



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Московский государственный технический университет имени
Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»
КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

Отчет по лабораторной работе №4 по дисциплине «Анализ Алгоритмов»

Тема Параллельные вычисления на основе нативных потоков

Студент Козырных А.Д.

Группа ИУ7-52Б

Преподаватель Волкова Л. Л., Строганов Д.В.

2024 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Входные данные	3
2 Выходные данные	3
3 Тестирование	4
4 Исследование	4
4.1 Технические характеристики	4
4.2 Исследование характеристик	5
ЗАКЛЮЧЕНИЕ	9
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	10

ВВЕДЕНИЕ

Многопоточность — способность центрального процессора или одного ядра в многоядерном процессоре одновременно выполнять несколько процессов или потоков, соответствующим образом поддерживаемых операционной системой.

Потоки — это задача виртуального процессора, точно так же, как виртуальный процессор является задачей ЦПУ. Потоки иногда называют облегченными процессами, так как они похожи на процессы, но предъявляют меньше требований к операционной системе [1].

Цель лабораторной работы — Получить навык организации параллельных вычислений на основе нативных потоков.

Для достижения поставленной цели необходимо выполнить следующие задачи:

- разработка алгоритма обработки данных;
- создание ПО, реализующего разработанный алгоритм;
- исследование характеристик созданного ПО.

1 Входные данные

Входными данными программы является URL ссылка пагинации.

2 Выходные данные

Выходные данные — директория с файлами, которые содержат скачанные данные со страниц параграфа.

3 Тестирование

В таблице 1 представлены функциональные тесты для разработанного программного обеспечения. Все тесты пройдены успешно.

Таблица 1 – Время работы алгоритмов (в секундах)

№ Теста	Входные данные	Полученные данные	Ожидаемые выходные данные
1	https://vkusnye-recepty.ru/page/	Директория с текстами рецептов	Директория с текстами рецептов
2	https://nevkusnye-recepty.ru/page/ — несуществующая ссылка	Error. Invalid URL: https://nevkusnye-recepty.ru/page/1	Error. Invalid URL: https://nevkusnye-recepty.ru/page/1
3		Error. Invalid URL: 1	Error. Invalid URL: 1

4 Исследование

В ходе исследования требуется получить характеристики созданного ПО в зависимости от количества потоков или от количества обрабатываемых страниц сайта.

4.1 Технические характеристики

Технические характеристики устройства, на котором выполнялось тестирование:

- Операционная система: Linux [2];
- Память: 16 GB;
- Процессор: AMD Ryzen 7 5800H [3, 4].

4.2 Исследование характеристик

В таблице 2 приведено время выполнения программного обеспечения в миллисекундах (далее — мс). На рисунке 1 показана зависимость времени работы от количества потоков без изменений количества обрабатываемых страниц сайта.

Таблица 2 – Время работы от количества потоков (в миллисекундах)

Количество потоков	Время работы
1	4494.6
2	2449.4
4	1862.2
8	1407.6
16	930.8
32	385.4
64	185.4

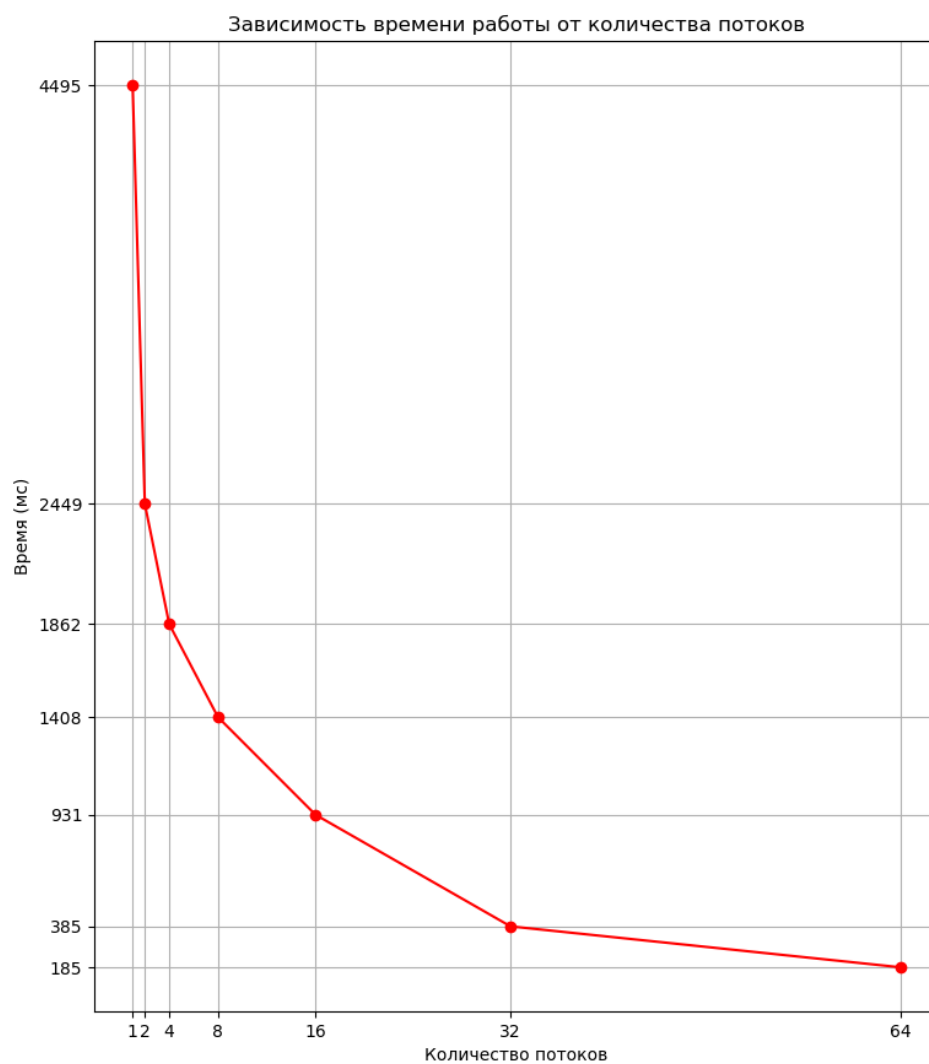


Рисунок 1 – Исследование характеристик созданного ПО от количества потоков

В таблице 3 приведено время выполнения в зависимости от количества обрабатываемых страниц. В отличие от предыдущего исследования, в таблице 3 отсутствует зависимость от количества потоков.

Таблица 3 – Время работы в зависимости от количества обрабатываемых страниц (в миллисекундах)

Количество страниц	Время для реализации	
	16-поточного	1-поточного
1	215.0	3437.4
2	716.4	10615.8
3	1372.6	21063.4
4	2222.2	34714.6
5	2821.6	46838.5

На рисунке 2 показана зависимость времени работы от количества обрабатываемых страниц.

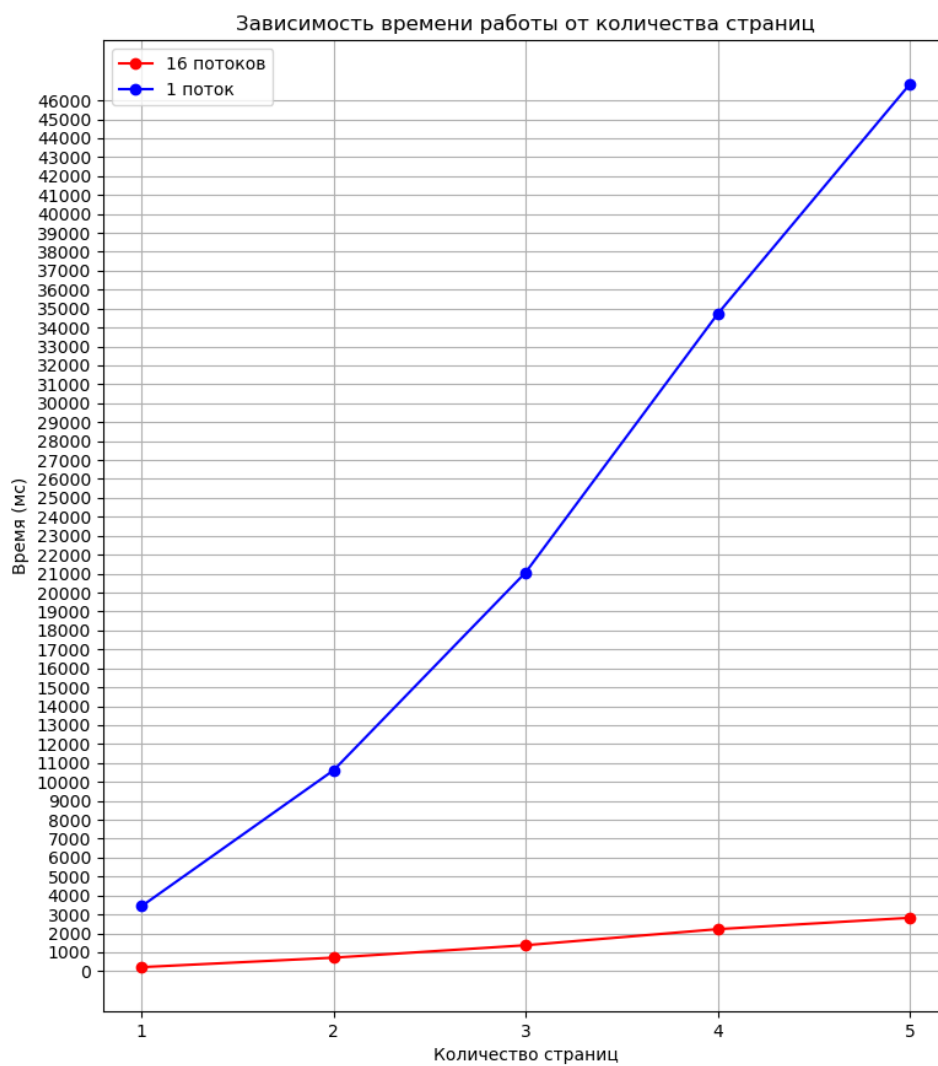


Рисунок 2 – Исследование характеристик созданного ПО от количества обрабатываемых страниц

ВЫВОД

По результатам проведенного исследования можно сказать, что время работы однопоточной реализации работа программы медленнее, чем в многопоточной реализации программы. Так, при фиксированном количестве потоков и при изменяющемся количестве обрабатываемых страниц время выполнения однопоточной реализации программы уступает в 16-17 раз по времени выполнения 16-поточной программы.

При росте количества потоков и при фиксированном количестве обрабатываемых страниц время выполнения уменьшается в 1.5-2.5 раза при увеличением количества потоков в два раза. Самая большая разница во времени выполнения между двух поточной программой и однопоточной программой, время выполнения которых различается в 1.8 раз.

ЗАКЛЮЧЕНИЕ

Цель работы достигнута. Решены все поставленные задачи:

- разработка алгоритма обработки данных;
- создание ПО, реализующего разработанный алгоритм;
- исследование характеристик созданного ПО.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- [1] Потоки. Документация IBM. [Электронный ресурс] URL: <https://www.ibm.com/docs/ru/informix-servers/12.10?topic=processors-threads> Дата обращения: 19.09.2024.
- [2] Операционная система Linux. Archlinux дистрибутив. [Электронный ресурс] URL: <https://archlinux.org/> Дата обращения: 19.09.2024.
- [3] AMD Ryzen7 5800H. Бенчмарк. [Электронный ресурс] URL: <https://technical.city/ru/cpu/Ryzen-7-5800H> Дата обращения: 19.09.2024.
- [4] AMD Ryzen7 5800H. Технические характеристики. [Электронный ресурс] URL: <https://www.notebookcheck-ru.com/AMD-Ryzen-7-5800H.519526.0.html> Дата обращения: 19.09.2024.