

BOINC – przetwarzanie rozproszone a pieniądze roztrwonione?

5 stycznia 2016

W poprzednim wpisie snułem rozważania o projektach naukowych, w których aktywnie uczestniczyć może każdy zainteresowany. Skupiłem się na programach poświęconych badaniu kosmosu, ale to nie jedyna tematyka wśród której można wybierać. Projekty znajdujące się pod adresem Zooniverse.org [1] wymagają aktywnego uczestnictwa, co oznacza potrzebę poświęcenia swojego czasu na analizę danych. Dziś chciałbym zapoznać was z nauką dla leniwych, czyli udziałem w projektach nie wymagających aktywnego uczestnictwa, a jedynie zainstalowania odpowiedniego oprogramowania na swoim komputerze lub smartphonie. Mowa tu o BOINC [2] (Berkeley Open Infrastructure for Network Computing) czyli sieciach komputerowych przetwarzania rozproszonego. Ten nieco enigmatycznie brzmiący skrót, oznacza udostępnianie swojego komputera, a dokładnie jego mocy obliczeniowej, do analizy danych w wybranym projekcie naukowym. Tu już nie ma potrzeby wpatrywania się w monitor, porównywania zdjęć i klikania. Wystarczy wybrać projekt na rzecz którego chcemy poświęcić gigaherce naszego procesora i gotowe! To tak jakby twój komputer był częścią superkomputera. Analiza wielkich ilości danych na pojedynczym superkomputerze nie jest tania, a nierzadko trzeba długo czekać w kolejce, aby mieć sposobność skorzystania z takiej potężnej maszyny obliczeniowej. Tak więc jakiś mózgowiec wpadł na pomysł, aby skorzystać z mocy obliczeniowej uwięzionej w milionach domowych komputerów osobistych; i to niemal za darmo. Genialne!

NAUKA DLA LENIWYCH?

Idea, jaka przyświecała autorom projektu, jest bardzo

wzniosła. W zamyśle oprogramowanie ma działać w ramach „wygaszacza ekranu”, czyli czasu, kiedy komputer nie jest przez jego właściciela używany. To dobre rozwiązanie dla wszelakich instytucji badawczych, uniwersytetów czy laboratoriów, gdzie istnieje wewnętrzna sieć z dużą ilością komputerów. Pozwala to na analizę danych bez zastosowania wydajnej ale drogiej jednostki centralnej. Oczywiście, jeżeli nie ma potrzeby szybkiego i wydajnego obliczania danych. Od strony technicznej przetwarzanie rozproszone polega na rozsyłaniu pakietów do poszczególnych komputerów – klientów. Pakiet taki jest wysyłany najczęściej tylko do jednego komputera i jest mu przydzielony okres ważności, czyli czas w jakim powinien być przeanalizowany. Jeżeli komputer nie zdąży z obliczeniami, pakiet jest usuwany u danego klienta i wysyłany do innego komputera. Czynność ta jest powtarzana do czasu, aż dane zostaną przeliczone i zwrócone do serwera. Dopiero teraz użytkownik, którego komputer zdołał przeliczyć dany pakiet w całości, jest nagradzany punktami. Widać tu zasadniczą różnicę między przetwarzaniem rozproszonym, a projektami z serii „Zooniverse”. W BOINC całe liczenie przypada komputerowi, więc nie ma możliwości, aby ten się pomylił. Zasadniczym minusem jest szybkość takiej analizy, albowiem należy czekać, aż pakiety zostaną policzone przez komputery klienckie. W „Zooniverse” to człowiek podejmuje decyzje, a że jest omylny, to te same dane są wielokrotnie analizowane przez wiele osób. To oczywiście tylko ogólne porównanie, albowiem idee, jakie stoją za poszczególnymi projektami, są odmienne. „Zooniverse” zostało stworzone ze względu na specyfikę analizowanych danych, z którymi algorytmy komputerów nie radzą sobie najlepiej. To co łączy te przedsięwzięcia, to możliwość oddania części zadań w ręce zwykłych użytkowników internetu, którzy być może poza samym dobrowolnym poświęceniem swojego czasu i komputera, postanowią poszerzyć swoją wiedzę w jakimś temacie z pożytkiem dla nich samych, a może nawet ogółu społeczeństwa.

Być może niektórzy z was kojarzą lub nawet sami brali, tudzież

nadal biorą, udział w programie SETI? [3] To przedsięwzięcie mające na celu poszukiwanie pozaziemskich cywilizacji drogą radiową. Program ma już swój dość spory staż (początek przypada na koniec lat 1950.), niemniej jednak jak dotąd nie przyniósł żadnego pozytywnego wyniku. Innymi słowy, nie odebrano żadnego radiowego sygnału od obcej cywilizacji, a w każdym razie nie takiego, który potrafimy odczytać □ Kiedy kilka lat temu brałem udział w tym programie, w czasie bezczynności komputera uruchamiał się „profesjonalny” wygaszacz ekranu z mnóstwem wykresów. Fajnie to wyglądało, ale wymiernych skutków nie miało żadnych – jak dotychczas □ Przedsięwzięcie z czasem ewoluowało i rozrastało się o kolejne projekty badawcze, których obecnie jest kilkanaście. Tak samo ja w „Zooniverse”, projekty rozpoczynają się i kończą. Jedne cieszą się ogromną popularnością, a inne już nie. Swojego czasu byłem uczestnikiem między innymi takich programów, jak Asteroids@home, Einstein@home, LHC@home, Milkyway@home, Radioactive@Home (ciekawy polski projekt), SETI@home. Już po samych nazwach można się domyśleć, czego dotyczą. Pełna lista wraz z opisem jest dostępna między innymi na stronach Boinc.berkeley.edu [4] oraz Boincatpoland.org [5].

RACHUNEK ZA PRĄD

I teraz uwaga! Zanim się zdecydujesz na uczestnictwo w BOINC dobrze przeczytaj „instrukcję obsługi” klienta, czyli oprogramowania, które masz zamiar zainstalować na swoim komputerze. Najważniejsze to dobrze go skonfigurować, aby później nie żałować kiedy przyjdzie rachunek za prąd! Tak, tak! W świecie w którym żyjemy, nic nie ma za darmo. Chcesz pomóc w rozwiązaniu jakiegoś naukowego „problemu” nie wysilając swojego mózgu? To płać! Niekiedy bardzo słono. Jak napisałem powyżej: udostępniasz moc obliczeniową swojego komputera, a ta idzie w parze ze zużyciem elektryczności, wszak wszystko we wszechświecie kręci się wokół energii pod wszelaką postacią.

W opcjach oprogramowania, jakie zainstalujesz, masz wiele

opcji do skonfigurowania. Między innymi kiedy komputer ma liczyć, czyli czy ma liczyć cały czas w tle, czy tylko w czasie, gdy jest bezczynny. Możesz ustawić procent mocy procesora, jaki ma być poświęcony na obliczenia, ilość pamięci RAM, ilość miejsca na dysku twardym, priorytety, którym projektom przyznajesz pierwszeństwo w obliczeniach itp. Pamiętaj, że najważniejsze to kiedy i z jaką mocą ma liczyć procesor. Pewnie niektórym z Was już zaświeciła się czerwona lampka – „A co ja z tego będę mieć?”. No cóż. Jeżeli nie wystarczy Ci satysfakcja, że masz swój mały udział w postępie naukowym ludzkości, to być może usatysfakcjonują Cię punkty, jakie są przyznawane za poświęconą moc komputera. W zależności od projektu, są okresy premiowane większą ilością punktów i tak dziś projekt X nagradza swoich uczestników mnożnikiem x4, a jutro projekt Y, będzie nagradzać x2. I tu zaczyna się coś, co uważam za przerost formy nad treścią.

PUNKTY. TYSIĄCE SETKI TYSIĘCY, MILIONÓW PUNKTÓW!

Decydując się na udział w BOINC zacząłem przeglądać internet w celu znalezienia dodatkowych, nieoficjalnych informacji o tym projekcie. Bardzo szybko zorientowałem się, że są ludzie, a nawet całe zespoły reprezentujące drużyny, uniwersytety czy nawet kraje w punktowej rywalizacji. Wkopując się głębiej w kolejne internetowe pokłady informacji, wkrótce trafiłem na fora poświęcone przetwarzaniu rozproszonemu. Czytam, czytam i zastanawiam się, o co tu tak naprawdę chodzi. Otóż okazuje się, że nie samym CPU człowiek żyje. Dawniej szybki komputer równał się szybki procesor główny. Te czasy już dawno minęły. Teraz się liczy GPU, czyli karta graficzna. Potrafi ona policzyć z nieporównywalnie większą wydajnością dane, z którymi CPU będzie się pocić znacznie dłużej. Specyficzna architektura z jaką zostały zaprojektowane niektóre karty graficzne powoduje, że ich wydajność jest wielokrotnie większa w liczeniu specyficznych danych. I tak oto już niemal normą stało się pisanie oprogramowania liczącego specjalnie pod karty graficzne a nawet projektowanie kart graficznych,

których głównym zadaniem nie jest wyświetlanie grafiki (!!!), ale właśnie liczenie. Kopanie bitcoinów, łamanie haszy czy właśnie liczenie z BOINC jest oparte w dużej mierze na kartach graficznych.

No dobrze, ale co z tego? – zapytacie. Są osoby, które konfiguruja swoje komputery specjalnie pod software obliczeniowy, w tym przypadku BOINC. Komputery takich osób składają się niejednokrotnie z czterech kart graficznych najwyższej klasy, które potrzebują duuuuużo watów energii elektrycznej. Myślisz, że jesteś kozak, bo twój PC bez problemów pociągnie BF4 na ULTRA detalach z zasilaczem 600 W? Schowaj się lepiej z Twoim przerośniętym kalkulatorem, bo są komputery które zjadają 1200W na rozgrzewkę. I tu dochodzimy do sedna idei, z jaką został pomyślany BOINC i każdy uczestnik musi sobie sam odpowiedzieć na pytanie – „Po co ja to robię?”. Czy to nadal chęć uczestniczenia w jakimś symbolicznym wymiarze w programie (mniej lub bardziej) naukowym, czy może celem samym w sobie są tak naprawdę... punkty? Kto tak naprawdę jest beneficjentem tej idei i czy jest ona dla Ciebie warta rachunku za prąd wyższego nawet o kilkaset złotych miesięcznie? Czy nie lepiej te pieniądze przeznaczyć na jakiś faktyczny program naukowy, gdzie widać w wymierny sposób efekty podarowanych pieniędzy? A może faktycznie lepiej przerobić swojego kompa na koparkę bitcoinów, ewentualnie dorobić sobie do kieszonkowego łamiąc hashe za pieniądze? To już zostawiam pod rozwagę każdemu uczestnikowi BOINC. Tymczasem ustawcie procesory na 5% w idle'u.

Autorstwo: Piotr Piekarczyk

Źródła: CosmicBulletin.wordpress.com, WolneMedia.net

PRZYPISY

[1] www.zooniverse.org

[2] <http://boinc.berkeley.edu>

[3] <http://setiathome.ssl.berkeley.edu>

[4] <http://boinc.berkeley.edu/projects.php>

[5] http://www.boincatpoland.org/wiki/Lista_projektów