

Assignment 2

Задача 1. Введение в гетерогенную параллелизацию

Теоретическое задание (25 баллов)

Объясните, что такое гетерогенная параллелизация.

В ответе раскройте следующие аспекты:

- различия между параллельными вычислениями на CPU и GPU;
- преимущества гетерогенной параллелизации;
- примеры реальных приложений, в которых используется гетерогенная параллелизация.

Задача 2. Работа с массивами и OpenMP

Практическое задание (25 баллов)

Реализуйте программу на C++, которая:

1. Создаёт массив из 10 000 случайных чисел.
2. Находит минимальное и максимальное значения массива:
 - в последовательной реализации;
 - с использованием OpenMP для параллельной обработки.
3. Сравнивает время выполнения обеих реализаций и формулирует выводы.

Задача 3. Параллельная сортировка с OpenMP

Практическое задание (25 баллов)

Реализуйте алгоритм **сортировки выбором** с использованием OpenMP:

- напишите последовательную реализацию алгоритма;
- добавьте параллелизм с помощью директив OpenMP;
- проверьте производительность для массивов размером 1 000 и 10 000 элементов.

Задача 4. Сортировка на GPU с использованием CUDA

Практическое задание (25 баллов)

Реализуйте параллельную **сортировку слиянием** на GPU с использованием CUDA:

- разделите массив на подмассивы, каждый из которых обрабатывается отдельным блоком;
- выполните параллельное слияние отсортированных подмассивов;
- замерьте производительность для массивов размером 10 000 и 100 000 элементов.

Контрольные вопросы к Assignment 2

(OpenMP, CUDA и гетерогенные вычисления)

1. Что понимается под гетерогенной параллелизацией?
2. В чём принципиальные различия архитектур CPU и GPU?
3. Какие типы задач лучше подходят для выполнения на GPU, а какие — на CPU?
4. Почему не все алгоритмы эффективно распараллеливаются с использованием OpenMP?
5. В чём заключается основная идея алгоритма сортировки слиянием?
6. Какие сложности возникают при реализации сортировки слиянием на GPU?
7. Как выбор размера блока и сетки влияет на производительность вычислений на GPU?
8. Почему гетерогенный подход может быть эффективнее использования только CPU или только GPU?