hodnot, tj. při uvážení datové struktury podobné `std::set`?

## 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. 52/2021, 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.

### Aktivita studenta

V souvislosti s průběhem a výsledkem práce posudte, zda byl student během řešení aktivní, zda dodržoval dohodnuté termíny, jestli své řešení průběžně konzultoval a zda byl na konzultace dostatečně připraven.

### Samostatnost studenta

V souvislosti s průběhem a výsledkem práce posudte schopnost studenta samostatně tvůrčí práce.

### 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.