

PRACA DYPLOMOWA MAGISTERSKA

Analiza wydajności i efektywności wybranych algorytmów sortowania równoległego na wielordzeniowych procesorach

*A performance and efficiency analysis of selected
parallel sorting algorithms on multicore processors*

Krzysztof Pielot

Nr albumu: 133537

Kierunek: Informatyka

Forma studiów: stacjonarne

Poziom studiów: II

Promotor pracy:

dr inż. Bartosz Kowalczyk

Praca przyjęta dnia:

Podpis promotora:

Częstochowa, 2025

Niniejszym chciałbym serdecznie podziękować

...

*za bezcenne wsparcie udzielone mi
w trakcie trwania studiów.*

Spis treści

Wstęp	3
Cel pracy	3
Zakres pracy	3
1 Wprowadzenie teoretyczne	5
1.1 Wprowadzenie do sortowania równoległego	5
1.2 Charakterystyka procesorów wielordzeniowych	5
1.3 Wydajność i efektywność	5
1.4 Skalowalność algorytmów równoległych	5
2 Opis wybranych algorytmów sortowania równoległego	7
2.1 Parallel Quicksort	7
2.2 Parallel Merge Sort	7
2.3 Parallel Bucket Sort	7
2.4 Odd-Even Transposition Sort	7
3 Środowisko badawcze i implementacja	8
3.1 Opis środowiska testowego	8
3.1.1 Konfiguracje sprzętowe	8
3.1.2 Konfiguracje systemowe	8
3.2 Wybór języka i bibliotek	8
3.3 Sposób implementacji algorytmów	8
3.4 Charakterystyka danych testowych	9
4 Metodyka badań	10
4.1 Scenariusze testowe	10
4.2 Wariant sekwencyjny jako punkt odniesienia	10
4.3 Parametry pomiarów	10
4.4 Narzędzia do pomiaru czasu, CPU i pamięci	10
4.5 Przebieg eksperymentów	10
5 Analiza wyników badań	11
5.1 Porównanie czasu sortowania	11

5.2	Wpływ liczby rdzeni na wydajność	11
5.3	Analiza zużycia pamięci RAM i obciążenia CPU	11
5.4	Skalowalność algorytmów	11
5.5	Wpływ typu danych	11
5.6	Porównanie efektywności w zależności od charakterystyki danych wejściowych	12
Podsumowanie		13
Bibliografia		15
Spis rysunków		16
Spis tabel		17
Listing		18
Streszczenie		19
Summary		20
Słowa kluczowe		21
Dodatek. Zawartość dołączonej płyty		22

Wstęp

W ostatnim czasie obserwuje się intensywny rozwój [Proszę uzupełnić].

Niestety [Proszę wpisać wadę lub niedogodność istniejących rozwiązań].

W związku z tym, ciekawym wydaje się zaprojektowanie i zrealizowanie [Proszę uzupełnić].

Cel pracy

Celem niniejszej pracy jest zaprojektowanie i zrealizowanie [Proszę uzupełnić].

Zakres pracy

Zakres pracy obejmuje:

- ▶ zebranie wiadomości z zakresu [Proszę uzupełnić],
- ▶ porównanie [Proszę uzupełnić],
- ▶ opracowanie założeń projektowych dotyczących [Proszę uzupełnić],
- ▶ implementację [Proszę uzupełnić],
- ▶ przetestowanie [Proszę uzupełnić].

W Rozdziale 1 pracy opisano/ podsumowano/ opisano/ rozważano/ skoncentrowano się na [Proszę uzupełnić]. Natomiast w Rozdziale [Proszę uzupełnić] opisano/ podsumowano/ opisano/ rozważano/ skoncentrowano się na [Proszę uzupełnić].
itd.

Rozdział 1

Wprowadzenie teoretyczne

1.1 Wprowadzenie do sortowania równoległego

Wprowadzenie w temat sortowania równoległego, dlaczego, kiedy oraz jakie korzyści. Jakie są ogólne strategie sortowania równoległego, czyli podział danych ich przetwarzanie oraz synchronizacja. Wyzwania z takiego podejścia jak komunikacja między wątkami, synchronizacja, narzut systemowy

1.2 Charakterystyka procesorów wielordzeniowych

Wyjaśnienie na jakiej architekturze sprzętowej będą testowane dane algorytmy, czyli czym jest procesor wielordzeniowy, jaka są różnice między rdzeniem fizycznym a wątkiem logicznym, jak przebiega komunikacja między rdzeniami, współdzielenie pamięci RAM, znaczenie równoległości sprzętowej i jej ograniczenia oraz wpływ liczby rdzeni na możliwości przyspieszania obliczeń

1.3 Wydajność i efektywność

Przedstawienie metryk używanych w analizie porównawczej algorytmów. Definicje wydajności oraz efektywności, czas przetwarzania, pomiary zużycia RAM, obciążenie CPU, Speedup oraz Efficiency

1.4 Skalowalność algorytmów równoległych

Omówienie, jak zmienia się wydajność algorytmu w zależności od liczby rdzeni oraz wielkości danych. Czym w ogóle jest skalowalności w informatyce, jakie są typy skalowalności czyli pozioma oraz pionowa, czynniki ograniczające skalowalność takie

jak synchronizacja, nierównomierny podział danych, dostęp do pamięci, zależności danych

Rozdział 2

Opis wybranych algorytmów sortowania równoległego

Dział poświęcony wybranym algorytmom zawierający opis, schemat działania, sposób dzielenia i łączenia, złożoność, potencjalne problemy oraz kiedy algorytm jest efektywny, a kiedy nie

2.1 Parallel Quicksort

2.2 Parallel Merge Sort

2.3 Parallel Bucket Sort

2.4 Odd-Even Transposition Sort

Rozdział 3

Środowisko badawcze i implementacja

3.1 Opis środowiska testowego

3.1.1 Konfiguracje sprzętowe

Szczegółowe dane dotyczące testowych maszyn takie jak model procesora, liczba rdzeni fizycznych i logicznych, pamięć RAM itp.

3.1.2 Konfiguracje systemowe

Szczegółowe dane dotyczące testowych maszyn takie jak system operacyjny, wersje sterowników i bibliotek systemowych itp.

3.2 Wybór języka i bibliotek

Informacje o wybranym języku programowania oraz bibliotek, uzasadnienie wyboru oraz opis

3.3 Sposób implementacji algorytmów

Opisanie w jaki sposób zostały zaimplementowane algorytmy sortujące. Ogólna architektura, w zależności od algorytmu jak następuje dzielenie, synchronizacja danych, wymiany danych, informacje jak został zaimplementowany waraint sekwencyjny

3.4 Charakterystyka danych testowych

Sposób przechowywania, typy danych(int, float), charakterystyka danych(losowe, częściowo posortowane, z dużą ilością duplikatów), rozmiar testowanych zbiorów danych, jak zostały wygenerowane dane do testów

Rozdział 4

Metodyka badań

4.1 Scenariusze testowe

Opis różnych warunków i przypadków testowych takich jak na jakich rozmiar danych, wariant algorytmu, liczba rdzeni, różne typy zbiorów(losowe itp.)

4.2 Wariant sekwencyjny jako punkt odniesienia

Wyjaśnienie, że dla porównania wyników algorytmów równoległych testowana jest także wersja sekwencyjna każdego algorytmu.

4.3 Parametry pomiarów

Parametry pomiarów co i jak będzie mierzone

4.4 Narzędzia do pomiaru czasu, CPU i pamięci

Opis wykorzystanych narzędzi i bibliotek użytych do pomiarów

4.5 Przebieg eksperymentów

Opis jak zostały przeprowadzone testy. Przygotowanie środowisk testowych. Sposób zapisywania i archiwizacji wyników. Procedura uruchomienia algorytmu itp.

Rozdział 5

Analiza wyników badań

5.1 Porównanie czasu sortowania

Prezentacja wyników pomiarów czasu wykonania dla każdego z algorytmów (dla różnych zbiorów danych, rozmiarów i typów danych). Omówienie i analiza

5.2 Wpływ liczby rdzeni na wydajność

Prezentacja i omówienie wyników pod względem użytych liczby rdzeni

5.3 Analiza zużycia pamięci RAM i obciążenia CPU

Prezentacja i omówienie zużycia pamięci i obciążenia w tym zestawienie średniego i maksymalnego zużycia

5.4 Skalowalność algorytmów

Ocena, jak dobrze algorytmy skalują się wraz ze wzrostem liczby rdzeni i rozmiarem danych

5.5 Wpływ typu danych

Porównanie wyników dla tych samych algorytmów, ale różnych typów danych wejściowych

5.6 Porównanie efektywności w zależności od charakterystyki danych wejściowych

Omówienie wyników dla danych losowych itp. Wskazanie, który algorytm radził sobie lepiej przy danym rodzaju danych

Podsumowanie

Bibliografia

- [1] Zasady pisania prac dyplomowych, <https://wiisi.pcz.pl/student-wiisi/vademecum-studenta/praca-dyplomowa>, stan na dzień: 26.10.2024
- [2] Strona internetowa Wydziału Informatyki i Sztucznej Inteligencji, <https://wiisi.pcz.pl>, stan na dzień: 26.10.2024

Spis rysunków

Spis tabel

Spis listingów

Streszczenie

Streszczenie

Summary

Tłumaczenie

Słowa kluczowe

informatyka; sortowanie równoległe; algorytmy sortowania; wielordzeniowe procesory; Python; wydajność algorytmów; skalowalność; pomiar czasu wykonania; zużycie pamięci RAM; obciążenie CPU; porównanie algorytmów; dane losowe; dane częściowo posortowane; duplikaty danych;

Dodatek. Zawartość dołączonej płyty

Do niniejszej pracy dołączono płytę z następującą zawartością:

- ▶ Dokument pracy w formatach tex i pdf.
- ▶ Kod źródłowy zaprojektowanego i zrealizowanego w ramach pracy systemu.

Częstochowa, dd.mm.20rr

Imię i nazwisko: Krzysztof Pielot

Nr albumu: 133537

Kierunek: Informatyka

Wydział: Informatyki i Sztucznej Inteligencji

Oświadczenie autora pracy dyplomowej*

Oświadczam pod rygorem odpowiedzialności karnej, że złożona przeze mnie praca dyplomowa pt. „*Analiza wydajności i efektywności wybranych algorytmów sortowania równoległego na wielordzeniowych procesorach*” jest moim samodzielnym opracowaniem i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami.

Jednocześnie oświadczam, że moja praca (w całości ani we fragmentach) nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w Politechnice Częstochowskiej.

Wyrażam/~~nie wyrażam~~** zgodę/zgody na nieodpłatne wykorzystanie przez Politechnikę Częstochowską całości lub fragmentów wyżej wymienionej pracy w publikacjach Politechniki Częstochowskiej.

.....
podpis studenta

*W przypadku zbiorowej pracy dyplomowej, dołącza się oświadczenia każdego ze współautorów pracy dyplomowej.

*Niepotrzebne skreślić.