



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

Группа ИУ7-51Б

Тема работы **Параллельные вычисления на основе нативных потоков**

Студент

Баранов Николай Алексеевич

Преподаватель

Волкова Лилия Леонидовна

2024 г.

# СОДЕРЖАНИЕ

<b>ВВЕДЕНИЕ</b>	<b>3</b>
<b>1 Входные и выходные данные</b>	<b>3</b>
<b>2 Преобразование входных данных в выходные</b>	<b>3</b>
<b>3 Примеры работы программы</b>	<b>4</b>
<b>4 Тестирование</b>	<b>5</b>
<b>5 Описание исследования</b>	<b>5</b>
<b>ЗАКЛЮЧЕНИЕ</b>	<b>6</b>

# ВВЕДЕНИЕ

В данной лабораторной работе рассматриваются параллельные вычисления на основе нативных потоков.

Цель работы – получение навыка организации параллельных вычислений на основе нативных потоков.

Задачи работы:

- анализ предметной области;
- разработка алгоритма загрузки данных со страниц сайта eda.ru [eda];
- создание ПО, реализующего разработанный алгоритм;
- исследование зависимости производительности разработанного ПО от количества дополнительных потоков.

## 1 Входные и выходные данные

Входными данными являются адрес главной страницы ресурса (eda.ru [eda], адрес главной страницы не может быть изменён пользователем), максимальное число скачиваемых страниц, количество дополнительных потоков (при получении 0 дополнительные потоки не создаются). Выходными данными является директория output с файлами, содержащими файлы с выгруженными данными со страниц в формате html.

## 2 Преобразование входных данных в выходные

Программа выгружает данные с главной страницы eda.ru [eda] в файл, после чего ищет в нём ссылки на другие страницы данного сайта. далее файл удаляется, а для найденных страниц процедура повторяется. Файлы, в которые были выгружены страницы с рецептами, не удаляются. Ссылкой на страницу с рецептом будет считаться ссылка вида eda.ru/recepty/категория/название. Процесс заканчивается либо после загрузки достаточного количества страниц, либо после загрузки всех страниц, если их оказалось меньше.

### 3 Примеры работы программы

Для реализации данной лабораторной работы был выбран язык *Java*, так как он содержит все необходимые средства для реализации алгоритмов. Нативные потоки создавались при помощи класса *Thread* [**thread**] через явный вызов конструктора.

На рисунке 3.1 представлен пример ввода данных для программы. На рисунке 3.2 представлена папка с выгруженными файлами с рецептами.

```
knife@knife-Swift-SF314-5106:~/labs/aa/lab4/lab4$ java -jar build/libs/lab4.jar
Ресурс: eda.ru
Ввод количества страниц.
Введите целое число не меньше 1: 0
Ошибка. Число меньше 1.
Введите целое число не меньше 1: один
Ошибка. Введено не целое число.
Введите целое число не меньше 1: 1.0
Ошибка. Введено не целое число.
Введите целое число не меньше 1: 10
Ввод количества дополнительных потоков.
Введите целое число в диапазоне от 0 до 16: 3
knife@knife-Swift-SF314-5106:~/labs/aa/lab4/lab4$
```

Рисунок 3.1 — Запуск программы

```

└─ output
  └─ recepty
    └─ osnovnye-blyuda
      <> cvetnaya-kapusta-s-syrom-187504.html
      <> kartofel-hasselbek-187463.html
      <> stebli-seldereya-s-sousom-k-rybe-187479.html
      <> steyk-iz-krasnoy-kapusty-187460.html
      <> tushenaya-tykva-s-rikottoy-187487.html
    └─ rizotto
      <> rizotto-s-belymi-gribami-187339.html
    └─ vypechka-deserty
      <> desert-yabloko-187210.html
      <> pirozhki-s-gribami-187248.html
    └─ zagotovki
      <> tykvennoe-pyure-187634.html
    └─ zakuski
      <> zapechennye-ovoschi-i-griby-so-smetannym-sousom-187527.html
```

Рисунок 3.2 — Папка с рецептами

## 4 Тестирование

Выполнено тестирование программы по методологии чёрного ящика. В таблице 4.1 представлены функциональные тесты. Первое число во вводе означает количество выгружаемых страниц, второе – число дополнительных потоков. Все тесты пройдены успешно.

Таблица 4.1 — Результаты выполнения функциональных тестов

№	Ввод	Ожидаемое количество страниц	Полученное количество страниц
1	10 0	10	10
2	10 1	10	10
3	10 3	10	10
4	25 0	10	10
5	25 1	10	10
6	25 3	10	10
7	100 0	10	10
8	100 1	10	10
9	100 3	10	10

## 5 Описание исследования

Было проведено исследование зависимости скорости работы программы от числа потоков в терминах числа выгруженных страниц с рецептами в минуту. Для замеров времени работы программа выгружала 1000 рецептов, после чего выводилось среднее количество выгруженных страниц в минуту. Время работы было замерено с помощью метода *System.nanoTime()* [nanotime]. Все замеры проводились на ноутбуке Acer Swift 3x, процессор 11th Gen Intel(R) Core(TM) i7-1165G7 [intel], 4 ядра.

Измерения проводились для 0, 1, 2, 4, 8 и 16 дополнительных потоков. Результаты представлены на рисунке 5.1.

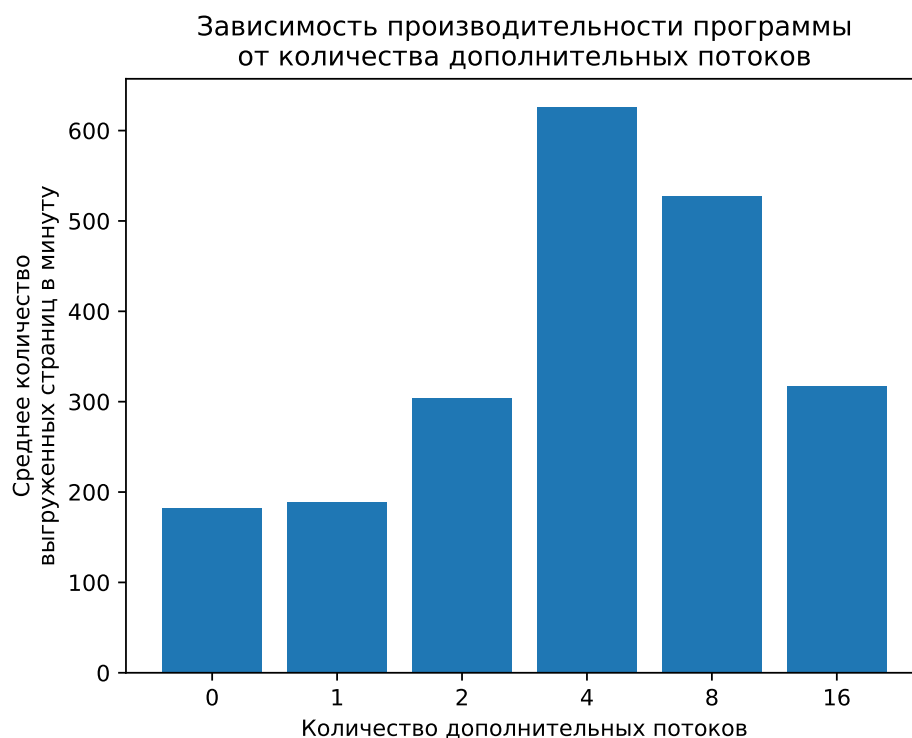


Рисунок 5.1 — Результаты измерений числа выгруженных страниц в минуту

Наибольшая скорость работы была достигнута при совпадении числа потоков с количеством ядер в процессоре. При меньшем числе потоков скорость работы прямо пропорциональна числу потоков (при этом в случаях отсутствия дополнительных потоков и наличия 1 дополнительного потока работает приблизительно с одной и той же скоростью, так как в каждом случае только один поток выполняет вычисления), а при большем – обратно пропорциональна, так как большее число потоков не может быть запущено, но при этом увеличиваются затраты на диспетчеризацию.

## ЗАКЛЮЧЕНИЕ

Цель работы достигнута. Решены все поставленные задачи:

- анализ предметной области;
- разработка алгоритма выгрузки данных со страниц сайта eda.ru [**eda**];
- создание ПО, реализующего разработанный алгоритм;
- исследование зависимости производительности разработанного ПО от количества дополнительных потоков.