



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»

Лабораторная работа №5 по дисциплине «Анализ Алгоритмов»

Тема Организация параллельных вычислений по конвейерному принципу

Студент Куликов Е. А.

Группа ИУ7-56Б

Преподаватель Волкова Л. Л.

Москва — 2024 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	2
1 Входные и выходные данные	2
2 Преобразование входных данных в выходные	2
3 Примеры работы программы	3
4 Тестирование	4
5 Описание исследования	4
ЗАКЛЮЧЕНИЕ	5
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	6

ВВЕДЕНИЕ

Цель работы — Получить навык организации параллельных вычислений по конвейерному принципу.

Задачи работы:

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

1 Входные и выходные данные

Входными данными программы является название папки, в которой находятся html-файлы с рецептами, полученные в результате работы программы из предыдущей лабораторной работы. Выходными данными программы являются: база данных, содержащая информацию о рецептах — порядковый номер, имя файла, заголовок рецепта, ингредиенты и способ приготовления в формате JSON, а также ссылку на основную картинку страницы.

2 Преобразование входных данных в выходные

Программа считывает название папки с рецептами и обрабатывает каждый рецепт по конвейерному принципу: создаются 5 рабочих потоков — поток, генерирующий заявки на обработку (создает структуру, в которой указан путь к файлу с рецептом); поток, выполняющий чтение данных из файла; поток, выполняющих поиск информации в файле; поток, выполняющий запись информации в базу данных *sqlite* с использованием библиотеки *sqlite3* [1]; поток, выполняющий логирование и уничтожение заявки.

После обработки в каждом из потоков заявка попадает в одну из четырех очередей, в которой ожидает обработки следующим потоком.

Последний поток производит уничтожение очередной задачи и логирование — подсчитывает среднее время обработки задачи в каждом потоке, а также время ожидания в каждой из очередей.

Кроме этого в специальной структуре на каждом этапе обработки фиксируется время начала обработки каждой задачи в каждом потоке, а также время постановки задачи в каждую из очередей. Эта информация сохраняется последним потоком и после обработки всех заявок записывается в файл логирования в хронологическом порядке.

3 Примеры работы программы

На рисунках 3.1- 3.2 представлен пример работы программы.

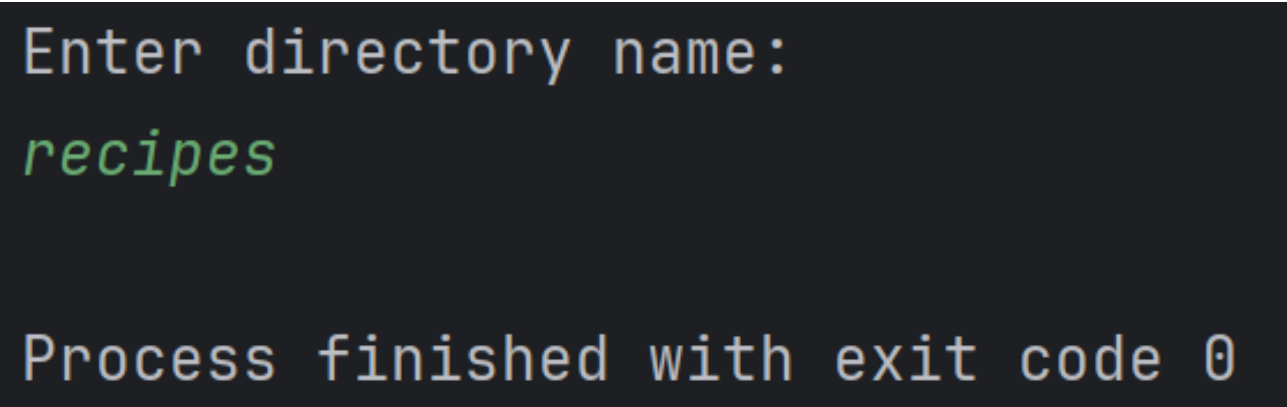


Рисунок 3.1 – Ввод входных данных

WHERE		ORDER BY							
ID	IS...	FI...	TITLE	INGREDIENTS	STEPS	IMAG...			
1	1	9257	recipes\re...	Пицца с яйцом, картоф...	[{"name": "Основа дл...	[{"name": "1. На осн...	https://cdn2...		
2	2	9257	recipes\re...	Смузи с кокосовым мол...	[{"name": "Авокадо -...	[{"name": "1. Разреж...	https://cdn2...		
3	3	9257	recipes\re...	Онигири – пошаговый р...	[{"name": "Ореховая ...	[{"name": "1. Для на...	https://cdn2...		
4	4	9257	recipes\re...	Марокканский морковны...	[{"name": "Очищенная...	[{"name": "1. Тонко ...	https://cdn2...		
5	5	9257	recipes\re...	Пита с куриной грудко...	[{"name": "Куриная г...	[{"name": "1. Для те...	https://cdn2...		
6	6	9257	recipes\re...	Крудите из свежих ово...	[{"name": "Морковь с...	[{"name": "1. Все ко...	https://cdn2...		
7	7	9257	recipes\re...	Русский салат с тепло...	[{"name": "Перепелки...	[{"name": "1. обрабо...	https://cdn2...		
8	8	9257	recipes\re...	Спринг роллы по-грузи...	[{"name": "Кориандр ...	[{"name": "1. Баклаж...	https://cdn2...		
9	9	9257	recipes\re...	Булочки для гамбургер...	[{"name": "Пшеничная...	[{"name": "1. Для оп...	https://cdn2...		
10	10	9257	recipes\re...	Спагетти айоли с черр...	[{"name": "Паста - 8...	[{"name": "1. Отвари...	https://cdn2...		

Рисунок 3.2 – Полученная база данных

4 Тестирование

В таблице 4.1 представлены функциональные тесты для разработанного ПО. Все тесты пройдены успешно. Входные данные: имя папки.

Таблица 4.1 – Функциональные тесты

№ теста	Входные данные	Полученные выходные данные	Ожидаемые выходные данные
1	recipes	база данных на 10 записей	база данных на 10 записей
2	notrecipes	Directory doesnt exist	Directory doesnt exist

5 Описание исследования

В ходе исследования определялось среднее время ожидания заявки в каждой из очередей, время обработки в каждом из потоков и связь этих времен. На рисунке 5.1 представлено среднее время обработки задачи в каждой стадии, а также среднее время ожидания в каждой из очередей.

1	Average existence time: 0.0252966 s.
2	Average generation time: 1.3e-07 s.
3	Average reading time: 0.00024415 s.
4	Average data extraction time: 7.632e-05 s.
5	Average <u>data base</u> writing time: 3.2e-06 s.
6	Average logging and destruction time: 0.00570618 s.
7	Average waiting in after-generation queue time: 0.0011733 s.
8	Average waiting in after-reading queue time: 7.8e-07 s.
9	Average waiting in after-extraction queue time: 9.8e-07 s.
10	Average waiting in after-writing queue time: 0.0237999 s.

Рисунок 5.1 – Характеристики среднего времени обработки и ожидания

По результатам проведенного исследования сделан вывод о том, что самыми долгими стадиями обработки являются чтение html-строки данных из файла, а также логирование и уничтожение заявки. Поэтому аналогично

самое долгое время ожидания наблюдается в очередях перед чтением и перед логированием и уничтожением.

ЗАКЛЮЧЕНИЕ

Цель работы достигнута. Решены все поставленные задачи:

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

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Документация библиотеки sqlite3 [Электронный ресурс]. — Режим доступа: <https://www.sqlite.org/cintro.html> (дата обращения: 22.11.2024).