Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”

Факультет інформатики і обчислювальної техніки Кафедра інформаційних систем та технологій

Лабораторна робота №2

З дисципліни «Технології паралельного програмування в умовах великих даних»

Тема: «Паралельні обчислення в моделі з розподіленою пам’яттю»

Виконали: Перевірив:

Студент групи ІА-11мн Жереб К. А.

Новиков Д.М.

Студент групи ІС-11мн

Іванов А.І.

Київ 2022

Тема: Паралельні обчислення в моделі з розподіленою пам’яттю.

Завдання: Необхідно реалізувати вирішення обраної задачі у вигляді послідовної програми та паралельної програми з розподіленою пам’яттю. Необхідно порівняти швидкодію послідовної та паралельної реалізації, а також за можливості продемонструвати залежність часу роботи від розміру задачі, кількості процесів та інших параметрів (залежно від задачі). Можна використовувати різні засоби паралелізму – зокрема, використати кілька процесів, або використати модель акторів в рамках одного процесу. Можна також реалізувати одразу декілька версій (як мультипроцесну, так і з моделлю акторів) та порівняти їх, за це можна отримати додаткові бали. Результатом виконання даної лабораторної роботи є працююча програма, а також звіт про використані технології та можливості, з результатами вимірів.

Репозиторій: <https://github.com/JokerFunny/parallel_programming>

Виконання роботи:

В якості корисного навантаження було обрано задачу: є великий файл, який зчитано рядково. В кожній строчці обчислити математичну формулу над кожним символом.

Реалізовано послідовну та паралельну версії. При написанні паралельної реалізації алгоритму з розподіленою пам’ятю використовувалась open-source бібліотека для роботи з паралельними та розподіленими застосунками [Orleans](https://github.com/dotnet/orleans) - це кросплатформна платформа для створення надійних, масштабованих розподілених програм. Створена Microsoft Research і надає можливість працювати з моделю віртуального актора як новий підхід до побудови нового покоління розподілених систем для епохи хмар. Основним внеском Orleans є її модель програмування, яка приборкує складність, притаманну високопаралельним розподіленим системам, не обмежуючи можливості та не накладаючи обтяжливих обмежень на розробника.

Результат роботи:

Sync version elapsed - 6028 ms.

Number of grains - 8, elapsed time - 1049 ms.

Sync version elapsed - 6018 ms.

Number of grains - 16, elapsed time - 777 ms.

Sync version elapsed - 6014 ms.

Number of grains - 32, elapsed time - 744 ms.

Sync version elapsed - 6047 ms.

Number of grains - 64, elapsed time - 727 ms.

Sync version elapsed - 6144 ms.

Number of grains - 128, elapsed time - 770 ms.

Sync version elapsed - 6015 ms.

Number of grains - 256, elapsed time - 684 ms.

Sync version elapsed - 6034 ms.

Number of grains - 512, elapsed time - 679 ms.

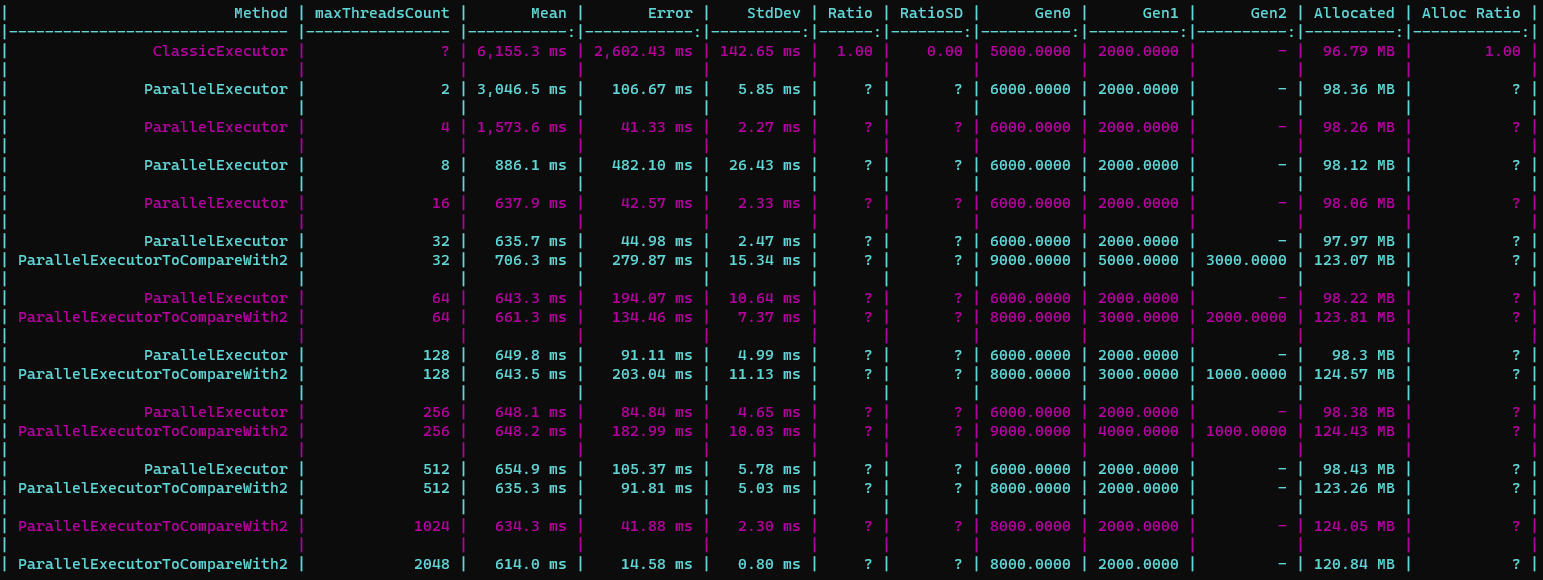
Sync version elapsed - 5994 ms.

Number of grains - 1024, elapsed time - 680 ms.

Sync version elapsed - 6049 ms.

Number of grains - 2048, elapsed time - 711 ms.

Паралельний варіант має кращий перформанс у випадках, коли ми створюємо велику кількість акторів (найкращий результат – при 512 акторах). Якщо дробити задачу занадто сильно, весь профіт від акторів втрачається, тому що на їх підтримку та отримання результату втрачається більше часу, ніж при послідовній версії.

В порівнянні з 1ю лабораторною (також додано ParallelExecutorToCompareWith2 бенчмарк який працює на тасках) маємо схожі результати (хоча й трохи гірше через інфраструктуру Orleans на створення і підтримку акторів):  


Висновки:

В результаті виконання лабораторної роботи було реалізовано алгоритм для вирішення поставленої задачі за допомогою послідовної та паралельної програми з розподіленою пам’яттю використовуючи модель акторів. Перевірено ефективність кожного підходу та проведено порівняння результатів між акторами та синхронним підходом та у порівнянні з ЛР1.