SADRŽAJ

Univerzitet u Kragujevcu
TEHNIČKI FAKULTET
ČAČAK

Milan Radosavljević

PARALELNO RAČUNANJE KORIŠĆENJEM GRAFIČKOG PROCESORA

Čačak
2012. godine

Razvoj računarskih komponenti je doživeo veliki napredak u protekloj deceniji. Sa napretkom računarskih resursa porasli su i zahtevi za programima koji ih mogu na najefektivniji način koristiti. Pošto su specijalizovani računari, sa više desetina ili stotina procesora skupi, nastale su mreže računara sličnih karakteristika koje efektivno oponašaju rad tih računara i poznate su kao računarski klasteri. Shodno tome klaster predstavlja skup računara istih ili sličnih karakteristika koji su povezani konvencionalnim mrežnim interfejsima.

Problem pisanja programa koji bi se izvršavali na paralelnim računarima je prisutan već decenijama. On se obično rešava na dva načina:

- *paralelizacijom postojećih sekvencijalnih programa ili*
- *pisanjem novih, u osnovi paralelnih, programa.*

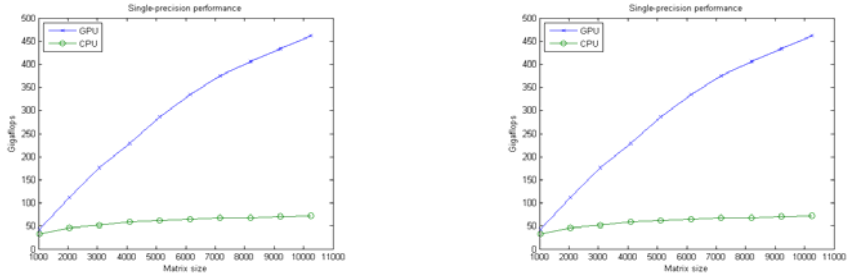
U oba slučaja javlja se isti problem komunikacije sa procesorima i optimizacije njihove aktivnosti. Razvojem standarda koji propisuju način komunikacije među procesorima, taj problem postaje rešiv, ali još uvek nedovoljno prihvatljiv, jer njegova implementacija zahteva dodatno zalaganje programera. Kao korak ka prevazilaženju novonastalog problema javlja se pojava skupa unapred isprogramiranih funkcija koje su već optimizovane za komunikaciju sa višeprocesorskim sistemima i omogućavaju programeru da se koncentriše samo na rešavanje problema koji je pred njim, a da pri tom ne brine o načinu ugradnje standarda i komunikaciji među procesorima.

Razvoj modernih grafičkih procesora je došao do tačke kada se računska moć ovih procesora može uporediti ili čak prevazilazi moć centralnih procesorskih jedinica. S tim u vezi, neminovno je došlo do toga da se grafički procesori koriste i za zadatke koji nisu striktno vezani za obradu grafike. Može se pokazati da je ubrzanje pri paralelnoj obradi na grafičkom procesoru nekoliko puta veće od paralelne obrade na centralnom procesoru. Ovakav pristup korišćenju grafičkog procesora naziva se GP GPU (General Purpose GPU).

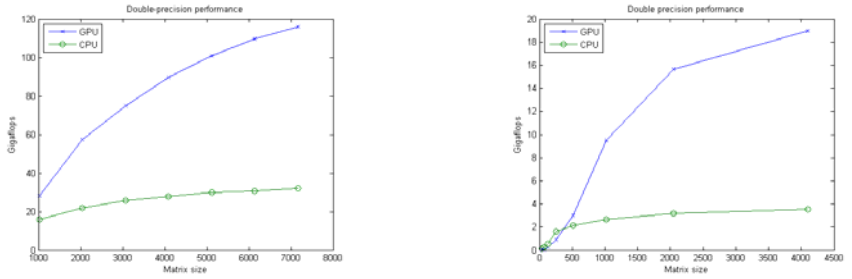
CUDA (skraćenica za Compute Unified Device Architecture) je arhitektura za paralelno računanje koju je razvila firma nVidia. CUDA je dostupna na nVidia GPU-ovima kroz standardne programske jezike. C for CUDA je ekstenzija programskog jezika C koja uključuje naredbe za kontrolu GPU-a. U CUDA programu, kod za CPU može biti u C++ jeziku, dok se za GPU može koristiti samo C. CUDA pruža programerima pristup nativnim setovima instrukcija i memoriji GPU-a. CUDA podržava čitav niz programskih interfejsa uključujući i OpenCL kao i najnoviji Microsoft API, DirectCompute koji je promovisan sa Windows 7 operativnim sistemom. Takođe podržani su i ostali programski jezici poput: Python, Fortran, Java i Matlab-a. Koristeći CUDA-u, programeri mogu razvijati aplikacije namenjene izvršavanju na GPU-u, kojem više nije potrebno pristupati kroz skupove instrukcija za obradu grafike. Tipična CUDA aplikacija je heterogena. Sastoji se od CPU i GPU funkcija. GPU i CPU izvode različite vrste koda. CPU izvodi glavni program i šalje GPU-u zadatke u obliku kernel funkcija. Više različitih kernel funkcija može biti deklarisano i pozvano, ali se samo jedan kernel može izvršavati u jednom trenutku.

U radu je izvršeno poređenje brzina izvršavanja programa pisanog za MATLAB razvojno okruženje za rešavanje sistema linearnih jednačina na Intel Pentium baziranim računarima. Korišćen je algoritam na principu "levog matričnog deljenja" tj. **x = A\b**. *Podrška za paralelno računanje ostvarena je na bazi GF108 i GF110 grafičkih procesora firme nVidia. U MATLAB postoji podrška za CUDA tehnologiju, a konkretni algoritam se realizuje preko bibliotečne funkcije –* ***paralldemo_gpu_backslash()***. *Argument ove funkcije maksimalni deo sistemske memorije koji je dostupan centralnom i grafičkom procesoru.*

Na slikama su prikazane performanse paralelnog računanja u slučaju jednostruke i dvostruke preciznosti.



Performanse jednostruke preciznosti (GTX570 vs GT440)



Performanse dvostruke preciznosti (GTX570 vs GT440)