



Instytut Automatyki i Informatyki Stosowanej
Politechnika Warszawska
ul. Nowowiejska 15/18, 00-665 Warszawa

Warszawa, 12 luty 2009 r.

Dr hab. inż. Ewa Niewiadomska-Szynkiewicz, prof. nadzw. PW
Wydział Elektroniki i Technik Informatycznych
Politechnika Warszawska
ens@ia.pw.edu.pl

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY
WYDZIAŁU INFORMATYKI
POLSKO-JAPOŃSKIEJ WYŻSZEJ SZKOŁY TECHNIK KOMPUTEROWYCH**

Tytuł rozprawy: Obiektowe i agentowe strategie przetwarzania wielkich zadań obliczeniowych w sieciach komputerowych

Autor rozprawy: mgr Marek Grochowski

1. Ogólna charakterystyka rozprawy. Jakie zagadnienie naukowe jest rozpatrzone w pracy? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Rozprawa doktorska mgr Marka Grochowskiego jest poświęcona rozważaniom dotyczącym rozwiązywania wielkich zadań obliczeniowych w sieciach komputerowych. Do rozwiązania tego typu zadań Autor proponuje wieloagentowy, rozproszony system informatyczny z algorytmem równoważenia obciążeń stosującym szeregowanie dyfuzyjne. Autor stawia tezę mówiącą o tym, że zaproponowany system jest wydajnym, efektywnym i wygodnym narzędziem do rozwiązywania zadań o znacznej złożoności obliczeń i stanowi alternatywę zarówno dla podejść stosujących biblioteki i pakiety programów do zrównoleglenia obliczeń jak i narzędzi do tworzenia obiektowych programów rozproszonych w środowiskach sieciowych. Tezę tę uzasadnia przedstawiając wyniki zastosowania systemu do rozwiązania dwóch przykładowych zadań obliczeniowych charakteryzujących się regularną i nieregularną równoległością. Zagadnienie będące przedmiotem rozprawy jest jasno przedstawione w części wstępnej pracy.

Zasadniczą część pracy podzielono na dziewięć rozdziałów. Opis rozważanych przykładów obliczeniowych oraz zastosowanych metod numerycznych do rozwiązania postawionych problemów zawarto w dwóch dodatkach.

Pracę rozpoczyna wprowadzenie w tematykę szeregowania zadań w systemach rozproszonych oraz systemów wieloagentowych. Podane są definicje i notacja stosowana w pracy. W następnych rozdziałach Autor przedstawia ewolucję koncepcji rozproszonych systemów obliczeniowych, rozpoczynając od omówienia popularnych bibliotek

komunikacyjnych i środowisk oprogramowania do realizacji programów rozproszonych, następnie narzędzi do tworzenia obiektowych programów rozproszonych w środowiskach sieciowych, kończąc na prezentacji koncepcji wieloagentowego systemu obliczeniowego. W rozdziale szóstym Autor koncentruje się na omówieniu architektury systemu wieloagentowego oraz proponowanych algorytmów dekompozycji, szeregowania, synchronizacji i migracji agentów. W rozdziale tym prezentowane są również modele formalne wykonanego systemu oraz dyfuzyjnego algorytmu szeregowania zadań. Dwa kolejne rozdziały koncentrują się na prezentacji rozwiązywanych przykładów obliczeniowych oraz omówieniu wyników badań numerycznych.

Podsumowując, rozprawa ma charakter teoretyczny, jak i doświadczalny oraz konstrukcyjny. W istocie dominuje charakter konstrukcyjny. Szczególnie istotne osiągnięcia Autora mają formę zaawansowanego wieloagentowego systemu oprogramowania. W tekście rozprawy było możliwe jedynie przedstawienie opisu architektury wykonanego systemu oraz kolejnych etapów jego konstrukcji, zaprezentowanie jego możliwości i wyników uzyskanych przy jego zastosowaniu.

Rozważania teoretyczne obejmują zagadnienia związane z systemami wieloagentowymi oraz technikami szeregowania zadań. Część doświadczalna zawiera liczne eksperymenty numeryczne pokazujące efekty rozproszenia obliczeń, tj. ich przyspieszenie poprzez dynamiczny rozdział zadań między komputery w sieci.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł świadczą o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Autor prezentuje dużą wiedzę w rozważanej problematyce, popartą bogatym zestawem cytowanej literatury. Wykaz cytowanej literatury liczy 134 pozycje. Wnioski sformułowane na podstawie przeglądu źródeł są przekonujące. Na wiedzy literaturowej w znacznym stopniu są oparte rozdziały 2, 4 i 5 rozprawy oraz dodatki. W rozdziale 4 omówione są: biblioteka MPI (Message Passing Interface) oraz środowisko PVM (Parallel Virtual Machine) i jego wersja udostępniająca migrację procesów MPVM. Brakuje w tym rozdziale wspomnienia o wygodnym rozwiązaniu – dyrektywach zrównoleglających OpenMP w klastrach. Omawiając środowiska wieloagentowe do obliczeń rozproszonych (rozdział 2) należałoby wspomnieć o systemach klastrowych realizujących paradygmat jednolitego obrazu systemu SSI (Single System Image) takich jak Mosix, OpenSSI czy Kerrighed.

3. Czy Autor rozwiązał postawione zagadnienia, czy użył właściwej to tego metody i czy przyjęte założenia są uzasadnione.

Autor rozwiązał postawione zagadnienia, użył właściwych metod teoretycznych i przedstawił przekonujące wnioski z badań symulacyjnych, które opracował w sposób syntetyczny. W zakresie konstrukcji proponowanego wieloagentowego systemu informatycznego przeprowadził rozważania teoretyczne i liczne testy potwierdzające efektywność przyjętych rozwiązań, w tym zastosowanego algorytmu do równoważenia obciążeń – metody dyfuzyjnej. Na podkreślenie zasługuje duży nakład pracy związany z tworzeniem złożonego systemu informatycznego i prowadzeniem eksperymentów symulacyjnych. Autor zaznacza, że prace te były wykonywane we współpracy z innymi osobami.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora?

Do oryginalnych rezultatów Autora zaliczam:

- Konstrukcję i wykonanie wieloagentowego systemu oprogramowania do obliczeń rozproszonych, w tym: opracowanie formalnego opisu wykonanego systemu, zdefiniowanie reguł współdziałania agentów obliczeniowych ze środowiskiem informatycznym, w którym się wykonują, propozycję architektury tzw. Smart Solid Agent-a, zaprojektowanie i realizację algorytmów sterowania ziarnem obliczeń oraz podziału, szeregowania, synchronizacji i migracji agentów.
- Analizę efektywności wykonanego systemu oprogramowania poprzez realizację licznych eksperymentów numerycznych w dość dużych klastrach obliczeniowych.
- Propozycję modyfikacji algorytmu dyfuzyjnego równoważenia obciążeń w systemach rozproszonych, jego eksperymentalną weryfikację i porównanie z innym rozwiązaniem proponowanym w literaturze.
- Zdefiniowanie nowych miar oceny efektywności wykorzystania mocy obliczeniowej komputerów przez systemy wieloagentowe.
- Zastosowanie systemu wieloagentowego do rozwiązania dwóch praktycznych zadań obliczeniowych: z regularną dekompozycją – metoda elementów skończonych oraz z nieregularną dekompozycją – hierarchiczny algorytm genetyczny.

5. Czy Autor wykazał umiejętność poprawnego przekonywującego przekazu przedstawienia uzyskanych przez siebie wyników?

Rozprawa liczy 128 stron, zawiera 27 rysunków i 18 tabel. Opracowanie zaopatrzone również w skorowidz oraz spis oznaczeń, które ułatwiają analizę jego treści. Praca jest zredagowana starannie zarówno pod względem językowym jak i graficznym. Błędy językowe są stosunkowo rzadkie. W pewnych miejscach występuje pewna niejednoznaczność notacji, np. na str. 47, T oznacza zadanie wykonywane przez agenta, a na str. 9, T oznacza opóźnienie wykonania zadania, podobnie różne znaczenie w rozdziałach dwa, cztery i sześć przyjmują symbole C i L . Autor w rozdziale 6 nie podaje objaśnień symboli pod formułami matematycznymi. Objaśnienia te znajdują się często kilka stron wcześniej. Utrudnia to analizę treści. Wskazane byłoby ich powtórzenie. W pewnych miejscach występują błędy w formułach, np. str. 47 mamy $A_i = (T_i, S_i)$, a na str. 49 $A_i = (S_{ij}, T_{ij})$, błąd we wzorze (6.10). Przy pierwszych pojawieniach się skrótów warto je rozwinąć. Na rysunku 4.1 opisy osi są nieczytelne. W rozdziale zawierającym spis literatury w niektórych pozycjach są podawane pełne imiona autorów, w innych inicjały.

Jak rozumiem, intencją Autora w rozdziale czwartym było pokazanie trudności w realizacji obliczeń z wykorzystaniem bibliotek komunikacyjnych, gdzie użytkownik musi zajmować się operacjami niskopoziomowymi, tj. formułowaniem komunikatów, synchronizacją, itd. Problemy te nie są podkreślone w tekście.

Omówienie niektórych wyników numerycznych jest zbyt skrótowe. Wskazane byłoby dokładniejsze wyjaśnienie w jaki sposób uzyskano tak duże przyspieszenia obliczeń dla przykładów prezentowanych w rozdziałach 4.3.1 oraz 5.5.3. Nie są również oczywiste wyniki zamieszczone w tabeli 8.8 (rozdział 8.5.2).

6. Jakie są słabe strony rozprawy i jej główne wady?

Rozprawa jako całość stwarza korzystne wrażenie. Wymienione poniżej wady mają drugorzędne znaczenie.

- Jak wspomniano wcześniej wskazane byłoby ustosunkowanie się do popularnych systemów klastrowych z grupy SSI (Single System Image).
- Omawiając biblioteki komunikacyjne do realizacji obliczeń rozproszonych Autor prezentuje wyniki numeryczne zastosowania pakietów PVM i MPVM. W obu przypadkach rozwiązywane są różne zadania. Wydaje się, że bardziej przydatne byłoby porównanie obu tych pakietów na przykładzie takiego samego zadania rozwiązywanego w tym samym środowisku sprzętowym. Ponadto większą wartość miałyby wykonanie testów dla interfejsu MPI, który stanowi obecnie standard przemysłowy, niż stosowanie raczej przestarzałego pakietu PVM (w 1995 roku przestano go rozwijać).
- Prace implementacyjne oraz testy numeryczne były wykonywane przez Autora przy współpracy innych osób. Z pracy nie wynika jasno jaki był udział Autora, w szczególności przy budowie wieloagentowego środowiska do przetwarzania rozproszonego Octopus.

7. Wniosek końcowy

Uważam, że rozprawa doktorska mgr Marka Grochowskiego spełnia w zupełności wymagania stawiane rozprawom doktorskim przez przepisy ustawy o tytułach i stopniach naukowych. Stawiam wniosek o przyjęcie rozprawy i dopuszczenie do obrony doktorskiej.



podpis