Protokol z profilovania aplikácie na výpočet smerodajnej odchýlky

1 Úvod

Podľa zadania projektu 2 predmetu IVS bolo potrebné vytvoriť matematickú knižnicu pre neskoršie využitie v kalkulačke. Táto matematická knižnica má byť však zároveň využitá pre vytvorenie programu schopného počítať smerodajnú odchýlku pomocou vzorca:

$$s = \sqrt{\frac{1}{N-1}\left(\sum_{i=1}^{N} x_i^2 - N\overline{x}^2\right)}$$

Na tomto vytvorenom programe má byť následne vykonaný profiling pre 10, 100 a 1000 vstupných hodnôt. Z merania je potrebné zistiť, v ktorých miestach program trávi najviac času a na čo sa pri optimalizácií kódu zamerať.

2 Testovacie prostredie

CPU: Intel i7-3632QM @ 2.2GHz x 4
RAM: 7.7 GiB
OS: Linux Mint 18.2 Cinnamon 64-bit
Profiler: Valgind 3.11.0 (nástroj Callgrind)
Analýza: KCachegrind 0.7.4

3 Testovacie dáta

Analýza bola vykonaná pre 3 sady čísel. Tieto sady mali 10, 100 a 1000 prvkov. Čísla boli vygenerované náhodným generátorom čísel v rozsahu 0-1000. Tieto čísla boli zaokrúhlené na 10 desatinných miest. Výsledky vypočítané programom vo všetkých prípadoch splnili testy s presnosťou na 6 miest.

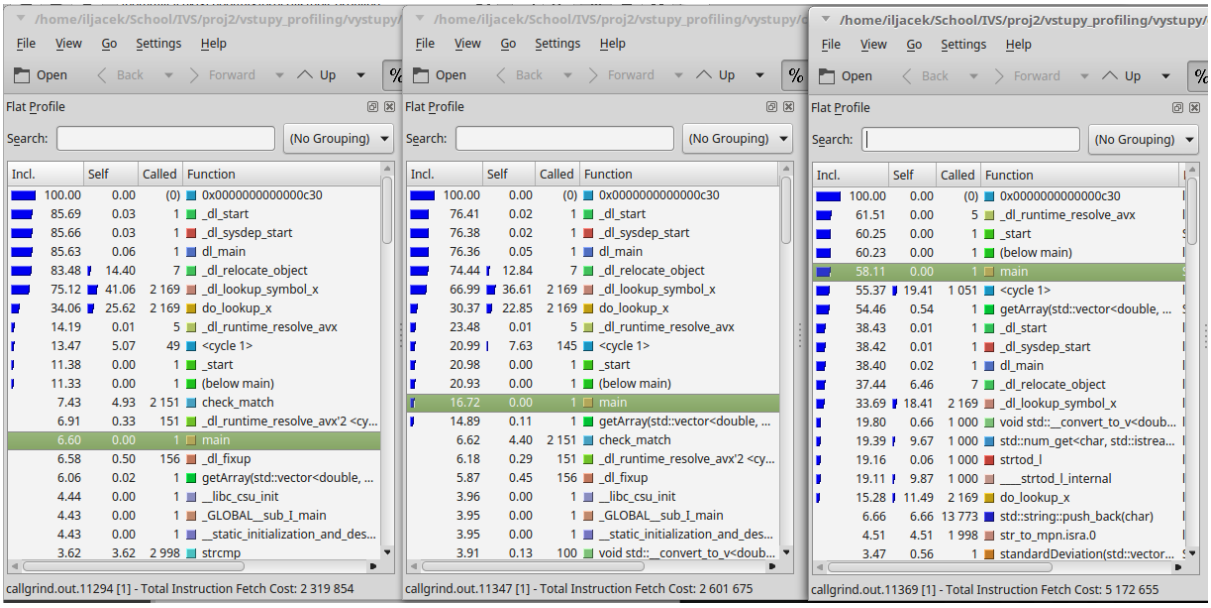
3 Výsledky

V prípade 10 vzorkov trávil program funkciou main len 6,6% času, v prípade 100 vzorkov to bolo 16,72% a v prípade 1000 vzorkov to bolo 58,11%. Vidíme teda že značná časť času bola spotrebovaná pre inicializáciu prostredia. Čo sa týka samotného behu programu vo všetkých troch prípadoch bola väčšina času strávená funkciou getArray slúžiacou na parsovanie dát (približne 90%). Čo sa týka ostatných funkcií z knižníc standard_deviation.h a mathematical_library.h, percento času potrebného pre ich výkon s rastúcim počtom vzorkov narastal. V prípade funkcie gerAverage to bolo 1.9% a v prípade funkcie standardDeviation 5.97%. Najviac volaní (2000 volaní pre 1000 vorkov) mala funkcia add z matematickej knižnice, ktorá však zaberala v prípade 2000 volaní len 0.53% z výpočetného času.

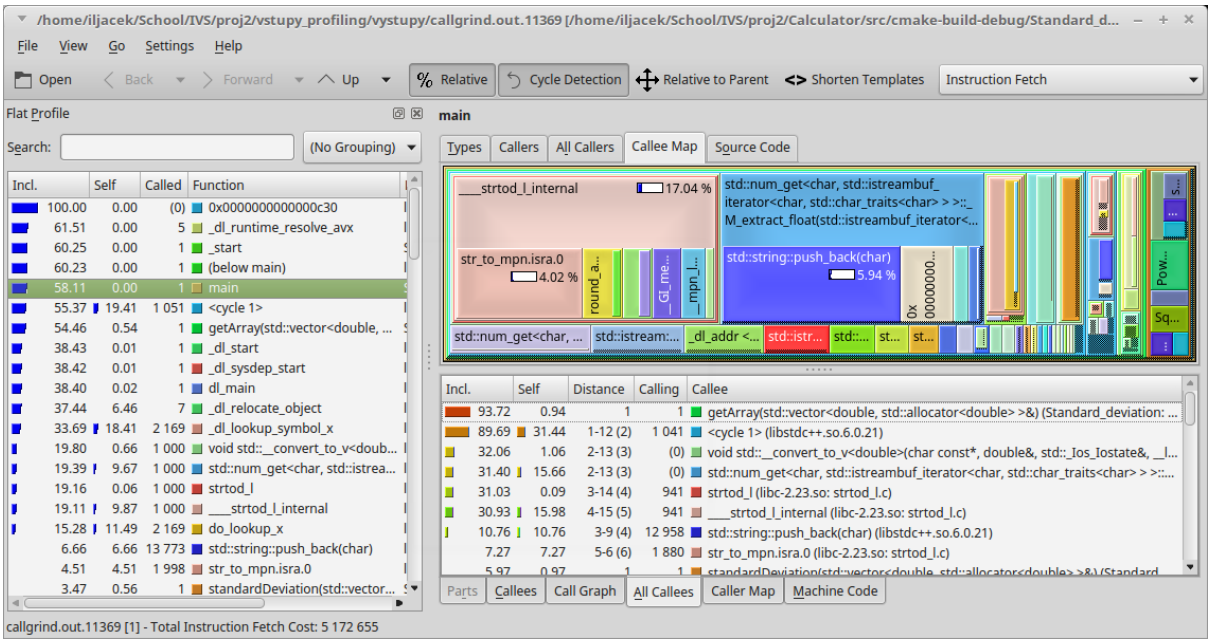
4 Záver

Z výstupu profilingu vidíme že program trávil najviac času parsovaním vstupných dát. Pre prípadné optimalizácie je teda vhodné zamerať sa práve na túto časť.

5 Prílohy



Obr5.1: Porovnanie profilingu pre 10, 100 a 1000 vzorkov (z ľava).



Obr 5.2: Detailejší pohľad na najnáročnejšie procesy vo funkcií main pre 1000 vorkov.