

Celem projektu jest udostępnienie w polskich centrach Komputerów Dużej Mocy nowych architektur obliczeniowych, które pozwolą na prowadzenie badań wymagających przetwarzania obliczeniowego o skali dotychczas zarezerwowanej dla nielicznych ośrodków zagranicznych.

Realizacja projektu umożliwi rozwój polskiej infrastruktury obliczeń wielkiej skali (HPC) dla bezpośredniego wykorzystania przez środowisko naukowe. Przeprowadzone inwestycje będą znacząco rozszerzać i wzbogacać dostępne zasoby centrów Komputerów Dużej Mocy, zdominowane w ciągu ostatnich lat przez rozwiązania klastrów.

Ponadto, w ramach projektu POWIEW utworzone zostały nowe grupy badawcze złożone z wysoko wykwalifikowanych analityków i programistów systemów wieloprocesorowych oraz naukowców

współpracujących z projektem. Powołane Centra Kompetencji pracują nad optymalizacją istniejących modeli obliczeniowych oraz tworzeniem nowych rozwiązań dedykowanych realizacji obliczeń wielkich wyzwań nauki i techniki na dużych systemach wieloprocesorowych. Systemy obliczeniowe użyte do realizacji projektu są najnowocześniejszymi rozwiązaniami generacji petaskalowej (o wydajności rzędu 10^{15} obliczeń na sekundę) ze szczególnym uwzględnieniem zastosowań wymagających zrównoległych obliczeń gruboziarnistych.

Projekt jest realizowany przez konsorcjum stworzone przez Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego UW (ICM), Akademickie Centrum Komputerowe CYFRONET AGH oraz Poznańskie Centrum Superkomputerowo-Sieciowe IChB PAN (PCSS).

System IBM Blue Gene/P w ICM

W lutym 2011 roku w ICM udostępniony został dla użytkowników system IBM Blue Gene/P. System składa się z jednej szafy obliczeniowej, która mieści 4096 rdzeni obliczeniowych oraz 4TB pamięci operacyjnej. Użytkownicy mają również do dyspozycji zintegrowaną z systemem przestrzeń pamięci masowej o pojemności ok. 170 TB. Główne przeznaczenie systemu to wysoce zrównoległone obliczenia w dziedzinach neuroinformatyki, dynamiki molekularnej, chemii oraz kosmologii. System otrzymał nazwę greckiego Boga wiatru południowego – Notosa.



Procesory Notosa chłodzi powietrze. Następny superkomputer stanie w serwerowni ICM na jesieni i będzie chłodzony strumieniami zimnej wody.

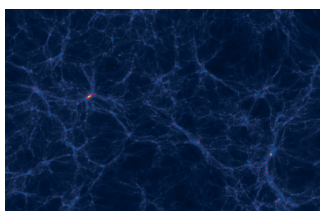
wykorzystywanego oprogramowania Gadget-3 dla mniejszych rozmiarów symulacji. Na podstawie otrzymanych wyników przygotowany został plan obliczeniowy dla Warsaw Universe.

Czy wiesz, kto był pierwszym użytkownikiem systemu Blue Gene/P w Polsce?

Pierwszym użytkownikiem systemu Blue Gene/P w Polsce była Helena Głąbska, doktorantka Instytutu Biologii Doświadczalnej im. M.Nenckiego PAN. Uruchamianie na systemie Notos obliczenia dotyczyły modelowania lokalnych potencjałów polowych w pętli korowo-wzgórzowej i realizowane były w ramach projektu badawczego prowadzonego w POWIEW przez dr. Daniela Wójcika.

Symulacja kosmologiczna „Warsaw Universe”

15 marca 2011 roku w ramach działalności w projekcie POWIEW dr Wojciech A. Hellwing uruchomił na komputerze Notos symulację kosmologiczną o wielkości 1024^3 cząstek. Wyniki symulacji opublikowane zostaną, wraz z wizualizacjami najciekawszych galaktyk, na specjalnie przygotowanej stronie internetowej. Symulacje kosmologiczne tych wielkości wykorzystywane są bardzo często przez naukowców do analizy dynamiki powstawania różnego rodzaju struktur kosmologicznych. Realizowana w projekcie POWIEW symulacja otrzyma, zgodnie ze zwyczajami panującymi w tej dziedzinie nauki, nazwę „Warsaw Universe”. Obliczenia wykorzystują 4096 rdzeni komputera Notos, a planowany czas symulacji to ponad 5 dni. Uruchomienie obliczeń poprzedzone zostało analizą wydajności



Tysiące rdzeni obliczeniowych pracuje nad bardzo obiecującym modelem kosmologicznym. W tym samym czasie inne maszyny ICM opracowują najnowszą prognozę pogody.

- Gorącym tematem ostatnich dni jest organizacja tutorialu „Best Practices to run applications in HPC environments” na konferencję Parallel Processing and Applied Mathematics (PPAM), która odbędzie się w dniach 11-14 września 2011 w Toruniu.
- 9 maja odbyła się długo oczekiwana konferencja prasowa ICM i IBM dotycząca zakupionej infrastruktury obliczeniowej.
- 16 czerwca 2011 w PCSS odbędzie się szkolenie „Wprowadzenie do równoległych bibliotek algebry liniowej”.
- 15 czerwca 2011 w ICM odbędzie się szkolenie „Uruchamianie i optymalizacja kodów na architekturze Blue Gene/P”.
- Na stronie projektu POWIEW ukazał się film przedstawiający instalację Notosa.

Celowość testów skalowalności

Maciej E. Marchwiany

Podczas pracy z pakietami obliczeniowymi bardzo ważna jest wiarygodność otrzymywanych wyników. Ale nie jest to jedyny element istotny w symulacjach komputerowych. Ważne są również ilość oraz sposób wykorzystania zasobów komputerowych, w szczególności czas obliczeń. Dla dużych i złożonych badań to właśnie kosztowność czasowa staje się kluczowa. Każdy użytkownik oprogramowania naukowego oczekuje możliwie dużo wyników w możliwie krótkim czasie. Czas oczekiwania można skracać na dwa sposoby:

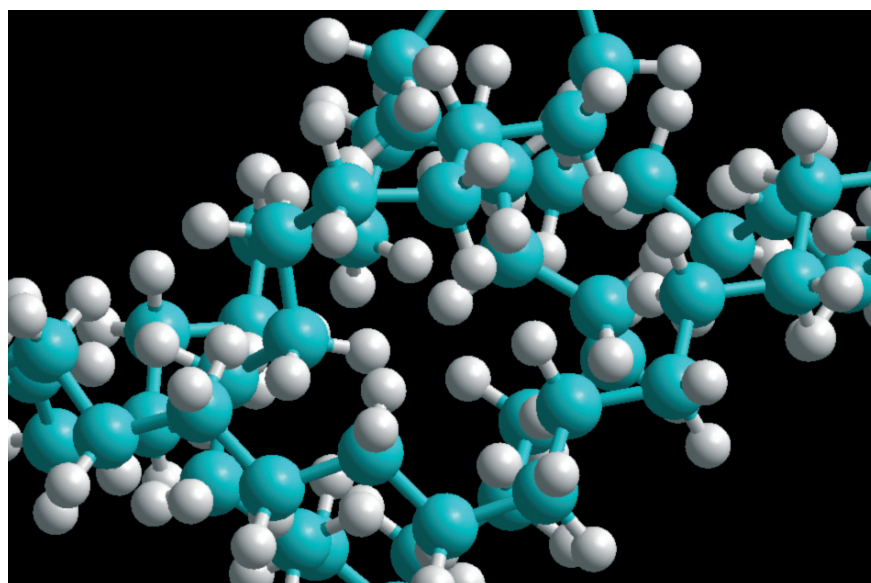
- używając komputera o większej mocy pojedynczego rdzenia obliczeniowego,
- uruchamiając aplikację na wielu rdzeniach obliczeniowych jednocześnie.

Użytkownik związany z konkretnym ośrodkiem obliczeniowym rzadko ma możliwość przeniesienia obliczeń na nową maszynę o większej mocy pojedynczego rdzenia, ale najczęściej nie ma większych trudności z dostępem do większej liczby procesorów, czy rdzeni obliczeniowych. Ponadto niemal wszystkie rozpowszechnione pakiety naukowe pozwalają na prowadzenie obliczeń równoległych w zależności od stosowanych algorytmów na kilku, kilkudziesięciu, lub nawet kilku tysiącach procesorów.

Niestety zwiększanie liczby stosowanych rdzeni nie skraca czasu obliczeń do zera. Dla każdego kodu pojawia się pewna liczba procesorów, powyżej której czas obliczeń się nie skraca, a może się nawet wydłużyć. Ponadto graniczna liczba procesorów jest bardzo trudna do oszacowania. W każdym przypadku konieczna jest bardzo dobra znajomość używanego kodu i stosowanych w nim algorytmów, co w przypadku komercyjnych kodów jest niemożliwe, a dla kodów otwartych zajmuje ogromną ilość czasu i jest bardzo pracochłonna.

Łatwiejszym sposobem badania wydajności pakietów dla wielu rdzeni jest wykonanie testów skalowalności. Pozwalają one na oszacowanie optymalnej liczby rdzeni dla danej aplikacji. Należy przy tym pamiętać, iż kod z racji swojej budowy, może różnie zachowywać się dla różnych układów, w zależności tak od ich wielkości, jak i symetrii czy złożoności oraz dla różnych rodzajów prowadzonych obliczeń.

Wykonanie testów skalowalności wydaje się koniecznością podczas pracy z nową aplikacją, dlatego tak ważne i potrzebne są wszelkiego rodzaju publikacje i raporty odnośnie skalowalności popularnych kodów. W kolejnych numerach postaramy się przedstawić wyniki testów skalowalności dla popularnych kodów. Mamy nadzieję, że pozwolą one zaoszczędzić czas i zwiększyć wydajność prowadzonych badań.



▲ Obliczenia *ab initio* dla dużych układów molekularnych zazwyczaj są bardzo czasochłonne i wymagają dużej ilości pamięci. Dlatego konieczne jest prowadzenie obliczeń na wielu procesorach jednocześnie.

Piszą o nas i mówią

[...] "Dzięki ogromnemu postępowi technologii komputerowych realne staje się podejmowanie wyzwań naukowych jeszcze niedawno daleko wykraczających poza granice wyobraźni. Poznanie natury układów o ekstremalnym poziomie złożoności, od skali atomowej do kosmologicznej, zaczyna być możliwe na drodze modelowania obliczeniowego. Rozwój technologii otwiera równocześnie nowe wyzwania wobec twórców algorytmów obliczeniowych, bez czego możliwości sprzętowe byłyby wykorzystywane jedynie w ograniczonym stopniu. To wszystko otwiera nowe wielkie szanse i stawia ambitne wyzwania przed zespołami ICM i naszymi współpracownikami. Możliwość współpracy z IBM, światowym liderem technologii w tej dziedzinie, inspiruje nas dodatkowo do intensywnego zaangażowania w tworzenie nowych rozwiązań i pozyskiwania nowej wiedzy" - powiedział Profesor Marek Niezgódka, Dyrektor ICM. [...]

Gazeta.pl Komputer w biurze 09.05.2011

[...] IBM Polska dostarczy sprzęt warty w sumie 15,2 mln zł dla Interdyscyplinarnego Centrum Modelowania Matematycznego i Komputerowego UW (ICM) w ramach części tzw. "Projektu POWIEW" (Program Obliczeń Wielkich Wyzwań Nauki i Techniki), poinformowali przedstawiciele obu instytucji na konferencji prasowej. [...]

Wyborcza.biz 09.05.2011

[...] W istniejącym na Uniwersytecie Warszawskim już od 18 lat ICM prowadzi się badania z wykorzystaniem złożonych, wieloprocessorowych obliczeń. Wraz z innymi ośrodkami, ICM wdraża na potrzeby polskiej nauki rozwiązania gridowe i chmurowe. Przedstawiciele ICM podkreślają, że trudno nawet mówić o "polskiej nauce", bo badane problemy to najpoważniejsze wyzwania, z którymi mierzą się zespoły z całego świata, a polscy naukowcy nierzadko prowadzą tu badania w ramach międzynarodowych projektów. [...]

Gazeta.pl Technologie 09.05.2011

[...] Wśród licznych obszarów aplikacyjnych system IBM Blue Gene/P wykorzystywany będzie przy modelowaniu potencjałów elektrycznych w mózgu, co pozwoli na weryfikację stawianych od kilkunastu lat hipotez dotyczących dynamiki przetwarzania informacji czuciowej. Superkomputer projektu POWIEW używany będzie w ICM też przy rozwijaniu nowych algorytmów oraz metod obliczeniowych dla architektur następnej generacji, które pojawią się w użyciu przed 2020 r. Będzie to kontynuacja współpracy ICM i IBM, w której budowa nowoczesnych architektur komputerowych i doskonalenie sposobów ich wydajnego wykorzystania stanowi klucz do skutecznego podejmowania największych wyzwań współczesnej nauki i techniki[...]

Computerworld 09.05.2011

Najważniejsze publikacje zagraniczne:

<http://www.hpcwire.com/hpcwire/2011-05-11/>
<http://classic.cnbc.com/id/42987992>

Audycja radia TOK FM:

<http://www.tok.fm/TOKFM/0,89523.html>