

Прозрачное использование массово-параллельных архитектур для локальных оптимизаций приложений на .NET

Михайлов Дмитрий
СПбГУ, Математико-Механический факультет, группа 545

Научный руководитель: магистр информационных технологий С.В. Григорьев
Рецензент: Н.М. Тимофеев

Предметная область

- Прикладные задачи
 - Важно ускорить исполнение медленного участка на порядок
 - Нет необходимости в low latency
 - Сложно целиком написать на C или Fortran
- Массово-параллельные архитектуры
 - General-purpose graphics processing unit

Постановка задачи

- Изучить особенности GPGPU и средства интеграции с .NET
- Доработать библиотеку Brahma.FSharp
 - Массивы
 - Локальная память
 - Дефекты
- Использовать библиотеку для решения задачи поиска подстрок в строке
 - Data carving

Требования и ограничения

- Online-задача
 - Входная строка и шаблоны не известны заранее
- Длинная входная строка
 - До сотен ГБ
- Длина шаблона не больше 32 байт
- Количество шаблонов не больше 512
- На хосте доступно не больше 256 МБ памяти

Обзор существующих решений

- Использование нативного кода из .NET
 - CUDA.NET
 - Cloo
 - OpenCLTemplate
- Библиотека типов и операций над ними
 - MS Accelerator
 - FCore
- Трансляция кода на языке приложения на лету
 - Alea.cuBase

Архитектура

- F# Quotations
 - Подмножество языка
- Abstract syntax tree
- Трансляция AST в OpenCL
- Очередь команд

Особенности реализации

- Доработка библиотеки
 - Узлы AST для инициализации массивов
 - Специальные функции для синхронизации и выделения локальной памяти
- Задача поиска подстрок
 - Несколько совпадений в одной позиции
 - Проблема дробления входной строки
 - Параллельные чтение и обработка файла
 - Вычисление количества рабочих групп и размера буфера

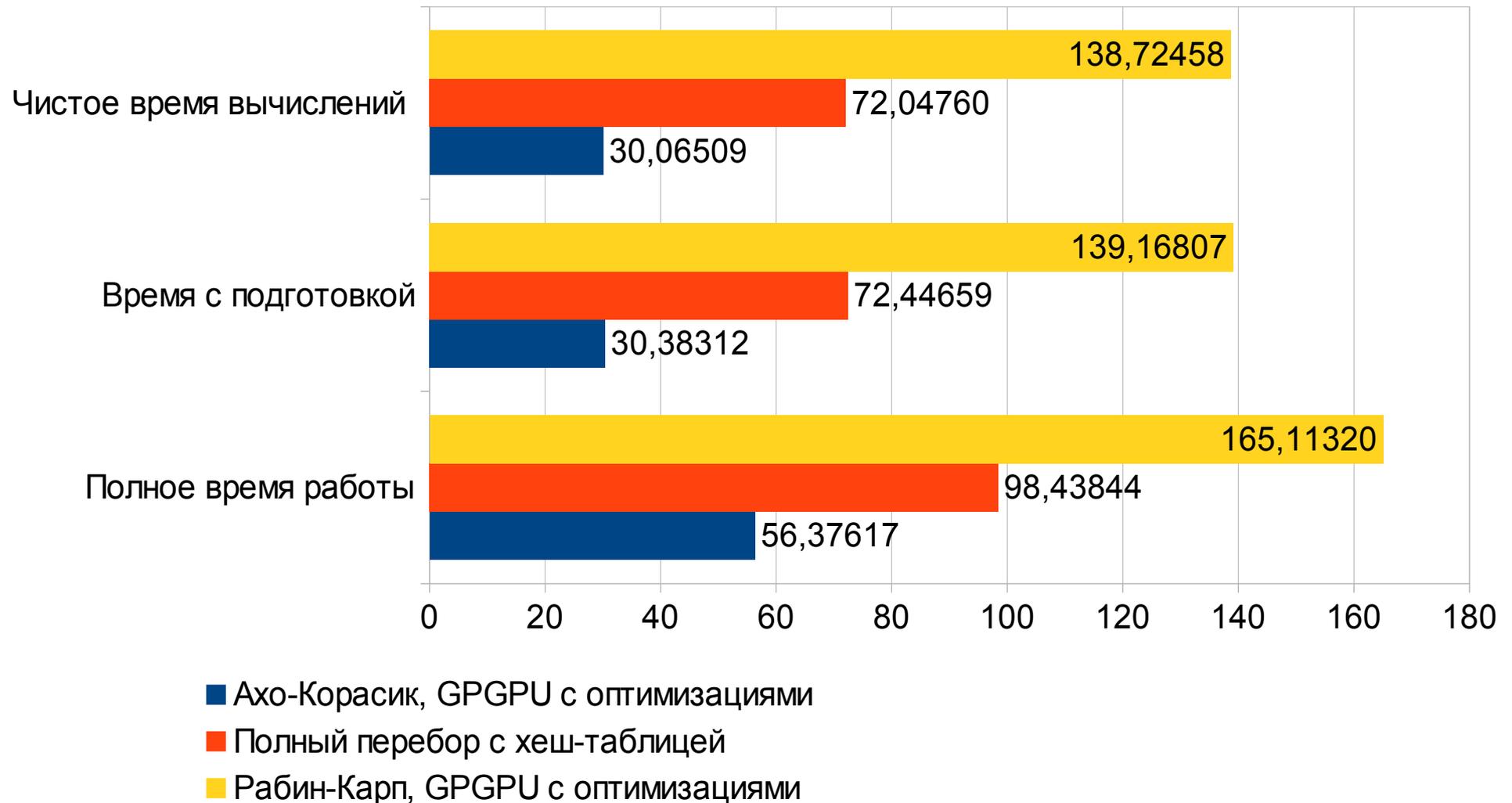
Алгоритмы

- Полный перебор
- Алгоритм Рабина-Карпа
 - Аддитивная хеш-функция
 - Быстрый пересчёт хешей подстрок при сдвиге
- Полный перебор с хеш-таблицей
- Алгоритм Ахо-Корасик
 - Инициализация лениво вычисляемой части автомата

Апробация

- Случайная входная строка размером 8 ГБ
- 512 случайных шаблонов длиной от 2 до 32
- NVIDIA GeForce GTX 670
- Intel Core i7-3770K

Сравнение алгоритмов



Результаты

- Реализована работа с массивами
- Реализована работа с локальной памятью
- Реализованы необходимые в нетривиальных задачах возможности
- Исправлены дефекты транслятора
- Реализован набор алгоритмов для решения выбранной задачи
 - Измерена производительность
 - Исследовано влияние новых возможностей и специфичных параметров