

Posudek na diplomovou práci

Parallel Data-processing on GPGPU

Bc. Radim Vansa

Hlavním cílem posuzované práce je analýza časově náročných operací, které jsou typické pro běžné databázové systémy a prozkoumání možností jejich akcelerace pomocí moderních grafických karet, které již zvládají i obecné výpočty. Autor se zaměřil především na problematiku operací třídění založeného na porovnávání a průniku dvou množin (databázovou operaci join). Zmíněná problematika je poměrně dobře prozkoumána a existuje řada publikací na dané téma, avšak dostupné články se zaměřují především na jednotlivé dílčí problémy. Práce nabízí čtenáři ucelený přehled různých možných přístupů a jejich výkonnostní srovnání. Práce je doplněna o prototypovou implementaci všech srovnávaných metod a experimentální data naměřená na soudobém hardware.

Práce je rozdělena přehledně do 6 kapitol. Po obecném úvodu následuje stručný a ucelený přehled vlastností dnešních GPU a paralelního frameworku OpenCL, který autor použil ve své implementaci. V práci chybí srovnání s alternativním frameworkem CUDA, avšak vzhledem k jeho restriktivním vlastnostem (zejména omezení na hardware jednoho výrobce) je výběr OpenCL opodstatněný. Ve třetí kapitole je popsána metodologie prováděných testů. Metodologie je netradičně představena před samotnými algoritmy a metodami, aby mohl autor lépe provázat popis algoritmů se souvisejícími experimentálními výsledky. Ve čtvrté kapitole jsou rozebrány nejběžnější třídící algoritmy, jejich možnosti paralelizace na GPU a experimentální výsledky. Pátá kapitola se pak stejným způsobem věnuje problému průniku dvou množin. Autor prozkoumal jak situace, kdy je alespoň jedna z množin setříděna (typické pro indexované tabulky), stejně jako situace, kdy jsou obě množiny nesetříděné (např. pokud vzniknou jako výsledky jiné mezioperace).

Celková analýza problému (zejména průniku dvou množin) je velice podrobná. Kromě srovnání několika algoritmů autor řeší i situace, kdy jsou vstupní množiny různě velké, obsahují různé počty společných prvků a také, když není možné je jednorázově umístit do paměti GPU. Přestože výkonnostní testy jsou z principu problematické a vždy vázané na konkrétní implementaci, autor se s nimi vypořádal korektně a nabízí velmi rozumné srovnání, které by mohlo být návodem pro programátora implementujícího databázový systém akcelerovaný na GPU.

Práce je psaná v anglickém jazyce pouze s drobnými chybami, které nebrání srozumitelnosti. Prototypová implementace je kvalitní a použitelná pro další experimenty.

Celkově práci považuji za velmi zdařilou a doporučuji ji k obhajobě.

15.5.2012

RNDr. Martin Kruliš
KSI, MFF UK