

Posudek dizertační práce

Student: Ing. Ladislav Dobrovský

Oponent: doc. RNDr. Jiří Filipovič, Ph.D.

Pracoviště opONENTA: Ústav výpočetní techniky, Masarykova Univerzita

Název práce: GPU akcelerované metaheuristiky ve vybraných úlohách globální a kombinatorické optimalizace

Shrnutí práce

Práce se zaměřuje na GPU akceleraci HC12 algoritmu s využitím tabu listu a diferenciální evoluce s více ostrovními populacemi. Přínos práce je demonstrován na problémech z oblasti spojité a kombinatorické (QAP, SAT) optimalizace. Práce se nejprve zabývá vysvětlením principu a existujících implementací optimalizačních metaheuristik a optimalizačních úloh. Následně pojednává o široké škále metod využívaných v HPC, s větším zaměřením na GPU akceleraci a efektivní využití vícejádrových CPU. Dále již práce prezentuje autorovy výsledky: implementaci HC12 s Tabu listem a diferenciální evoluce, implementaci konkrétních optimalizačních úloh, a experimentální evaluaci implementovaných metod.

Zhodnocení práce

Práce se zaměřuje na téma, které je aktuální, a její přínos je netriviální (dosažené zrychlení pomocí GPU je významné, vzorné je také škálování autorovy CPU implementace). Práce celkově prokazuje schopnost autora implementovat poměrně komplexní kód a hluboké porozumění problematice optimalizace kódu s ohledem na specifika hardware a omezení existujících překladačů.

Z více kritického pohledu práci sráží fakt, že autor (především v části GPU a CPU implementace) příliš rychle přechází k vysvětlení svého řešení, aniž by detailně argumentoval proč je toto řešení vhodné či nezbytné a jaké slabiny by měl naivní přístup (je sice diskutován problém spojitého přístupu do paměti a využití explicitního cacheování ve sdílené paměti, bez obecnější diskuse toho jak celý algoritmus funguje, jaká omezení jsou fundamentální a jaká omezení naopak můžeme překonat, je pro čtenáře obtížné uvažovat o alternativách autorova přístupu). Celá práce se tak více blíží učebnímu textu (části popisující metaheuristiky, optimalizační problémy a metody využívané v HPC) a dokumentaci software (popis autorovy implementace), než standardní vědecké publikaci obsahující zevrubný popis problému a jeho specifik (v tomto případě možností paralelizace a akcelerace) a vymezení autorova přístupu oproti existujícím řešením.

Vyhodnocení je provedeno poctivě z hlediska hodnocení konvergence a přesnosti implementovaných metod. Významným cílem práce bylo zrychlení metaheuristik na GPU. To bylo vyhodnoceno pouze pro HC12 a QAP. Bylo by nepochybně zajímavé vyhodnotit i další optimalizační úlohy.

Poznámky k textu

- v kapitole 2.2.2 by bylo vhodné uvést pseudokód prezentovaného algoritmu, bez něj je obtížnější pochopit následující texty, i interpretovat obrázek 1
- popis gridu v kapitole 4.1.4 je zavádějící – popisuje pouze distribuované prostředí typu folding@HOME, avšak velké množství gridových aplikací pracuje se vzájemně propojenými clustery poskytujícími relativně homogenní prostředí
- kapitola 4.3.3 chybně uvádí, že AVX instrukce jsou 128bitové
- 5.1.1 – algoritmus 2 by zasloužil obsáhlejší komentář, není jasné, co vyjadřují proměnné X_a , X_b , navíc se jedná o poměrně složitý kód, slovní okomentování algoritmu by zvýšilo jeho srozumitelnost
- kapitola 5.2 uvádí, že "Pro implementaci čistě na GPU není základní verze DE příliš vhodná, často je však vhodné využít GPU pro urychlení výpočtu optimalizované funkce." -- chybí zde rozbor proč není DE vhodná, zároveň chybí vysvětlení proč pomůže konstantní počet jedinců v různých populacích

Otázky k obhajobě

1. Algoritmus 1 (vlození řádku do nepřímo seřazeného pole) je prováděn paralelně, zatímco Algoritmus 2 (nalezení řádku v nepřímo seřazeném poli) sériově (paralelizmu je pak dosaženo spuštěním více instancí tohoto algoritmu). Jak funguje synchronizace a předávání dat mezi těmito dvěma algoritmy?
2. Dá se stanovit teoretický limit rychlosti implementovaných algoritmů (tj. minimální počet aritmetických operací či přenesených bytů na jednu iteraci optimalizační metody)? Pokud to lze, jak blízký je dosažený výkon teoretickému maximu použitých CPU a GPU?
3. Liší se nějak výkon či škálování v případě spojitě optimalizace či SAT?

Závěr

Práce se zaměřuje na aktuální téma GPU akcelerace metaheuristik, přináší novou implementaci tabu prohledávání pro HC12 algoritmus a optimalizovanou diferenciální evoluci s více ostrovními populacemi. Slabším aspektem práce je prakticky absentující porovnání s GPU akcelerací podobných problémů a relativně stručný popis širšího kontextu prezentovaných řešení (nevýhody naivního přístupu, fundamentální omezení rychlosti algoritmů). Vytvořená implementace je dobře provedená, přináší významné zvýšení výkonu a dobře škáluje, dokazuje autorovu schopnost kreativně řešit výzkumný úkol. Popisem metaheuristik a technologií využívaných v HPC autor prokazuje vynikající šířku i hloubku znalostí této problematiky.

Celkově přes výše zmíněné výhrady práce prokazuje schopnost doktoranda samostatné pedagogické i tvůrčí vědecké činnosti, je na slušné odborné úrovni a splňuje požadavky kladené na dizertační práci. Práci proto **doporučuji** přijmout k obhajobě.