

Projekt interdyscyplinarny

fizyka-informatyka

PRZETWARZANIE ROZPROSZONE

Streszczenie

Uczniowie korzystają z Internetu zapoznając się z projektami naukowymi realizowanymi na zasadzie obliczeń rozproszonych. Wybierają jeden z projektów (Einstein@home, ClimatePrediction, cosmology@home, milkyway, orbit, cleanenergy...), instalują oprogramowanie i biorą udział w projekcie. Zbierają informacje na temat celów projektu, tworzą sprawozdanie z opisem projektu i przedstawiają go innym. Poznają i opisują metodę przetwarzania rozproszonego.

Czas realizacji

Ok. 4-6 tygodni

Etapy realizacji

- 1 Lekcja wprowadzająca. Przedstawienie projektów, podział na grupy (1 x 45 minut)
- 2 Praca zespołowa. Zbieranie informacji i udział w projekcie (4-6 tygodni)
- 3 Lekcja podsumowująca. Przedstawienie sprawozdań grup z udziału w projekcie (1 x 45 minut)

Podstawa programowa

Poziom podstawowy

Udział w projektach przetwarzania rozproszonego umożliwi realizację kilku celów kształcenia. W zakresie informatyki będzie to przede wszystkim wykorzystanie komputera do poszerzania wiedzy i rozwijania zainteresowań. Cele realizowane w zakresie innych przedmiotów zależą od wybranego projektu. Zawsze będzie to jednak związane

z posługiwaniem się informacjami zawartymi w opracowaniach popularnonaukowych, a czasem nawet naukowych.

Cele kształcenia – wymagania ogólne

Fizyka

IV. Posługiwanie się informacjami pochodzącymi z analizy przeczytanych tekstów (w tym popularnonaukowych).

Informatyka

IV. Wykorzystanie komputera oraz programów i gier edukacyjnych do poszerzania wiedzy i umiejętności z różnych dziedzin oraz do rozwijania zainteresowań.

V. Ocena zagrożeń i ograniczeń, docenianie społecznych aspektów rozwoju i zastosowań informatyki.

Treści nauczania

Fizyka

Oprócz wiedzy z wybranych działów fizyki, uczeń:

8) przedstawia własnymi słowami główne tezy poznanego artykułu popularnonaukowego z dziedziny fizyki lub astronomii.

Informatyka

3. Uczeń wykorzystuje technologie komunikacyjno-informacyjne do komunikacji i współpracy z nauczycielami i innymi uczniami, a także z innymi osobami, jak również w swoich działaniach kreatywnych.

7. Wykorzystywanie komputera i technologii informacyjno-komunikacyjnych do rozwijania zainteresowań, opisywanie zastosowań informatyki, ocena zagrożeń i ograniczeń, aspekty społeczne rozwoju i zastosowań informatyki. Uczeń:

1) opisuje szanse i zagrożenia dla rozwoju społeczeństwa, wynikające z rozwoju technologii informacyjno-komunikacyjnych;

3) zapoznaje się z możliwościami nowych urządzeń i programów związanych z technologiami informacyjno-komunikacyjnymi, zgodnie ze swoimi zainteresowaniami i potrzebami edukacyjnymi.

ETAP 1

Lekcja wprowadzająca.

Przedstawienie projektów, podział na grupy

(45 minut)

Nauczyciel przedstawia metodę przetwarzania rozproszonego. Uczniowie korzystają z Internetu zapoznając się z projektami naukowymi realizowanymi tą metodą. Wybierają jeden z projektów, tworzą grupy, które będą realizować wybrany projekt, organizują pracę w grupie.

Nauczyciel powinien zapoznać się z dwoma podstawowymi stronami poświęconymi przetwarzaniu rozproszonemu:

<http://www.boincatpoland.org> – polska strona platformy BOINC;

<http://boinc.berkeley.edu> – główna strona platformy BINC (ang.);

<http://distributedcomputing.info> – zbiór informacji na temat wszystkich projektów realizowanych metodą przetwarzania rozproszonego (ang.).

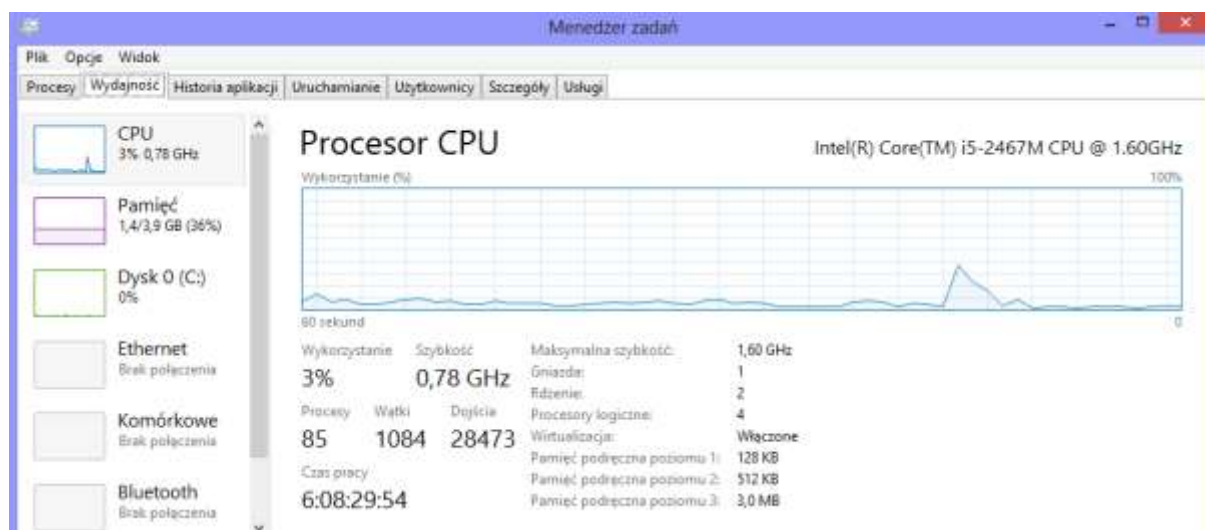
Nauczyciel może w trakcie lekcji wykorzystać prezentację – **OblRozproszonePrzykłady.ppt** oraz listę aktywnych projektów – **ListaAktywnychProjektów.xls**.

Przebieg zajęć

Wprowadzenie (15 minut)

Nauczyciel przedstawia metodę przetwarzania rozproszonego.

Dobrym wstępem jest przedstawienie Menedżera Windows. Najszybciej można go uruchomić naciskając trzy klawisze Ctrl Alt i Del. Widać, że komputer wykonuje równocześnie wiele procesów oraz że wykorzystanie mocy procesora jest zazwyczaj niewielkie. Rzadko kiedy procesor zajęty jest w więcej niż 50%. Ma więc sporo wolnego czasu. A jeśli odejdziemy od włączonego komputera, będzie miał go jeszcze więcej. Aż się prosi, żeby ten czas wykorzystać.



Najstarszy i najpopularniejszy projekt realizowany metodą obliczeń rozproszonych to Seti@home. Poszukiwania cywilizacji pozaziemskich zaczęły się dawno temu. Polegają one na prowadzeniu nasłuchu specyficznych sygnałów radiowych przy pomocy radioteleskopów. Tak powstał program SETI (Search for ExtraTerrestrial Intelligence). W roku 1999 wszedł on w kolejny etap rozwoju. Powstał projekt SETI@home (Seti w domu), który wykorzystuje wolną moc obliczeniową komputerów podłączonych do Internetu, do przetwarzania próbek danych, pochodzących z radioteleskopu w Arecibo. W projekcie SETI@home może wziąć udział każdy posiadacz komputera podłączonego do Internetu, po zainstalowaniu na swoim komputerze bezpłatnego oprogramowania, dostępnego na serwerach projektu. O szczegółach projektu można przeczytać na stronach: <http://setiathome.berkeley.edu>.

Na czym polega przetwarzanie rozproszone?

Jest wiele problemów, których nie udało się jeszcze rozwiązać ze względu na ogromną liczbę obliczeń do wykonania lub z powodu wielkiej liczby danych, które należałoby przetworzyć, aby znaleźć rozwiązanie. Wraz z powstaniem komputerów pojawiła się nadzieja, że część tych problemów da się rozwiązać – wystarczy do obliczeń zaprząć maszyny. Jednak czasem nawet superkomputer nie wystarczy. Niektóre projekty liczyłyby się dziesiątki, a nawet setki lat. Naukowcy wpadli więc na inny pomysł. Jeśli nie można zadaniem obciążyć jednego komputera, to może podzielić problem na małe zadania i dać je do wykonania wielu słabszym maszynom. Nie każdy problem da się tak „pokawałkować”, ale dla wielu jest to możliwe. Po połączeniu w całość rozwiązań tych

małych zadań mamy szansę rozwiązać problem. Komputery są wszędzie, na przykład w naszych domach. W dodatku istnieje sposób, aby mogły porozumiewać się z centrum wspólnie wykonywanego projektu – komunikacja przez Internet. Pozostaje tylko jedno – zaprosić właścicieli komputerów do współpracy.

Obecnie nawet największe na świecie superkomputery są budowane na zasadzie równoległego połączenia tysięcy procesorów podobnych do tych, które znajdują się w komputerach domowych. Sumowanie się mocy obliczeniowej wielu małych jednostek prowadzi do bardzo szybkich, wydajnych obliczeń.

Od strony technicznej, realizacja projektów przetwarzania rozproszonego okazała się stosunkowo prosta. Organizatorzy projektu przygotowują oprogramowanie, które wykonuje obliczenia. Udostępniony w sieci program może ściągnąć na swój komputer każdy użytkownik łączący się z Internetem. Oprogramowanie jest zazwyczaj tak skonstruowane, aby jego działanie jak najmniej przeszkadzało właścicielowi komputera. Na przykład program pracuje tylko wtedy, gdy nie pracuje właściciel – czyli działa jak wygaszacz ekranu. Właściciele komputerów o większej mocy obliczeniowej, mogą pozwolić takim programom na pracę ciągłą w tle innych czynności, które wykonuje komputer. Jest to sposób, w którym moc obliczeniowa komputera jest wykorzystywana w sposób efektywny.

Właściciel komputera rejestruje się jako użytkownik projektu. Oprogramowanie zainstalowane na jego komputerze łączy się z serwerem projektu i pobiera do opracowania porcję danych. Po wykonaniu obliczeń – przesyła wyniki do centrum i pobiera następną porcję i tak dalej. Wyniki uczestnika projektu są sprawdzane, wysyłane do dalszej obróbki i odnotowywane. Może on na stronach projektu sprawdzić ile danych już opracował, które miejsce zajmuje w rankingu wszystkich uczestników. Może też z nimi podyskutować na forum poświęconym projektowi. Wyniki i kolejne etapy obliczeń są publikowane na stronach projektu, aby każdy z uczestników wiedział, na jakim etapie jest problem, w którego rozwiązaniu uczestniczy.

Definicję przetwarzania rozproszonego i więcej wyjaśnień można znaleźć na stronie internetowej polskiego zespołu zaangażowanego w projekty przetwarzania rozproszonego: <http://www.boincatpoland.org>. Na tej witrynie znajduje się też sporo informacji o różnych projektach i o tym, jak wziąć w nich udział.

W grudniu 2013 roku na platformie BOINC było około 70 aktywnych projektów o charakterze przetwarzania rozproszonego, do których można się przyłączyć. „Około”, gdyż w sieci wszystko jest zmienne – niektóre projekty zawieszają chwilowo swoją działalność (np. z powodu zmiany oprogramowania) inne dopiero startują, jeszcze inne kończą... Pełną listę aktywnych projektów można znaleźć na stronie (w języku angielskim) <http://distributedcomputing.info/projects.html>.

Większość projektów posługuje się własnym oprogramowaniem, napisanym specjalnie pod kątem projektu. Powstają jednak platformy (jak BOINC), na bazie których można realizować różne projekty. Wspólna platforma oprogramowania łączy projekty o podobnej tematyce i pozwala uczestnikowi na branie udziału w kilku projektach równocześnie. Najbardziej popularna już od kilku lat jest wspomniana platforma BOINC – Berkeley Open Infrastructure for Network Computing (*Otwarta Infrastruktura Przetwarzania Rozproszonego Berkeley*).



BOINC <http://boinc.ssl.berkeley.edu> (ang.)

<http://www.boincatpoland.org> (pol.)

BOINC umożliwia zarówno uczestnictwo w wielu projektach (to wariant dla zwykłego użytkownika), jak i przygotowanie własnego projektu pracującego pod tą platformą (to dla tych, którzy wiedzą, co i jak należy zbadać). Grupa BOINC w Polsce współpracuje z Wikipedią. W tej encyklopedii można znaleźć opisy wielu projektów.

Całkowita moc obliczeniowa realizowana przez projekty platformy BOINC w 2013 roku to ok. 24 PFlops, a największy superkomputer osiąga niewiele więcej – ok. 34 PFlops (dane ze strony Top500: <http://www.top500.org>).

Wybór projektów (15 minut)

Nauczyciel poleca uczniom wyszukanie interesujących projektów. Omawia kilka przykładowych projektów na platformie BOINC.

Poszukiwanie inteligencji pozaziemskiej – SETI (Search for ExtraTerrestrial Intelligence).

- <http://setiathome.berkeley.edu> (niektóre teksty po polsku, większość po angielsku)
- <http://www.setiathome.pl> (pol.)



Projekt SETI@home (Seti w domu), przetwarza próbki danych, pochodzących z radioteleskopu w Arecibo. Zyskał dużą popularność między innymi dzięki pomysłowi, którym zainteresowała się duża grupa ludzi (obecnie, po 14 latach działania, w projekcie pracuje prawie 1,5 miliona uczestników, w tym 62 tys. zespołów na ponad 3,5 mln komputerów). Zespół projektu SETI@home rozpropagował ideę przetwarzania rozproszonego na specjalnej platformie, którą udostępnił wszystkim chętnym do współpracy.

Długofalowe prognozowanie zmian klimatycznych



- www.climateprediction.net (ang.)
- Wikipedia (<http://pl.wikipedia.org/wiki/ClimatePrediction.net>) (pol.)

Projekt, którego celem jest prognozowanie zmian klimatycznych w długich okresach. W swoich obliczeniach posługuje się metodą Monte Carlo opracowaną i pierwszy raz zastosowaną przez wybitnego polskiego matematyka Stanisława Ulama. Służy ona do modelowania matematycznych procesów zbyt złożonych, aby można było przewidzieć ich wyniki za pomocą standardowych obliczeń. Do takich właśnie problemów należy prognozowanie pogody – bardzo wiele czynników wpływa na całokształt zmian klimatycznych. Istotną rolę w tej metodzie odgrywa losowanie wielkości charakteryzujących proces i przeprowadzanie obliczeń dla różnych wyników losowania. Po zebraniu dosta-

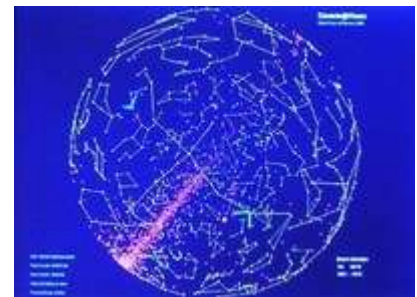
tecznie dużej liczby takich wyników można zestawić je z obserwowanymi wynikami doświadczalnymi, potwierdzając lub negując słusność poczynionych w całym badaniu założeń.

Einstein@home – poszukiwanie fal grawitacyjnych

- <http://einstein.phys.uwm.edu/> (ang.)
- Wikipedia: <http://pl.wikipedia.org/wiki/Einstein@home> (pol.)

Projekt ma na celu poszukiwanie fal grawitacyjnych, których istnienie przewidział Albert Einstein w ogólnej teorii

względności. Einstein@Home jest projektem przeznaczonym do przetwarzania danych z Obserwatoriów Fal Grawitacyjnych LIGO w USA i GEO 600 w Niemczech. Obserwatoria rejestrują sygnały wytwarzane przez obracające się szybko neutronowe gwiazdy, znane jako pulsary. Naukowcy sądzą, że niektóre pulsary nie mogą być idealnie kuliste, a skoro tak, powinny emitować charakterystyczne fale grawitacyjne.



rosetta@home – poszukiwanie kształtu łańcuchów białkowych

- <http://boinc.bakerlab.org/rosetta>
- Wikipedia: <http://pl.wikipedia.org/wiki/Rosetta@home> (pol.)



Celem projektu jest przewidzenie kształtu, w jaki białko układa się w naturze. Białka są cząsteczkami o budowie łańcuskowej, złożonymi z aminokwasów, które można sobie wyobrazić jako ogniwka łańcucha. W skład łańcuchów białkowych może wchodzić 20 różnych aminokwasów. Połączone aminokwasy oddziałują wzajemnie na siebie co powoduje, że cały łańcuszek przyjmuje kształt nazywany „złożeniem”. W projekcie poszukuje się „złożenia” o najniższej energii – takie białka tworzą się w naturze.

primeGrid – poszukiwanie liczb pierwszych



- <http://www.primegrid.com> (ang.)
- <http://pl.wikipedia.org/wiki/PrimeGrid> (pol.)

Grupa projektów poświęconych poszukiwaniu liczb pierwszych. Proces szukania liczb pierwszych jest zazwyczaj dzielony na 2 etapy:

- przesiewanie – czyli szybką eliminację z dużego zakresu tych liczb, których podzielność łatwo wykazać,
- testowanie pierwszości wybranych kandydatów.

Jednym z podprojektów (Prime Sierpinski) jest poszukiwanie liczb pierwszych Sierpińskiego.

Word Community Grid – grupa projektów (dawniej osobna platforma)



<http://www.worldcommunitygrid.org> (ang.)

[http://pl.wikipedia.org/wiki/World Community Grid](http://pl.wikipedia.org/wiki/World_Community_Grid) (pol.)

Projekty aktywne w 2013 roku:

- FightAIDS@Home, przeliczając jednostki tego projektu pomagamy w szukaniu lekarstwa na AIDS
- Human Proteome Folding – faza 2, kolejny „białkowy” projekt, w chwili obecnej przewiduje struktury białek
- Help Conquer Cancer, badania prowadzące do lepszego rozumienia raka
- Help Fight Childhood Cancer
- Discovering Dengue Drugs – Together – faza 2, szukanie leku na chorobę Dengue’a
- Computing for Clean Water
- The Clean Energy Project – faza 2
- Drug Search for Leishmaniasis
- GO Fight Against Malaria
- Say No to Schistosoma

Uczniowie wyszukują interesujący ich projekt.

Praca w grupach (15 minut)

Uczniowie znajdują interesujące ich projekty i zbierają na ich temat informacje.

Z pomocą nauczyciela tworzą zespoły – każdy zespół zajmuje się wybranym projektem. Już na tej lekcji mogą z udziałem nauczyciela zainstalować projekt na komputerze w pracowni. Dzielą się również rolami. W każdym zespole powinni być uczniowie:

- zajmujący się informatyczną stroną projektu, instalacją oprogramowania, śledzeniem postępu obliczeń (statystyki),
- zajmujący się dziedziną projektu – opisem zagadnień, które są obliczane,
- zajmujący się dokumentowaniem pracy – tworzeniem sprawozdania z udziału grupy w projekcie.

ETAP 2

Praca zespołowa. Zbieranie informacji i udział w projekcie

(4-6 tygodni)

Po zainstalowaniu oprogramowania projektu, zespoły śledzą przebieg obliczeń. Równocześnie zbierają informacje o realizowanym projekcie, jego celach i uzyskanych dotąd wynikach. Wszystkie projekty są dobrze udokumentowane. Naukowcy przygotowujący projekt starają się przedstawić go w sposób prosty, aby zachęcić ludzi do udziału.

Rola nauczyciela ogranicza się do kontrolowania postępów pracy zespołów.

ETAP 3

Lekcja podsumowująca. Przedstawienie sprawozdań grup z udziału w projekcie

(45 minut)

Ta lekcja może mieć charakter otwarty, można zaprosić na nią innych uczniów, a nawet rodziców. Każdy zespół przedstawia sprawozdania z realizowanego projektu. Sprawozdanie powinni zaprezentować wszyscy członkowie grupy np. pokazując różne aspekty projektu. Uczniowie powinni podać statystykę pracy swojego zespołu (liczbę przeliczonych jednostek).

Ocena pracy zespołu może być dokonana przez nauczyciela, ale lepiej, gdy dokona jej komisja złożona z nauczycieli różnych przedmiotów oraz uczniów starszych klas.

Dostępne pliki

ListaAktywnychProjektów.xls – spis aktywnych projektów realizowanych na platformie BOINC.

OblRozproszonePrzykłady.ppt – prezentacja przykładowych projektów realizowanych metodą obliczeń rozproszonych.