

PRACA DYPLOMOWA MAGISTERSKA

Tytuł pracy

Title of the Thesis

Jan Kowalski

Nr albumu: XXXXX

Kierunek: Informatyka

Forma studiów: stacjonarne

Poziom studiów: II

Promotor pracy:

dr inż. Tomasz Olas

Praca przyjęta dnia:

Podpis promotora:

Częstochowa, 2023

Spis treści

1	Wstęp	4
1.1	Cel i zakres pracy	4
1.2	Struktura pracy	4
2	MES i jej implementacja w systemach równoległych i rozproszonych	5
2.1	Przykładowy rozdział - Architektury systemów równoległych i rozproszonych .	5
2.2	Klastry komputerów PC i stosowane w nich sieci komunikacyjne	6
2.2.1	Architektura klastrów obliczeniowych	7
2.2.2	Sieci komunikacyjne stosowane w klastrach komputerów PC	9
	Bibliografia	11

1. Wstęp

To jest wstęp pracy magisterskiej

1.1. Cel i zakres pracy

Cel i zakres pracy

1.2. Struktura pracy

Praca składa się z X rozdziałów. Rozdział pierwszy pokazuje w jaki sposób pisać pracę magisterską wraz z najczęściej wykorzystywanymi elementami - tworzenie akapitów, wypunktowania, dodawanie tabel, rysunków i odniesień do nich, umieszczanie w tekście odwołań do literatury. Pozycje literatury znajdują się w pliku mgr.bib - do jego utworzenia i modyfikacji polecam program jabref.

2. Metoda elementów skończonych i jej implementacja w systemach równoległych i rozproszonych

2.1. Przykładowy rozdział - Architektury systemów równoległych i rozproszonych

W niniejszej pracy definiuje się pojęcie obliczeń równoległych jako proces równoległego rozwiązywania problemu z wykorzystaniem p procesorów, przy czym zadania uruchamiane w poszczególnych procesach są od siebie zależne i niezbędna jest wymiana danych pomiędzy nimi [1]. W szczególności przetwarzanie równoległe dotyczy obliczeń, w których procesy przyporządkowywane są procesorom w komputerze równoległym, natomiast w przypadku obliczeń rozproszonych procesy przypisywane są poszczególnym komputerom w systemie wielomaszynowym.

Równoległy system komputerowy (lub po prostu komputer równoległy) definiowany jest jako system zawierający szereg procesorów, zdolnych do wspólnego rozwiązywania pewnego zagadnienia obliczeniowego. Określenie to jest na tyle szerokie, że obejmuje superkomputery masywnie równoległe, składające się z dziesiątków, setek, a nawet tysięcy procesorów, wieloprocessorowe stacje robocze i serwery, systemy wbudowane (*embedded systems*) itp.

Rozproszone systemy komputerowe, zwane również architekturami wielomaszynowymi, składają się z wielu samodzielnych komputerów (procesor, pamięć, wejście-wyjście) połączonych siecią komunikacyjną.

Komputery równoległe można podzielić według dwóch głównych klasyfikacji. Pierwsza z nich pochodzi od M. Flynna [2]. Podstawą tej klasyfikacji jest liczba strumieni danych oraz liczba strumieni rozkazów jakie mogą zostać wyróżnione w pracy systemu komputerowego. Zostały wyróżnione cztery typy organizacji komputerów:

- SISD - systemy z pojedynczym strumieniem rozkazów i danych (*Single Instruction stream - Single Data stream*),
- SIMD - systemy z pojedynczym strumieniem rozkazów i wielokrotnym strumieniem danych (*Single Instruction stream - Multiple Data stream*),
- MISD - komputery z wielokrotnym strumieniem rozkazów i pojedynczym strumieniem danych (*Multiple Instruction stream - Single Data stream*),
- MIMD - maszyny z wielokrotnym strumieniem rozkazów i wielokrotnym strumieniem danych (*Multiple Instruction stream - Multiple Data stream*).

Druga klasyfikacja systemów równoległych opiera się na dostępie do pamięci. Wyróżnia się tutaj systemy z:

- pamięcią wspólną lub inaczej współdzieloną (*shared-memory parallel computers*), pozwalające na jednoczesny i swobodny (*random*) dostęp do wspólnej pamięci przez wszystkie procesory,
- pamięcią rozproszoną (*distributed memory*), w których każdy procesor posiada bezpośredni dostęp tylko do swojej lokalnej pamięci, oprócz tego może się odwoływać do modułów pamięci innych procesorów w sposób pośredni za pomocą odpowiedniej sieci połączeń (*interconnection network*).

2.2. Klastry komputerów PC i stosowane w nich sieci komunikacyjne

Idea klastra obliczeniowego stanowi rozszerzenie idei sieci stacji roboczych (*Network Of Workstations - NOW*). Pod względem sprzętowym oba składają się z pewnej liczby komputerów (stacji roboczych) połączonych siecią komunikacyjną. Różnica polega na logicznej reprezentacji systemu: w przypadku sieci stacji roboczych każda maszyna widziana jest jako niezależny system i fakt ten musi znaleźć odbicie w stosowanych tam rozwiązaniach programistycznych. W przypadku klastra natomiast pojedyncze maszyny stanowią część wirtualnego komputera udostępnionego dla użytkowników w sposób analogiczny do rzeczywistych superkomputerów.

Jedną z podstawowych metod klasyfikacji klastrów stanowi podział na [3, 4]:

1. klastry wysokiej wydajności lub obliczeniowe (*high performance clusters*),
2. klastry wysokiej dostępności (*high availability clusters*).

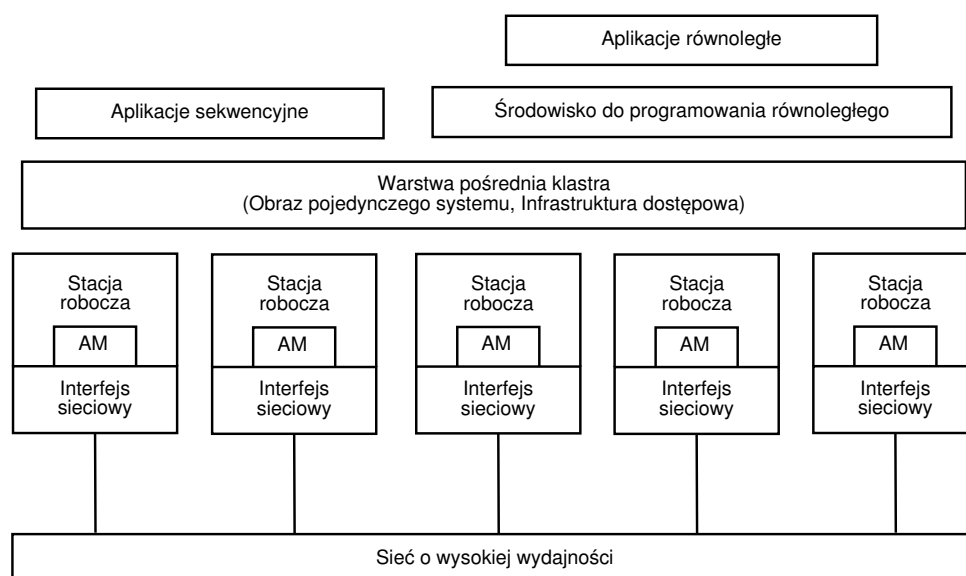
W przypadku klastrów obliczeniowych głównym celem twórców jest osiągnięcie jak najwyższej maksymalnej wydajności maszyny równoległej. Powoduje to stosowanie komponentów o najlepszych, pod względem szybkości, parametrach, wykorzystanie dedykowanych sieci komunikacyjnych oraz bazowanie na zabezpieczeniach programowych w zakresie stabilności pracy wirtualnego komputera.

Klastry wysokiej dostępności wykorzystywane są w zadaniach wymagających niezwykle wysokiej niezawodności i gwarancji nieprzerwanej pracy komputera. Często komponenty stosowane w nich należą do starszych generacji, mających za sobą długi okres testów i udoskonaleń oraz gwarantujących bezawaryjną współpracę z warstwą programową. Powszechnie spotykaną praktyką jest również wykonywanie pojedynczego zadania jednocześnie przez wiele węzłów klastra, co umożliwia łatwe wykrycie przekłamań w wynikach generowanych przez wadliwą jednostkę jak również natychmiastowe zastąpienie jej przez inną, należącą do tej samej grupy. Praktykowane jest również pozostawienie grupy maszyn nieprzydzielonych do żadnego zadania jako zabezpieczenia w przypadku poważniejszej awarii. Typowym przykładem zastosowania klastrów wysokiej dostępności są serwery internetowe oraz systemy zarządzające przedsiębiorstwami.

2.2.1. Architektura klastrów obliczeniowych

Oprogramowanie instalowane na klastrach obliczeniowych można podzielić na kilka grup pod względem spełnianej przez nie funkcji. Podział ten odpowiada warstwowemu schematowi budowy klastra. Kolejnymi warstwami są (rys. 2.1):

- podstawowa warstwa komunikacyjna
- środowiska programowania równoległego
- warstwa pośrednia (*middleware*)
- aplikacje użytkowe wykorzystujące poprzez wszystkie niższe warstwy moc obliczeniową klastra.



Rysunek 2.1. Architektura klastrów obliczeniowych

Niejaką obok tej hierarchii znajdują się środowiska programistyczne i narzędzia wykorzystywane w procesie tworzenia i testowania aplikacji użytkowych oraz system operacyjny będący podstawą działania wszystkich wymienionych składowych. W chwili obecnej większość klastrów pracuje pod kontrolą systemów UNIX'owych. Często wykorzystywany jest system operacyjny Linux, będący darmową implementacją systemu UNIX, ze względu na jego niezawodność i łatwość dostosowania do różnorodnego sprzętu i potrzeb użytkowników.

Do podstawowej warstwy komunikacyjnej zaliczamy sterowniki urządzeń sieciowych, zarówno podstawowych, wykorzystywanych głównie do zarządzania klastrem, jak i kart wysokiej wydajności stosowanych w trakcie uruchamiania zadań obliczeniowych.

Zadaniem zaawansowanej warstwy komunikacyjnej jest utworzenie środowiska pracy dla programu równoległego oraz zapewnienie łączności pomiędzy poszczególnymi jego procesami. Ponadto w skład tej warstwy będą wchodziły systemy emulujące pamięć współdzieloną na architekturze z pamięcią rozproszoną jakim jest klaster.

Określenie middleware dotyczy aplikacji i narzędzi służących ułatwieniu przydziału zasobów maszyn równoległych użytkownikom. Pozwalają one uzyskać efekt reprezentacji klastra jako pojedynczej maszyny równoległej z punktu widzenia użytkownika. W chwili obecnej istnieje wiele systemów oferujących zróżnicowane podejście do zagadnienia przydziału zasobów klastra. Niektóre z nich stanowią zespół programów umożliwiających uruchamianie aplikacji użytkowych na klastrze zgodnie z zadaną polityką przydziału zasobów bez konieczności ich

modyfikacji. Systemy takie noszą nazwę systemów kolejkowania. Do tej grupy zaliczają się między innymi systemy LSF [5], PBS [6] i Condor [7].

Istnieją również aplikacje stanowiące połączenie systemu przydziału zasobów oraz oprogramowania podstawowej i zaawansowanej warstwy komunikacyjnej. Niejednokrotnie są one wyposażone również w pewne cechy narzędzi do oceny wydajności programu. Ich poważną wadą jest fakt wysokiej złożoności i głębokiej ingerencji w system operacyjny. Przykładem takiego systemu jest SCore [8].

Ostatnim sposobem dystrybucji zasobów klastra jest rozwiązanie całkowicie lub niemal całkowicie przeźroczyste dla użytkownika. Wymaga ono modyfikacji systemu operacyjnego w taki sposób, że procesy tworzone na jednym węźle przenoszone są niejawnie na inny węzeł dysponujący w danej chwili największą ilością wolnych zasobów. Zaawansowaną implementacją tej koncepcji jest system Mosix [9] stanowiący modyfikację jądra systemu Linux.

2.2.2. Sieci komunikacyjne stosowane w klastrach komputerów PC

Sieć komunikacyjna jest jednym z istotniejszych czynników wpływających na ogólną wydajność klastra. W chwili obecnej istnieje szereg rozwiązań, które mogą być z powodzeniem zastosowane przy budowie klastrów. Koszt tych rozwiązań, w przeliczeniu na pojedynczy węzeł, waha się od prawie zera, w przypadku rozwiązań stosowanych powszechnie w sieciach lokalnych, aż do kwoty zbliżonej do ceny pojedynczego węzła klastra, w przypadku wysokowydajnych sieci komunikacyjnych przeznaczonych specjalnie dla klastrów obliczeniowych (Tabela 2.1). Dlatego też wybór sieci komunikacyjnej, jaka ma zostać wykorzystana do budowy klastra jest zadaniem trudnym, wymagającym uwzględnienia wielu czynników, nie tylko kosztu całego rozwiązania, należy mieć na uwadze również ogólną wydajność klastra. Wśród najczęściej wykorzystywanych sieci komunikacyjnych w klastrach można wymienić sieci Ethernet, Fast Ethernet, Gigabit Ethernet, Myrinet, SCI (Scalable Coherent Interface), Giganet (cLAN), QsNet, ServerNet, ATM (Asynchronous Transmission Mode). Obiecująco wygląda rozwój technologii InfiniBand, która może stać się standardem w zakresie rozwiązań klastrowych.

Tabela 2.1. Porównanie wybranych sieci komunikacyjnych stosowanych w klastrach komputerów PC (Opracowane na podstawie [10])

	Myrinet	SCI	Gigabit Ethernet	Infiniband
Przepustowość (MBytes/s)	140 - 33 MHz 215 - 66 MHz	80	30 - 50	740 - 840
Opóźnienie - MPI (μs)	14 - 33 MHz 5 - 66 MHz	6	50 - 200	6 - 7
Minimalna cena za port	\$1000	\$1000	\$150	\$1000
Wsparcie dla systemu Linux	tak	tak	tak	tak
Maksymalna liczba węzłów	1024	1000's	1000's	64K
MPI	LAM, MPICH-GM	3rd Party	MPICH	MVAPICH, MIBAPICH, MVAPICH2

Bibliografia

- [1] R. Wyrzykowski. Organizacja procesów obliczeniowych w równoległych i rozproszonych systemach komputerowych. <http://www.k2.pcz.czyst.pl/~roman/book1/spis.html>.
- [2] S. Kozielski and Z. Szczepiński. *Komputery równoległe: architektura, elementy programowania*. Wydawnictwa Naukowo-Techniczne, Warszawa, 1993.
- [3] R. Buyya ed. *High Performance Cluster Computing: Programming and Applications*. Prentice Hall Inc., 1999.
- [4] J. Kitowski. *Współczesne systemy komputerowe*. CCNS, Kraków, 2000.
- [5] <http://www.platform.com>. Platform computing - accelerating intelligence - grid computing. Web Page, 2004.
- [6] <http://www.openpbs.org>. OpenPBS. Web Page, 2003.
- [7] Douglas Thain, Todd Tannenbaum, , and Miron Livny. Distributed Computing in Practice: The Condor Experience. *Concurrency and Computation: Practice and Experience*, 2004.
- [8] <http://cluster.mcs.st-and.ac.uk/score/en>. SCore Cluster System Software Documents. Web Page, 2003.
- [9] <http://www.mosix.org>. MOSIX. Web Page, 2004.
- [10] M. Baker. Cluster Computing White Paper, 2000. [cite-seer.ist.psu.edu/baker00cluster.html](http://seer.ist.psu.edu/baker00cluster.html).