оглавление

ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ 5

Введение 7

Глава 1. Анализ исходных данных и проектирование 9

1.1 Постановка и анализ задачи 9

1.1.1 Назначение и предметная область 9

1.1.2 Анализ предполагаемых достоинств и недостатков тестирующей системы 10

1.1.3 Выбор средств разработки 11

1.2 Выбор исходного вида системы, ее структуры 12

1.3 Выбор архитектур серверной части для анализа 15

Глава 2. Разработка реализаций сервера и клиента 19

2.1 Установка и настройка необходимых компонентов 19

2.2 Конструирование внешнего вида клиентского приложения 20

2.3 Реализация многопоточного и многопоточного асинхронного серверов 25

2.3.1 Многопоточный сервер 26

2.3.2 Многопоточно-асинхронный сервер 45

Глава 3. Тестирование программного продукта 52

3.1 Тестирование функционирования системы 52

3.2 Анализ результатов 58

3.2.1 Тестирование многопоточного сервера 60

3.2.2 Тестирование многопоточно-асинхронного сервера 67

3.2.3 Общие итоги тестирования, сравнение результатов 74

Заключение 77

Список использованных источников 79

Приложение A. Файл main.cpp генератора нагрузки 80

Приложение Б. Файл server.cpp многопоточного сервера 96

Приложение В. Файл main.cpp многопоточно-асинхронного сервера 100

Приложение Г. Файл server.cpp многопоточно-асинхронного сервера 103

Приложение Д. Файл mainwindow.cpp клиентского приложения 127

Приложение Е. Файлы потока чтения клиентского приложения 134

# ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ

В тексте выпускной квалификационной работы введены специальные термины на русском и английском языках:

| Термин, сокращение | Полная форма |
| --- | --- |
| ПО | Программное обеспечение |
| Лог | Записи событий или действий, сделанные в программе или системе. Логи содержат информацию о произошедших событиях, таких как ошибки, предупреждения, запросы и другие активности в системе. |
| Логгирование | Запись логов |
| Поток | Поток в программировании — это выполнение одного или более процессов в одной программе параллельно или конкурентно. |
| Линкер | Компонент компиляции программного кода, который выполняет связывание объектных файлов, которые были созданы компилятором, в единую исполняемую программу или библиотеку. |
| Проблема гонок | Ситуация, когда несколько процессов или потоков имеют доступ к изменению участка памяти одновременно, что приводит к непредсказуемому поведению программы при сложно прогнозируемых ситуациях. |
| Труба данных | Механизм межпроцессного взаимодействия в операционных системах, который позволяет передавать данные между двумя процессами. |
| Докер | Программное обеспечение для автоматизации развёртывания и управления приложениями в средах с поддержкой контейнеризации, контейнеризатор приложений. |
| Образ | Статический файл, содержащий все необходимые компоненты для запуска приложения – его код, зависимости, библиотеки и операционную систему. |
| Контейнер | Легкий автономный исполняемый пакет программного обеспечения, который включает в себя все необходимые компоненты, необходимые для запуска приложения в изолированной среде. |
| Контейнеризация | Процесс создания контейнера приложения. |
| Тепловая карта (Heatmap) | Графическое представление данных, в котором числовые значения таблицы отображаются при помощи цвета. |
| Корутина | Блок кода, работающий асинхронно с возможностью остановки работы с последующим возобновлением с сохранением состояния. |
| Конвейерная архитектура | Простая, но надежная и эффективная архитектура, состоящая из множества этапов обработки, которые последовательно взаимодействуют с данными. |
| Пул потоков | Набор из работающих потоков, каждый из которых используется для выполнения асинхронных вызовов приложения по необходимости. |
| Заморозка потока | Остановка эффективной работы потока по причине попадания в цикл ожидания выполнения какой-либо операции. |

# Введение

Бэкенд – важнейшее направление IT, на котором базируется все, с чем взаимодействует пользователь. Оно динамично развивается, что не является чем-то необычным, т.к. к нему предъявляются высокие требования по производительности, которые, к тому же, постоянно растут.

Одними из ключевых технологий, применяемых в бекэнде являются многопоточность и асинхронность. Современные языки программирования активно их используют и внедряют как в стандартные, так и в сторонние библиотеки их реализации для упрощения разработки и увеличения производительности программного продукта. Эти две технологии позволяют создавать крайне высокоэффективные программы, рассчитанные на огромное количество пользователей. При их комбинации можно достигнуть еще большей производительности. Любая компания использует их в разработке своего ПО и постоянно стремится найти оптимальное решение своих задач с их помощью. В частности, большинство серверов и сервисных структур основаны на них.

Многопоточность – особенность платформы или приложения, состоящая в разбиении процесса в ОС на некоторое количество потоков, выполняющихся параллельно, т.е. не последовательно друг за другом, что может помочь получить лучшее использование ресурсов компьютера.

Асинхронность – использование событий, которые возникают вне зависимости от основного потока программы и выполнение действий для их обработки в неблокирующем режиме, т.е. основной поток не останавливает свою работу на время их выполнения.

Целью разработки данного программного продукта является создание платформы для сравнения производительности и анализа преимуществ и недостатков относительно друг друга для многопоточной и многопоточно-асинхронной серверных архитектур.

Разработка описанного ПО является уникальной и как позволит разработчикам выбрать необходимую архитектуру для своего продукта, опираясь на визуализированные данные тестирования, так и может стать образцом для создания аналогичных систем. Данный проект будет актуален и крайне полезен начинающим разработчикам, которые смогут при его использовании изучить эти технологии, а также для опытных программистов, которые получат возможность использовать этот подход для своих крупных проектов.

В будущем разрабатываемое ПО может получить дополнительные функции и анализируемые параметры, благодаря легким и приспособляемым инструментам анализа, которые будут доступны пользователю для настройки под свои нужды.

Для создания программного продукта требуется разработать тестируемые реализации серверов – асинхронная и многопоточно-асинхронная, которые после будут встроены в тестирующую систему при помощи программы, имитирующей клиентскую нагрузку, логгирования и системы визуализации собранных данных.

# Анализ исходных данных и проектирование

## Постановка и анализ задачи

### Назначение и предметная область

Разрабатываемое программное обеспечение для сравнения производительности многопоточной и многопоточной асинхронной реализаций серверов с целью нахождения оптимальной технологии для выбранной задачи.

Предметная область, для которой разрабатывается ПО, связана с процессом нагрузочного тестирования и отображает ее с точки зрения профессионального разработчика.

Целью разработки данного программного продукта является создание платформы для сравнения производительности и анализа преимуществ и недостатков относительно друг друга для многопоточной и многопоточно-асинхронной серверных архитектур.

Также стоит принять во внимание необходимость наглядности и понятности замеряемых параметров для конечного пользователя для удобства и облегчения анализа.

Перед разрабатываемым программным продуктом ставится задача реализации следующего функционала:

* запуск, эффективная работа и логгирование тестируемых серверов;
* создание выбранной нагрузки при помощи клиента;
* функционал графического отображения результатов тестирования.

Также необходимо создать реальное клиентское приложение для иллюстрации правдивости работоспособности и применимости тестируемых систем. Однако вне демонстративной части это приложение не будет использоваться, поэтому следует на данном этапе не следует делать его сложнее необходимого.

Для гибкости и универсальности использования ПО дополнительно необходимо использование средства визуализации, способного к изменению формул измеряемых параметров, за счет которой разработчики будут иметь возможность:

* создавать свои новые отслеживаемые параметры для анализа и сравнения тестируемых программ;
* изменять и удалять отслеживаемые изначально параметры;
* настраивать варианты иллюстрации: виды графиков, цвета иллюстрируемых данных и т.д.

### Анализ предполагаемых достоинств и недостатков тестирующей системы

Для оценки актуальности темы и при разработке ПО необходимо предусмотреть его возможные достоинства и недостатки заранее перед началом работы над ним.

Разрабатываемое ПО позволит разработчикам:

* получить платформу для проведения нагрузочного тестирования продукта;
* наглядно увидеть недостатки и преимущества конкретной технологии;
* выбрать для себя идеально подходящий набор архитектурных подходов к созданию ПО;

Тестирующий программный продукт может создавать дополнительную вычислительную нагрузку на серверное оборудование при использовании, однако оно не будет принципиально влиять на его работу в силу некрупного требуемого объема выполняемых подзадач. Новые возможности безусловно перекрывают непринципиальные недостатки, возникающие от использования тестировочного подхода.

Благодаря уникальности и актуальности функционала программного продукта, предполагается, что достоинства разрабатываемого программного продукта более весомы по своей значимости. Таким образом, проект может в будущем стать отличной платформой для образования начинающих разработчиков.

### Выбор средств разработки

Важно выбрать средства для реализации поставленной задачи, соответствующей функциональным требованиям.

**Выбор целевой операционной системы**

Операционная система — комплекс взаимосвязанных программ, предназначенных для управления устройством и организации взаимодействия с пользователем, который предоставляет базовый набор функций по управлению аппаратными средствами устройства.

Для разработки программного обеспечения было решено выбрать ОС Ubuntu Linux, т.к. она проста в использовании для разработчиков, т.к. позволяет быстро и без множества проблем устанавливать необходимы библиотеки, и гибка в настройке при сохранении простоты использования в данном дистрибутиве. Данные достоинства дадут возможность получить необходимый результат при выполнении задачи.

**Выбор языка программирования**

Для разработки любого программного обеспечения необходимо выбрать язык программирования. Так как изучаемое ПО требует максимальной эффективности, необходимо использование высокопроизводительного языка программирования для бекэнд разработки для создания серверов. Был выбран С++ в силу его широких возможностей.

Клиентские приложения, как правило, требуют обширного использования графики, пользовательских интерфейсов, мультимедиа и интерактивности. QT, благодаря обширной библиотеке компонентов и удобной среде разработки, обеспечивает простоту и скорость разработки на разных операционных системах для таких приложений. QT также предоставляет возможности для создания взаимодействующих компонентов, а также инструменты для сборки и развертывания приложений.

В результате, использование QT для разработки клиентского приложения упрощает и ускоряет процесс разработки, обеспечивает высокий уровень функциональности и гибкость пользовательского интерфейса, что делает приложение более привлекательным для пользователей.

Для скриптов, отправляющих данные на сервисы иллюстрации, будет использоваться Python, по причине простоты разработки и наличия удобных библиотек.

## Выбор исходного вида системы, ее структуры

Тестирующая система будет состоять из многих компонентов, связанных друг с другом посредством передаваемых данных.

Основой будет взят докерный контейнер с запущенным серверным приложением. Докер был выбран в силу его простоты и удобства:

А) сервер будет запускаться в легко настраиваемой среде контейнера

Б) проблемы несовместимости будут обходиться изолированностью контейнера от запускающей среды

В) контейнер позволяет выбрать выводящий данные порт, что даст возможность продолжать как считывать, так и отправлять данные

Г) логгируемые данные также возможно считывать, поэтому проблем с их обработкой не должно быть

Д) вносимые в настройки контейнера изменения будут требовать минимального времени на сохранение и использование благодаря “слоям“ контейнера

Далее данные логов будут захвачены и обработаны запущенным python скриптом, который будет пересылать формируемые данные через отдельный порт на сервис Prometheus, замеряющий нужные метрики и запущенный в своем контейнере, чьи данные будут передаваться еще одному контейнеру с запущенным сервисом Grafana.

Prometheus необходим для структуризации данных для последующего анализа в Grafana, который еще и может выводить наглядные и понятные иллюстрации обработанных данных.

Подробное описание серверных архитектур будет приведено в последующем разделе.

**Получение данных**

Сервер будет получать и обрабатывать данные после запуска сессии из двух источников:

* получение запросов от клиента;
* обработка данных из файлов чатов.

Первый способ будет реализован через считывание и обработку клиентских данных из сокета сессии.

Второй способ – будет применяться только к пользователям, находящимся в чате в нынешний момент:

1. при добавлении нового пользователя ему будут отправлены все сохраненные сообщения
2. при отправке нового сообщения сервер отправит всем клиентам все недостающие данные

Хранение чатов должно быть осуществлено при помощи стандартных файлов ОС Linux. Сообщения будут разделяться специальным символом, для каждого пользователя будет храниться позиция в файле, далее которой сообщения еще не отправлены. Для корректной обработки сообщений без опасности появления ошибок в силу многопоточности и асинхронности, необходима синхронизация соответствующего функционала. Она будет обеспечиваться интерфейсом соответствующего класса, что упростит разработку и понимание модели.

Получаемые сервером данные следует зафиксировать определенным набором для стабильности работы системы и простоты обработки. Серверу необходимы тип запроса и данных, отправленных вместе с ним.

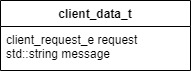


Рисунок 1 – Структура представления данных сервера

**Передача данных**

Для оптимального использования данных следует выбрать единую единицу информации, отправляемую сервером клиенту. Клиенту необходимо знать успешность выполнения операции на сервере, возможные возвращаемые им данные и конкретный тип клиентского запроса, на который этот ответ был сформирован. Альтернативная форма данных создаст проблемы при отладке продукта и усложнит его структуру.

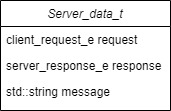


Рисунок 2 – Структура представления данных для клиента

Если обобщать пользователь может получать лишь результат своего запроса и новые сообщения, которые будут обработаны интерфейсом клиентского приложения. Созданная версия клиента будет базовой и упрощенной до необходимого минимума, но сервер должен иметь возможность простой доработки и расширения функционала.

Вся передача данных будет выполняться асинхронно для многопоточно-асинхронной архитектуры и с ожиданием отправки для многопоточного приложения. Такая разница в логике обмена сообщениями очевидно повлияет на производительность, и разница между двумя реализациями будет заметна в итогах анализа.

Пользователь не сможет получить доступ к чатам, пока не займет свободное имя пользователя, после чего ему станет доступен список чатов для подключения, каждый из которых будет иметь свой список пользователей.

## Выбор архитектур серверной части для анализа

Под архитектурой сервера следует понимать устройство компонентов, на которых он базируется, и взаимосвязей между ними. Для решения выбранной задачи следует разработать две архитектуры серверов.

Многопоточный сервер следует свести к максимально простой модели – конвейерная обработка запросов. Каждое из состояний клиента подразумевает различную логику обработки. При переходе из одного состояния в другое клиент будет переходить с одного «конвейера» к другому. Основными этапами будут: получение имени, получение списка чатов и выбор одного из них, нахождение в конкретном чате, отключение от сервера.

Необходимо отметить, что на каждый отдельный чат необходим отдельный поток обработки, который будет запускаться при подключении первого пользователя к нему и останавливаться при отключении от него последнего пользователя. В силу такой логики во избежание проблемы гонок необходимо произвести синхронизацию проверок для запуска и остановки потока.

Также следует выделить потенциальные преимущества и недостатки разрабатываемой архитектуры.

Достоинства:

* минимальное использование ООП результирует в минимальном использовании памяти
* использование проверенных и стабильных библиотек линукса дает уверенность в стабильности и прогнозируемости работы системы
* разделение этапов обработки данных в разные потоки облегчает отладку логики программы и позволяет использовать дополнительную производительность от многопоточности

Недостатки:

* передача данных от потока к потоку потенциально связана с проблемой гонок и ошибками в передаче данных
* запуск нового потока для каждого чата – ресурсозатратное действие
* остановка потока для ожидания нового сообщения уменьшает эффективность использования ресурсов потоков

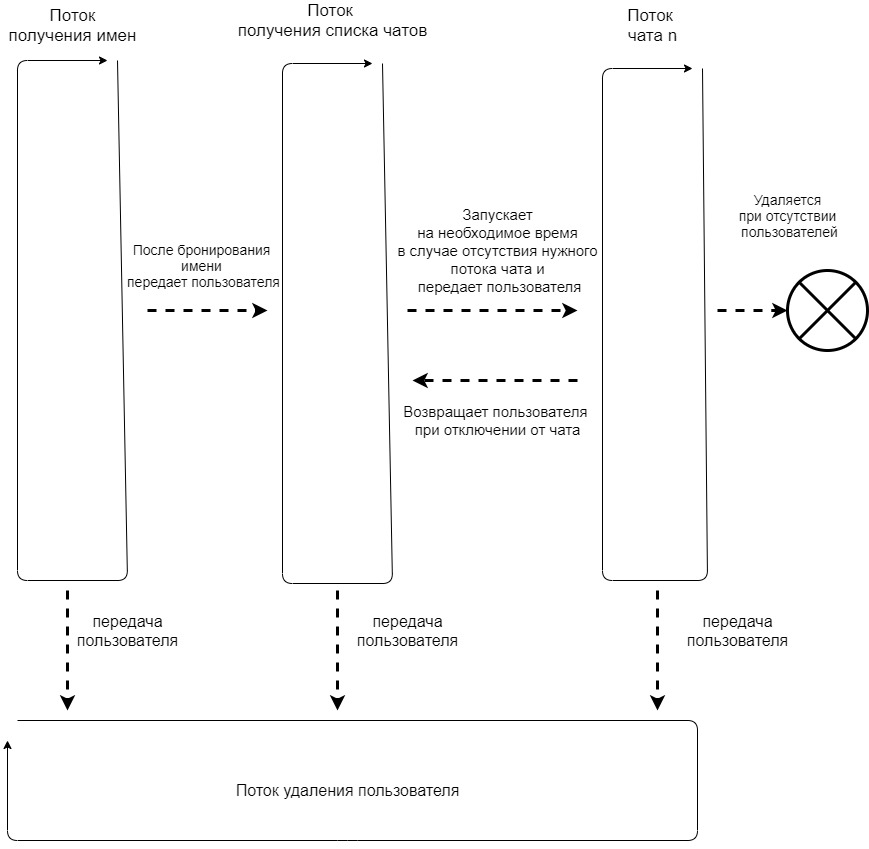


Рисунок 3 – Архитектура многопоточного сервера

Многопоточный асинхронный сервер имеет несколько более комплексную логику относительно предыдущего. Асинхронные задачи должны выполняться синхронизировано относительно задач своих групп, при том нужно различать что нужно синхронизировать относительно выбранной задачи, а что – нет. Поэтому для наиболее логичной организации был выбраны ООП в совместном использовании с корутинами.

Это позволяет создать необходимые интерфейсы классов, которые автоматически будут производить необходимые синхронизации в нужных контекстах, а корутины дают возможность простой разработки асинхронной логики выполнения задач.  
Модель устроена следующим образом: объекты класса чата хранятся и корректно управляются классом менеджера чатов, оба этих класса имеют свои уровни синхронизации. Класс сервера хранит в себе менеджер чатов, к которому будет обращаться создаваемый и запускаемый при подключении к серверу объект корутины сессии, который будет заниматься обработкой запросов пользователя асинхронно и параллельно работе сервера.

Такой архитектурный подход дает набор преимуществ:

* относительная простота разработки логики
* оптимизированность работы сервера за счет минимизации периодов ожидания какого-то действия благодаря использованию асинхронности
* безопасное использование ресурсов памяти – С++ известен возможными проблемами с ошибками памяти, максимальный упор на ООП дает возможность использовать преимущества высокопроизводительного языка с минимальным риском столкнуться с вышеописанным недостатком при правильной практике программирования – использованием умных указателей в качестве автоматизации защиты памяти
* использование пула потоков для асинхронной работы объединяет преимущества многопоточности и асинхронности в плане производительности
* инкапсуляция позволяет скрыть детали реализации от других объектов, обеспечивая простой и безопасный доступ к методам и свойствам объектов

Однако стоит отметить и потенциальные недостатки:

* большой упор на использование библиотек потенциально опасен сложностью в отладке ошибок
* ООП, дающий понятность и простоту модели, может увеличить расход памяти

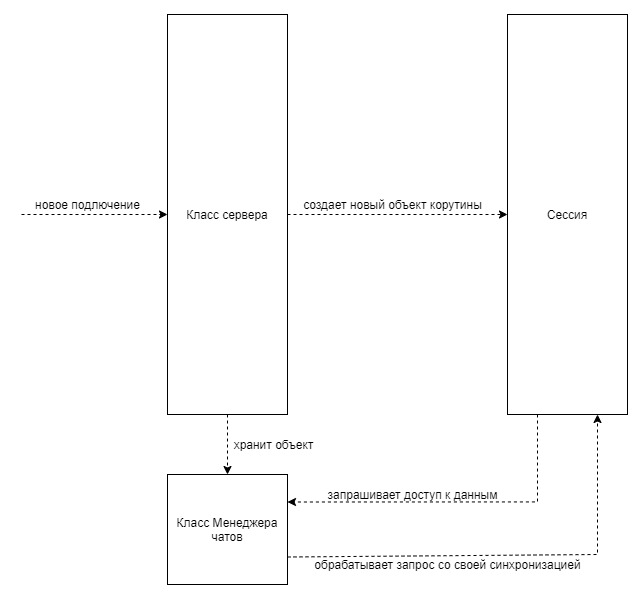


Рисунок 4 – Архитектура многопоточно-асинхронного сервера

# Разработка реализаций сервера и клиента

## Установка и настройка необходимых компонентов

Перед непосредственной разработкой программного обеспечения требуется в первую очередь установить и настроить к работе необходимые компоненты – библиотеку BOOST и линкер CMake.

BOOST устанавливается при использовании менеджера зависимостей для С/С++ - Conan. Удобство данного инструмента главным образом заключается в полной автоматизации настройки среды разработки: достаточно прописать необходимые для разработки библиотеки и используемый линкер – в нашем случае это CMake.Также Conan приспособлен к кроссплатформенности, поддерживает большинство популярных библиотек, позволяет создавать свои.

Он устанавливается через скачивание необходимой версии с необходимого репозитория через терминал Linux.

Основные преимущества CMake:

1. Кроссплатформенность. CMake позволяет создавать кроссплатформенные проекты, а это многократно упрощает разработку, а также поддержку программного обеспечения относительно различных операционных систем, архитектур и компиляторов.
2. Простота и гибкость. Простой и гибкий синтаксис позволяет быстро и легко описывать собираемые проекты и их зависимости.
3. Автоматизация сборки.
4. Интеграция с другими инструментами. Многие IDE, менеджеры зависимости и другие инструменты разработки интегрируют CMake, а это дает возможность существенно ускорить и упростить процесс разработки.
5. Поддержка модульной структуры.

Преимущества библиотеки BOOST:

1. Широкий спектр функций, в том числе работа с сетью, вводом/выводом, строками, файлами и многим другим.
2. Обширная документация.
3. Проверенность и надежность – библиотека прошла множество тестов и признана качественным инструментов для разработчиков.
4. Поддерживает последние стандарты С++, а это очень пригодится мне в разработке.

## Конструирование внешнего вида клиентского приложения

При разработке внешнего вида демонстрационной версии клиентского приложения было нужно отметить, что клиент в описанной системе находится в трех ключевых состояниях:

1. Без имени
2. Вне чата
3. В чате

Каждому из этих состояний должен соответствовать свой режим работы приложения.

Для первого состояния достаточно указать требования к формату вводимого имени, выделить поле для ввода и обозначить название приложения наверху.

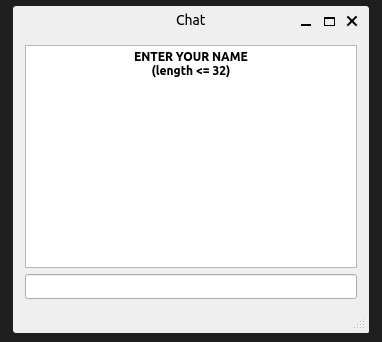


Рисунок 5 – Иллюстрация GUI регистрации

Если пользователь пытается занять уже используемое имя, он будет уведомлен об этом всплывающим окном. После его закрытия поле ввода имени будет очищено для нового ввода нового варианта имени пользователя.

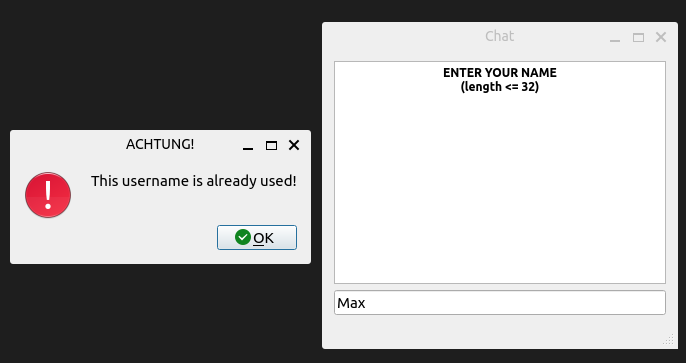


Рисунок 6 – Всплывающее окно при отказе в регистрации имени

В обратном случае приложение зарегистрирует пользователя и перейдет ко второму из режимов работы.

Для второго состояния необходимо рассмотреть чуть более сложную концепцию: постоянно обновляющийся список доступных чатов с полем для создания своего нового.

Каждый чат в списке должен быть представлен динамически созданной кнопкой подключения. После нажатия на любую из них на сервер будет передан запрос на подключение к выбранному чату, а приложение сменит свой режим работы на третье.

Стоит отдельно отметить, что при создании нового чата список чатов должен быть обновлен для каждого из пользователей, который зарегистрирован, но все еще находится вне какого-то либо чата.

Таким образом, пока пользователь не войдет в чат, список доступных, т.е. таким образом список кнопок подключения должен иметь возможность постоянного обновления.

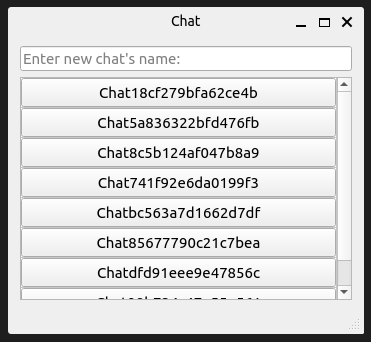


Рисунок 7 – GUI выбора чата из списка доступных

Если же пользователь попытается создать чат с неправильным (содержащим пробелы или излишне длинным) названием – он будет уведомлен о невозможности выполнения такого действия всплывающим окном. Поле ввода названия нового чата будет очищено после закрытия всплывающего окна.

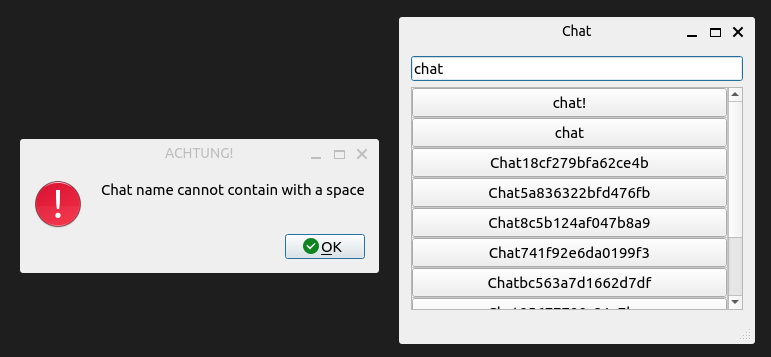


Рисунок 8 – Всплывающее окно при отказе в создании чата с неверным именем

Если же будет предпринята попытка создания чата с уже занятым именем, то запрос будет проигнорирован, а пользователь все еще сможет подключиться к чату с желаемым им именем.

При нажатии на кнопку выбранного чата приложение перейдет в последний, третий режим работы.

В чате пользователю необходимо предоставить следующий функционал:

1. Возможность покинуть чат
2. Возможность прочитать сообщения и пролистать их до более старых или новых
3. Возможность отправки своего сообщения

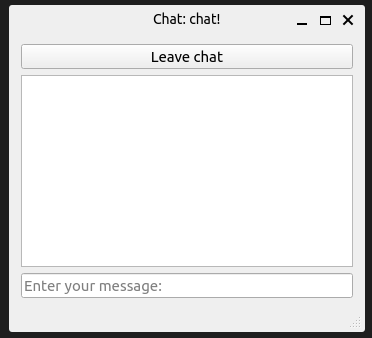


Рисунок 9 – GUI чата

Cверху будет указано название чата, к которому подключен пользователь

Сообщение вводится в нижнее поле и отправляется нажатием Enter. После отправки сообщение рассылается всем пользователям, подключенным к чату, и отражается в списке сообщений следующим образом:

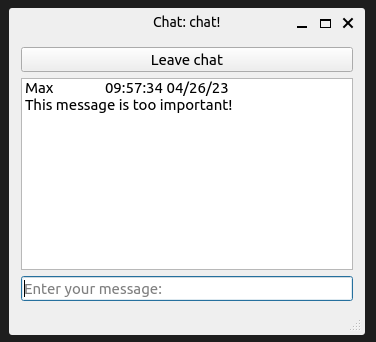


Рисунок 10 – Отображение сообщения в GUI чата

У каждого сообщения отмечаются автор, время и дата отправки помимо самого содержания.

Нажатие на кнопку выхода из чата пользователь переходит обратно во второй режим работы приложения и получает актуальный список чатов, откуда может как снова вернуться в изначальный чат, так и подключиться к другому.

При закрытии приложения пользователь отключается и выходит из системы. Его пользовательское имя освобождается для остальных и может быть уже занято при последующем подключении.

## Реализация многопоточного и многопоточного асинхронного серверов

|  |
| --- |
| class Logger {  Logger() = default;  Logger(const Logger&) = delete;  public:  static Logger& GetInstance() {  static Logger obj;  return obj;  }  // Выведите в поток все аргументы.  void Log(const std::string\_view, boost::json::value);  void InitBoostLog(); |

**В обоих реализациях** серверов используется класс Logger, отвечающий за создание и вывод логов:

Данный класс является ярким примером принципа Singleton: невозможно получить в распоряжение более одного объекта класса, таким образом все настройки, заданные при помощи метода **InitBoostLog**, будут легко применимы по всему проекту:

|  |
| --- |
| void Logger::InitBoostLog(){  //setting up logging  logging::add\_common\_attributes();  logging::add\_console\_log(  std::cout,  keywords::format = &logger::MyFormatter,  keywords::auto\_flush = true  );  } |

В данном случае настраивается использование атрибутов системой логгирования, печать логов в стандартный поток вывода без задержки по настроенному формату.

### Многопоточный сервер

**Основные рутины потоков сервера**

**Рутина user\_name\_enter отвечает за регистрацию пользователей в системе.**

На вход ей подаются такие параметры, как входной и выходной конец трубы данных (pipe). Рутина получает новых пользователей из входного конца при их подключении к серверу, после чего добавляет их в список отслеживаемых и переходит к их обработке запросов на регистрацию.

|  |
| --- |
| while((bytes = read(end\_to\_read\_from, &new\_socket, sizeof(int))) > 0) {  ev.data.fd = new\_socket;  epoll\_ctl(no\_name\_epoll\_fd, EPOLL\_CTL\_ADD, ev.data.fd, &ev);  } |

При удачной регистрации пользователь удаляется из отслеживаемых и по выходному концу передается в следующую рутину – list\_of\_chats.

|  |
| --- |
| user\_data[client\_sockfd] = received.message\_text;  used\_usernames.insert(received.message\_text);  resp.responce = s\_success;  epoll\_ctl(no\_name\_epoll\_fd, EPOLL\_CTL\_DEL, client\_sockfd, &ev);  if(!send\_resp\_to\_client(&resp, client\_sockfd)) write(disco\_pipe[1], &client\_sockfd, sizeof(int));  std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(end\_to\_write\_to, &client\_sockfd, sizeof(int)); |

Если пользователь отключается на моменте регистрации – он передается рутине disconnect для проведения процедур отключения от системы.

|  |
| --- |
| if(received.request == c\_disconnect || err == 0){  epoll\_ctl(no\_name\_epoll\_fd, EPOLL\_CTL\_DEL, client\_sockfd, &ev);  std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(disco\_pipe[1], &client\_sockfd, sizeof(int));  boost::json::value data = {  {"socket", client\_sockfd}  };  logger::Logger::GetInstance().Log("User disconnected"sv, data);  continue;  } |

Обе этих рутины будут также рассмотрены чуть позже.

Бесконечный цикл обеспечивает беспрерывную обработку запросов, получение новых сообщений опирается на использование epoll.

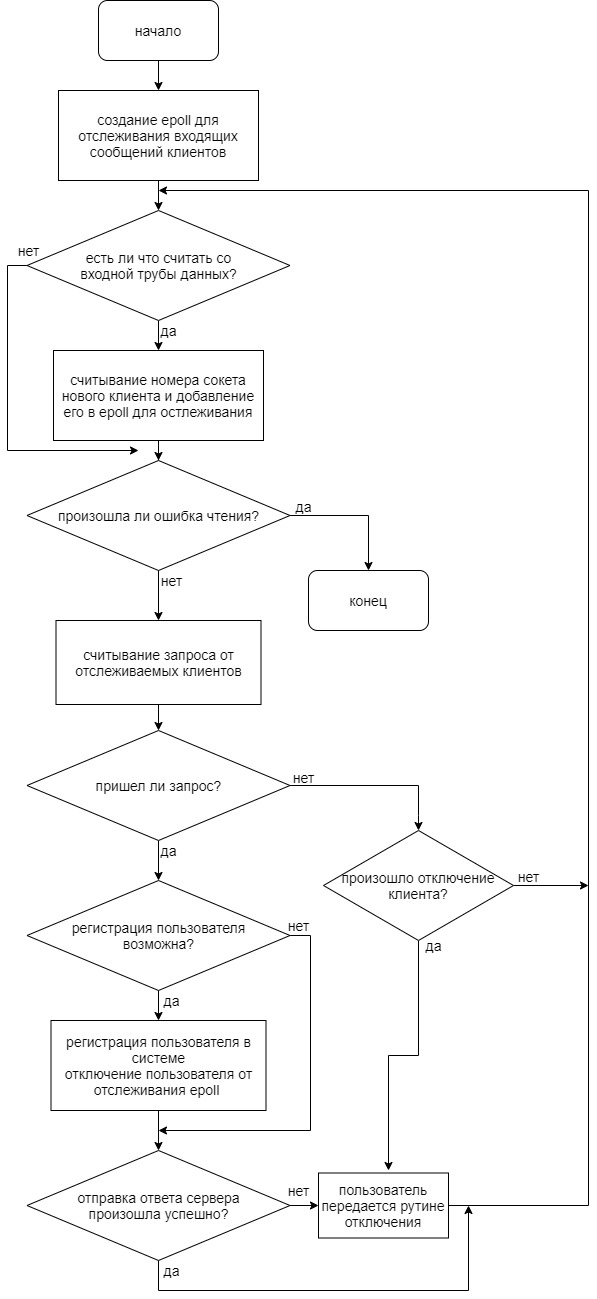


Рисунок 11 - Структурная схема алгоритма user\_name\_enter

**Рутина list\_of\_chats отвечает за предоставление зарегистрированным пользователям вне чатов актуального списка доступных.**

В этой рутине предусмотрены обработка подключения к чату, создания нового чата, отключения от системы.

Подключение к чату:

|  |
| --- |
| case c\_connect\_chat:  {  resp.request = c\_connect\_chat;  bool start\_ok = true;  if (chats.count(received.message\_text) == 0){  resp.responce = s\_failure;  resp.message\_text = "Can't connect to a non-existant chat";  break;  }  //if chat was unused -> setup  chats[received.message\_text].mtx.lock();  if(chats[received.message\_text].pipe.in = -1){  int tmp[2];  for(auto socket : closed\_sockets){ // handle closed sockets  epoll\_ctl(no\_chat\_epoll\_fd, EPOLL\_CTL\_DEL, socket, &ev);  clients\_without\_chat.erase(socket); std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(disco\_pipe[1], &socket, sizeof(int));  }  }  break;  } |

Создание нового чата:

|  |
| --- |
| case c\_create\_chat:  {  resp.request = c\_create\_chat;  int check = creat((received.message\_text + ".chat").c\_str(), S\_IRWXG | S\_IRWXO | S\_IRWXU);  if (errno == EEXIST) {  resp.responce = s\_failure;  resp.message\_text = "Chat with such name already exists";  if(!send\_resp\_to\_client(&resp, client\_sockfd)) goto c\_disconnect\_list\_chats;  }  else if (check < 0){  resp.responce = s\_failure;  resp.message\_text = "Server error: could not create chat, try again later";  boost::json::value data = {  {"chat", received.message\_text},  {"error", std::error\_code{errno, std::generic\_category()}.message()},  };  logger::Logger::GetInstance().Log("Chat creation error"sv, data);  if(!send\_resp\_to\_client(&resp, client\_sockfd)) goto c\_disconnect\_list\_chats;  }  else{  chats[received.message\_text].subs = {};  chats[received.message\_text].pipe.in = -1;  chats[received.message\_text].pipe.out = -1;  boost::json::value data = {  {"chat", received.message\_text}  };  logger::Logger::GetInstance().Log("Chat created"sv, data);  std::unordered\_set<int> closed\_sockets;  for (int socket : clients\_without\_chat){  if(!send\_available\_chats(socket)){ // if client disconnects while being written into  closed\_sockets.insert(socket);  }  }  for(auto socket : closed\_sockets){ // handle closed sockets  epoll\_ctl(no\_chat\_epoll\_fd, EPOLL\_CTL\_DEL, socket, &ev);  clients\_without\_chat.erase(socket); std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(disco\_pipe[1], &socket, sizeof(int));  }  }  break;  } |

Отключение от системы:

|  |
| --- |
| case c\_disconnect:  c\_disconnect\_list\_chats:  epoll\_ctl(no\_chat\_epoll\_fd, EPOLL\_CTL\_DEL, client\_sockfd, &ev);  clients\_without\_chat.erase(client\_sockfd);  std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(disco\_pipe[1], &client\_sockfd, sizeof(int));  break; |

Новые пользователи считываются из рутины username\_enter, если же пользователь отключается от системы, он все так же передается в рутину disconnect.

При подключении к чату возникает необходимость в синхронизированной проверке наличия в нем пользователей. Если чат пуст, то его отдельный поток не запущен, тогда нужно провести процедуры подготовки к запуску рутины чата в отдельном потоке и только после передать ей пользователя.

В этой рутине предусмотрены обработка подключения к чату, создания нового чата, отключения от системы.

Новые пользователи считываются из рутины username\_enter, если же пользователь отключается от системы, он все так же передается в рутину disconnect.

При подключении к чату возникает необходимость в синхронизированной проверке наличия в нем пользователей. Если чат пуст, то его отдельный поток не запущен, тогда нужно провести процедуры подготовки к запуску рутины чата в отдельном потоке и только после передать ей пользователя.

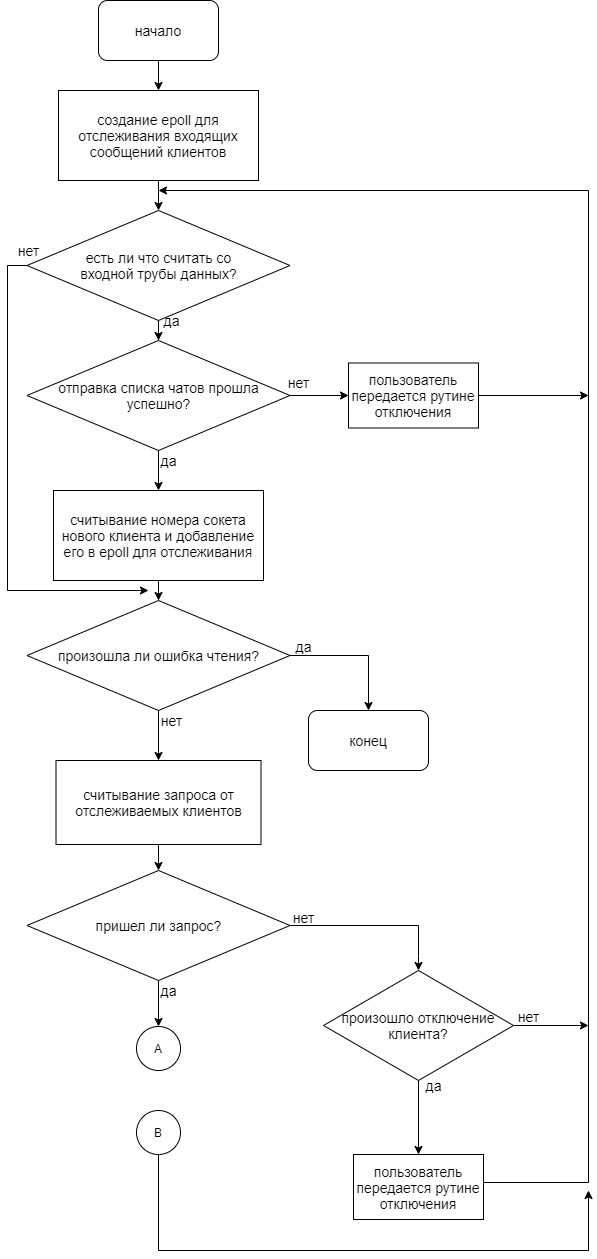


Рисунок 12 - Структурная схема алгоритма list\_of\_chats без обработки данных

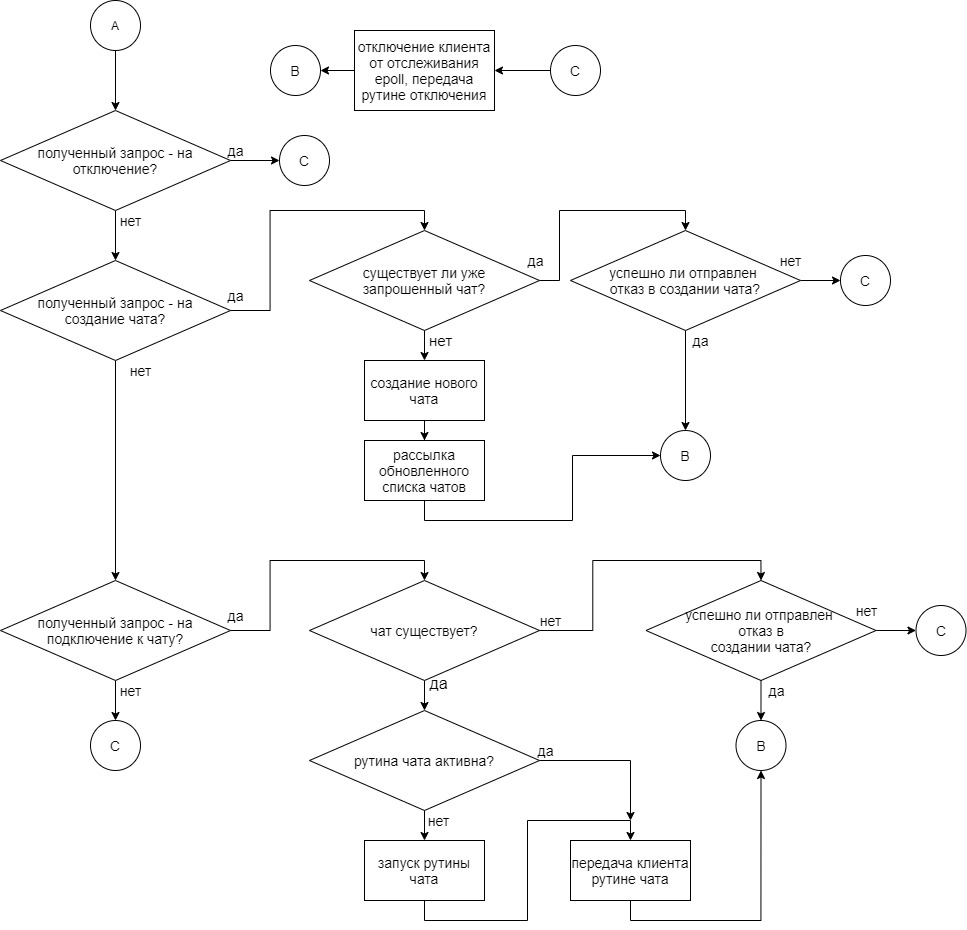


Рисунок 13 - Структурная схема алгоритма обработки запросов list\_of\_chats

**Рутина chat\_message отвечает за обработку логики каждого конкретного чата в отдельном потоке.**

В этой рутине предусмотрен функционал выхода из чата, отключения от системы и отправки нового сообщения.

Выход из чата:

|  |
| --- |
| case c\_leave\_chat:  {  if(epoll\_ctl(chat\_epoll\_fd, EPOLL\_CTL\_DEL, fd, &ev) == -1){  boost::json::value data = {  {"error", std::error\_code{errno, std::generic\_category()}.message()},  {"socket", fd},  {"chat", chat\_name}  };  logger::Logger::GetInstance().Log("Failed to remove socket from epoll"sv, data);  }  messages\_offset.erase(fd);  chats[chat\_name].subs.erase(fd); //unsubscribing  std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  if(write(list\_chats\_pipe[1], &fd, sizeof(int)) < 0){  boost::json::value data = {  {"error", std::error\_code{errno, std::generic\_category()}.message()},  {"socket", fd},  {"chat", chat\_name}  };  logger::Logger::GetInstance().Log("Failed to send client's fd to no\_chat routine"sv, data);  }  boost::json::value data = {  {"chat", chat\_name},  {"socket", fd}  };  logger::Logger::GetInstance().Log("User left chat"sv, data);  }  goto OUT\_OF\_CYCLE; |

Отключение от системы:

|  |
| --- |
| case c\_disconnect:  {  boost::json::value data = {  {"chat", chat\_name},  {"socket", fd}  };  logger::Logger::GetInstance().Log("Disconnecting user"sv, data);  if(epoll\_ctl(chat\_epoll\_fd, EPOLL\_CTL\_DEL, fd, &ev) == -1){  boost::json::value data = {  {"error", std::error\_code{errno, std::generic\_category()}.message()},  {"socket", fd},  {"chat", chat\_name}  };  logger::Logger::GetInstance().Log("Failed to remove socket from epoll"sv, data);  }  messages\_offset.erase(fd);  chats[chat\_name].subs.erase(fd); //unsubscribing  std::atomic\_thread\_fence(std::memory\_order::memory\_order\_seq\_cst);  write(disco\_pipe[1], &fd, sizeof(int));  }  goto OUT\_OF\_CYCLE; |

Отправка сообщения:

|  |
| --- |
| case c\_send\_message:  {  boost::json::value data = {  {"chat", chat\_name},  {"socket", fd},  {"message", client\_data.message\_text}  };  logger::Logger::GetInstance().Log("Sending message"sv, data);  chat\_file << user\_data[fd] << '\t';  chat\_file.write(timedate\_string, strlen(timedate\_string));  chat\_file << std::endl << client\_data.message\_text << "\r";  }  break; |

После каждой итерации цикла рутины происходит отправка недостающих сообщений для каждого пользователя чата, а также отключение пользователей, закрывших соединение.

Стоит отметить, что во избежание ошибок синхронизации отправки сообщений при каждой итерации цикла алгоритма проходит отправка потенциально неотправленных сообщений всем участникам чата.

В этой рутине, как и во всех остальных, можно встретить использование std::atomic\_thread\_fence. Этот класс используется для того, чтобы предотвратить перестановку операций компилятором во время оптимизаций.

Для понятной иллюстрации алгоритма работы этой рутины достаточно показать часть с обработкой запросов, т.к. основная его часть практически неотличима от предыдущей рутины, за исключением того, что для успешного запуска рутины чата необходимо успешное открытие файла, и того, что при отсутствии подключенных пользователей рутина остановится.

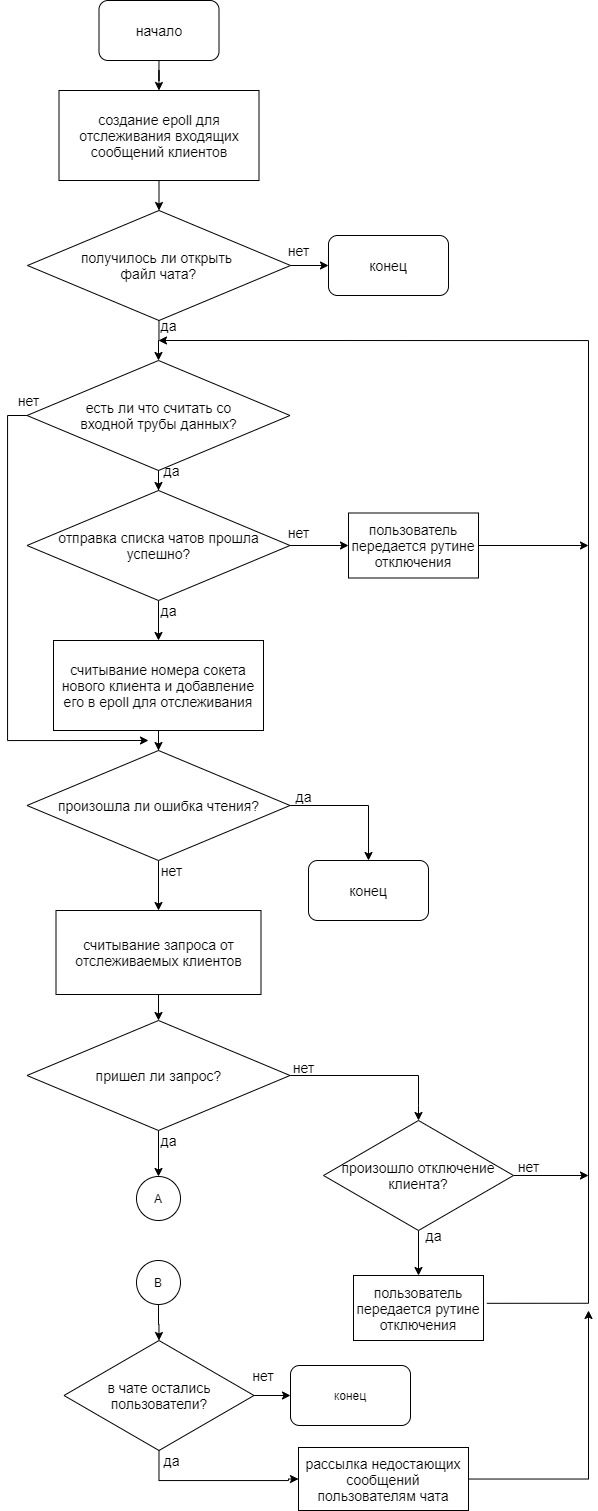


Рисунок 14 - Структурная схема алгоритма chat\_message без обработки данных

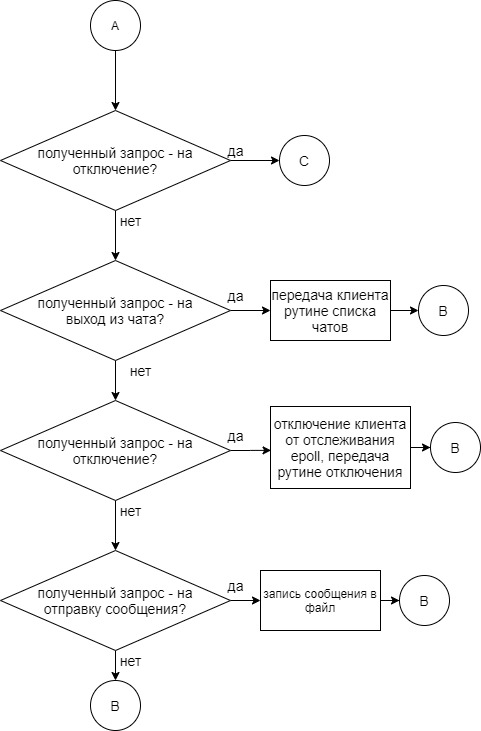


Рисунок 15 - Структурная схема алгоритма обработки запросов рутины chat\_message

**Последняя рутина, disconnect, отвечает за отключение пользователя:**

|  |
| --- |
| void disconnect(){  int read\_bytes, fd;  while (true) {  while ((read\_bytes = read(disco\_pipe[0], &fd, sizeof(int))) > 0) {  boost::json::value data = {  {"socket", fd},  {"name", user\_data[fd]}  };  logger::Logger::GetInstance().Log("User was disconnected"sv, data);  end\_resp\_to\_client(fd);  used\_usernames.erase(user\_data[fd]);  user\_data.erase(fd);  }  if (read\_bytes < 0 && errno != EAGAIN && errno != EINTR) {  boost::json::value data = {  {"error", std::error\_code{errno, std::generic\_category()}.message()}  };  logger::Logger::GetInstance().Log("Pipe read error disconnected"sv, data);  }  }  } |

В этой рутине происходит освобождение занятого имени пользователя и удаление промежуточных данных о нем из системы.

**Во всех** рутинах, описанных выше, применяются трубы (pipes). Они имеют максимальный размер данных, защищенных от проблем гонок. Нам не требуется синхронизация относительно них т.к. передаваемые данные пользователей меньше заданного размера.

Для данной рутины структурная схема алгоритма не является необходимой в силу ее относительной простоты.

**Функционал отправки и получения данных**

Функция read\_request\_from\_client используется для простоты и единообразия считывания данных во всей реализации сервера.

|  |
| --- |
| int read\_request\_from\_client(client\_data\_t\* received, int sockfd){  size\_t length;  int err;  err = recv(sockfd, (void\*)(&(received->request)), sizeof(client\_request\_e), MSG\_WAITALL);  if (err != sizeof(client\_request\_e))  return err;  err = recv(sockfd, (void\*)(&length), sizeof(size\_t), MSG\_WAITALL);  if (err != sizeof(size\_t))  return err;  received->message\_text.resize(length);  err = recv(sockfd, (void\*)(received->message\_text.data()), length, MSG\_WAITALL);  if (err != length)  return err;  boost::json::value data = {  {"request",  {  {"client\_request\_e", StringifyRequest(received->request)},  {"message", received->message\_text}  }  }  };  logger::Logger::GetInstance().Log("Read request from client"sv, data);  return length + 1;  } |

Функция send\_resp\_to\_client выполняет аналогичные задачи относительно отправки данных пользователю.

|  |
| --- |
| bool send\_resp\_to\_client(const server\_data\_t\* resp, int sockfd){  #if DEBUG\_TRACE  printf("%d : - send\_resp\_to\_client()\n", getpid());  #endif  int written\_bytes;  written\_bytes = send(sockfd, (void\*)(&resp->request), sizeof(client\_request\_e), MSG\_WAITALL | MSG\_NOSIGNAL);  if(written\_bytes < 0) return false;  written\_bytes = send(sockfd, (void\*)(&resp->responce), sizeof(server\_response\_e), MSG\_WAITALL | MSG\_NOSIGNAL);  if(written\_bytes < 0) return false;  size\_t length = resp->message\_text.size();  written\_bytes = send(sockfd, (void\*)(&length), sizeof(size\_t), MSG\_WAITALL | MSG\_NOSIGNAL);  if(written\_bytes < 0) return false;  written\_bytes = send(sockfd, (void\*)(resp->message\_text.c\_str()), resp->message\_text.size(), MSG\_WAITALL | MSG\_NOSIGNAL);  if(written\_bytes < 0) return false;  boost::json::value data = {  {"response",  {  {"client\_request\_e", StringifyRequest(resp->request)},  {"server\_response\_e", StringifyResponse(resp->responce)},  {"message", resp->message\_text}  }  }  };  logger::Logger::GetInstance().Log("Sent response to client"sv, data);  return true;  } |

**Инициализация сервера**

После настройки параметров вроде адреса и порта сервера нужно начать отслеживать уже созданные ранее чаты, за это отвечает отдельная функция – server\_starting.

|  |
| --- |
| int server\_starting(){  namespace fs = std::filesystem;  boost::json::value data = { {"pid", getpid()} };  logger::Logger::GetInstance().Log("Server started"sv, data);  fs::path path\_of\_chats(".");  for (const auto& entry : fs::directory\_iterator(path\_of\_chats)) {  if (entry.is\_regular\_file() && fs::path(entry).extension() == ".chat") {  chats[fs::path(entry).stem()].subs = {};  }  }  return 1;  } |

Если файл, обнаруженный в выбранной папке чатов, имеет расширение .chat, то файл чата учитывается системой, его название записывается, тем самым делая его доступным при дальнейшей работе.

### Многопоточно-асинхронный сервер

Эта версия сервера в куда большей степени опирается на ООП, поэтому необходимо в первую очередь описать ключевые классы

**Класс Chat** – помощь в упрощении синхронизации работы с чатами и удобный интерфейс.

Конструктор класса выполняет не только стандартную задачу создания объекта класса, но и создает файл чата при его отсутствии на момент запуска.

Метод AddUser отвечает за добавление нового пользователя в чат. После подключения ему будут все сообщения, отправленные до запуска метода.

Метод SendMessage отправляет сообщение, если указанный пользователь находится в чате. Все пользователи чата получат отправленное сообщение.

Метод ContainsUser отвечает за проверку наличия пользователя в чате. Используется по большей части описанным позже классом ChatManager для корректной работы всей системы.

Метод DeleteUser отвечает за отключение пользователя из чата, если он вообще там находится.

**Каждый** из этих методов синхронизируется через инкапсулированный объект strand[1] таким образом, чтобы все асинхронные запросы к методам класса выполнялись отдельно и последовательно.

|  |
| --- |
| class Chat{  public:  Chat(net::io\_context& io, const std::filesystem::path& chat\_path) : strand(net::make\_strand(io)), path(chat\_path){  // file.exceptions(std::ifstream::failbit | std::ifstream::badbit );  try{  std::ofstream {path, std::fstream::app};  // If no file is created, then show the error message.  } catch(std::ifstream::failure e){  throw e;  }  }  boost::asio::awaitable<void> AddUser(std::shared\_ptr<tcp::socket> socket, std::string username);  boost::asio::awaitable<void> SendMessage(std::shared\_ptr<tcp::socket> socket, std::string message);  boost::asio::awaitable<bool> ContainsUser(std::shared\_ptr<tcp::socket> socket);  boost::asio::awaitable<void> DeleteUser(std::shared\_ptr<tcp::socket> socket);  private:  net::strand<net::io\_context::executor\_type> strand;  std::unordered\_map<std::shared\_ptr<tcp::socket>, std::streamoff> messages\_offset; // Session to offset in chat file  std::unordered\_map<std::shared\_ptr<tcp::socket>, std::string> usernames; // Session to username  std::filesystem::path path; // Path to the file  }; |

Также класс хранит имена пользователей и соответствующие каждому пользователю сдвиг внутри файла чата, а также путь до самого файла.

**Класс ChatManager** – основа логики сервера, он содержит занятые имена пользователя, управляет созданием новых чатов, обновлением их списка и подключением к любому из них.

|  |
| --- |
| class ChatManager{  public:  ChatManager(net::io\_context& io, std::string path) : io\_(io), chats\_strand\_(net::make\_strand(io)), usernames\_strand\_(net::make\_strand(io)), path\_of\_chats(path.data()){  namespace fs = std::filesystem;  if(!fs::exists(path\_of\_chats) || !fs::is\_directory(path\_of\_chats)) fs::create\_directory(path\_of\_chats);  for (const auto& entry : fs::directory\_iterator(path\_of\_chats)) {  if (entry.is\_regular\_file() && fs::path(entry).extension() == ".chat") {  chats\_.insert({fs::path(entry).stem().string(), Chat(io, entry.path())});  }  }  }  boost::asio::awaitable<bool> SetName(std::string, std::shared\_ptr<tcp::socket>);  boost::asio::awaitable<std::string> ChatList();  boost::asio::awaitable<void> UpdateChatList(std::shared\_ptr<tcp::socket>);  boost::asio::awaitable<void> ConnectChat(std::string, std::shared\_ptr<tcp::socket>);  boost::asio::awaitable<bool> CreateChat(std::string);  boost::asio::awaitable<bool> LeaveChat(std::shared\_ptr<tcp::socket>);  boost::asio::awaitable<void> SendMessage(std::string, std::shared\_ptr<tcp::socket>);  boost::asio::awaitable<void> Disconnect(std::shared\_ptr<tcp::socket>);  private:  std::unordered\_map<std::string, Chat> chats\_; // чаты с юзернеймами  std::unordered\_map<std::shared\_ptr<tcp::socket>, User> users\_; // юзеры  net::io\_context& io\_;  net::strand<net::io\_context::executor\_type> chats\_strand\_, usernames\_strand\_;  std::filesystem::path path\_of\_chats;  }; |

Конструктор помимо основной своей функции также выполняет задачу учета всех файлов чатов, созданных ранее.

Метод SetName отвечает за установку имени пользователя при регистрации. До прохождения этого этапа пользователь не может взаимодействовать с чатами.

Метод ChatList выдает актуальный список чатов при использовании. Применяется в UpdateChatList и не только.

Метод UpdateChatList рассылает актуальный список чатов всем зарегистрированным пользователям, находящимся вне чатов.

Метод ConnectChat подключает пользователя к запрошенному чату, если он существует.

Метод CreateChat создает новый чат, если он еще не существует.

Метод LeaveChat удаляет пользователя из выбранного чата, если он находится в нем.

Метод SendMessage позволяет пользователю отправить сообщение, если он на момент запроса находится в чате.

Метод Disconnect отвечает за отключение отсоединившихся пользователей от системы. Их данные будут удалены, а имя пользователя – освобождено.

Все методы используют синхронизацию через два strand – относительно имен пользователей и относительно списка чатов.

**Оба предыдущих класса во всех методах используют корутины, т.к. нужно асинхронно выполнять запросы к ним из корутины сессии.**

|  |
| --- |
| boost::asio::awaitable<void> Session(tcp::socket&& socket\_, ChatManager\* chatManager); |

Использование корутин дает возможность применить всю мощь асинхронности: в нужный момент выполнение блока кода корутины приостанавливается, сохраняя последнее состояние. После того, как придет очередь вернуть управление остановленной корутине, она продолжит свою работу с сохраненными свойствами.

**Класс Server** – класс, на который опирается вся логика подключения к системе новых пользователей.

Класс сервера постоянно находится в динамической памяти, продолжая принудительно поддерживать свой срок жизни, захватывая при каждом вызове методов shared\_ptr на самого себя. Как только сервер остановится, его объект автоматически удалится.

|  |
| --- |
| class Server : public std::enable\_shared\_from\_this<Server> {  public:  explicit Server(net::io\_context& ioc, const tcp::endpoint& endpoint, std::string chats\_path)  : ioc\_(ioc)  , acceptor\_(net::make\_strand(ioc))  , chat\_manager\_(ioc, chats\_path)  {  acceptor\_.open(endpoint.protocol());  acceptor\_.set\_option(net::socket\_base::reuse\_address(true));  // Привязываем acceptor к адресу и порту endpoint  acceptor\_.bind(endpoint);  acceptor\_.listen(net::socket\_base::max\_listen\_connections);  }  void Run();  private:  net::io\_context& ioc\_;  tcp::acceptor acceptor\_;  ChatManager chat\_manager\_;  std::shared\_ptr<Server> GetSharedThis();  void DoAccept();  void OnAccept(sys::error\_code ec, tcp::socket socket);  void AsyncRunSession(tcp::socket&& socket);  }; |

Метод Run запускает работу сервера, он начинает принимать соединения, впервые запуская метод DoAccept.

Метод DoAccept вызывается для приема нового соединения. Происходит его асинхронное ожидание, к которому привязывается обработчиком OnAccept.

Метод OnAccept вызывается предыдущим методом, когда соединение произошло, чтобы запустить для него отдельную сессию и переключить сервер к ожиданию нового соединения вызовом метода DoAccept..

Метод AsyncRunSession вызывается методом OnAccept чтобы запустить отдельную корутину сессии от нового соединения. Она стартует в режиме detached, т.е. не имеет обработчика при остановке работы и идет отдельным потоком действий в контексте асинхронности.

Полученная система позволяет динамически настраивать расходуемые на работу ресурсы при сохранении производительности. Классовая структура многократно упрощает читаемость и тестируемость кода.

Для еще более глубокого понимания принципов работы системы следует проиллюстрировать их диаграммой классов. Использование структурной схемы алгоритма нецелесообразно, т.к. процессы происходят асинхронно, вызываясь по отдельности объектами классов.

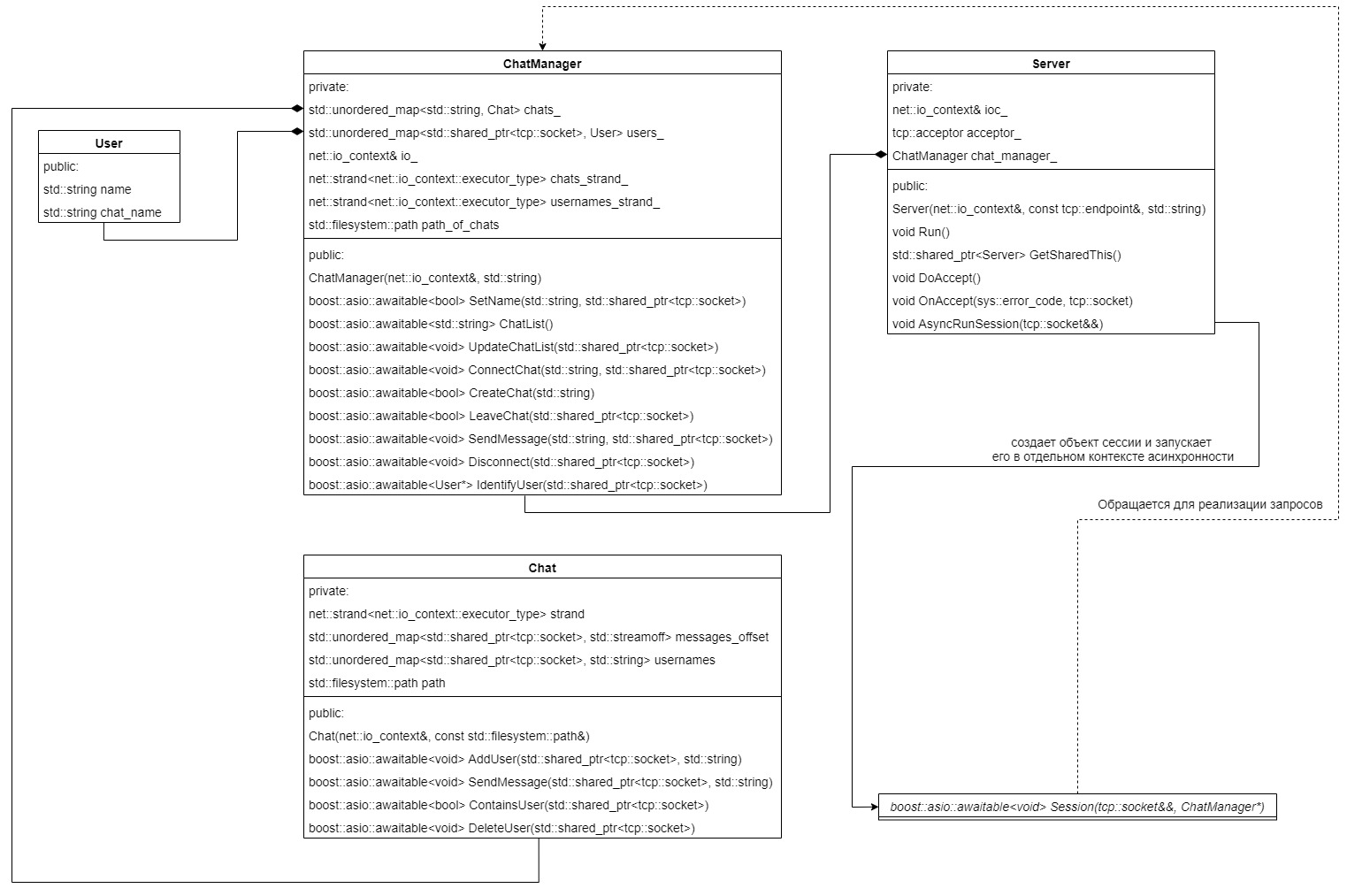


Рисунок 16 - Диаграмма классов многопоточно-асинхронного сервера

# Тестирование программного продукта

В данном разделе будут проанализированы работоспособность тестирующей системы и производительность разбираемых архитектур серверов.

При иллюстрации данных тестирования, графическое отображение результатов может быть крайне полезным инструментов для

## Тестирование функционирования системы

Т.к. система в своей основе базируется на использовании докера, для иллюстрирования корректной работы в первую очередь следует пояснить скрипты запуска и показать список запущенных контейнеров.

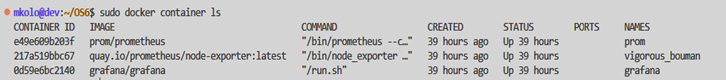


Рисунок 17 – Скрин необходимых запущенных контейнеров

Т.к. мы тестируем сетевое приложение, генератор нагрузки будет запускаться на отдельном устройстве.

Эти три ключевых контейнера обеспечивают работоспособность нашей системы.

Контейнер prom/prometheus отвечает за:

* сбор метрик – сервис собирает из с различных источников – в нашем случае с приложений, запущенных в своих контейнерах. В дальнейшем эти метрики будет использовать уже Grafana
* хранение данных – Prometheus хранит собранные метрики в своей временной базе данных
* создание возможности обработки данных при помощи специального языка для запроса и анализа метрик – PromQL

Контейнер quay.io/prometheus/node-exporter отвечает за запуск важнейшего инструмента мониторинга – Node-exporter. Он предоставляет метрики об использовании как системных ресурсов, так и производительности операционной системы Linux. Он устанавливается на каждом хосте, который мониторится Prometheus.

Последний из трех, но такой же важный, как и остальные, контейнер – grafana/grafana. Этот популярный в индустрии инструмент визуализации и анализа данных используется нами в целях визуального мониторинга производительности сложных систем. В нашей системе он подключается к Prometheus как к источнику данных, после чего необходимые данные передаются на гибкую систему дашбордов - наборов графиков и других средств визуализации данных.

Запуск Prometheus происходит по команде

|  |
| --- |
| sudo docker run -d --name prom --network="host" prom/prometheus |

Докеру передаются параметры:

1. -d : “detached” режим – поток вывода сервера не направлен в консоль, он работает отдельно от нее
2. –network=”host” : сети контейнера будут доступны все порты устройства

Далее для корректной работы сервиса ему нужно передать следующий файл настроек:

|  |
| --- |
| global:  scrape\_interval: 1m  evaluation\_interval: 5s  scrape\_configs:  - job\_name: 'node'  static\_configs:  - targets: [ 'localhost:9100' ]  - job\_name: 'server'  static\_configs:  - targets: [ 'localhost:9200' ] |

В нем указывается, что данные, анализируемые Prometheus, будут собираться относительно периоду 1 минуты, максимальная задержка считывания параметра – 5 секунд.

Также настраиваются источники данных – данные от Node-exporter читаются с порта 9100, а работа сервера ожидается на порте 9200.

Для того, чтобы запущенный контейнер учел параметры из этого файла, необходимо скопировать его в контейнер и перезапустить его:

|  |
| --- |
| sudo docker cp prometheus.yml prom:/etc/prometheus  sudo docker restart prom |

Докеру передаются параметры для запуска Node-exporter:

|  |
| --- |
| sudo docker run -d --name prom --network="host" prom/node-exporter |

Они аналогичны предыдущему запуску. После запуска все необходимые системные данные будут передаваться уже запущенному Prometheus.

Таким же образом запускаем Grafana:

|  |
| --- |
| sudo docker run -d --name prom --network="host" grafana/grafana |

Результатом запуска всех трех сервисов получим возможность просматривать по ip сервера данные с порта 9090 для Prometheus и с порта 3000 для Grafana.

Далее следует собрать на двух различных устройствах образы серверов и тестировщика соответственно, чтобы потом на них запустить необходимые контейнеры.

Dockerfile для многопоточного и асинхронного серверов будет отличаться лишь набором файлов для сборки, поэтому проиллюстрируем лишь тот, что отвечает за многопоточный вариант.

|  |
| --- |
| # Создать образ на основе базового слоя gcc (там будет ОС и сам компилятор).  # 11.3 — используемая версия gcc.  # Не просто создаём образ, но даём ему имя build  FROM gcc:11.3 as build  # Выполнить установку зависимостей внутри контейнера.  RUN apt update && \  apt install -y \  python3-pip \  cmake \  && \  pip3 install conan==1.59  # Копируем conanfile.txt в контейнер и запускаем conan install  COPY conanfile.txt /app/  RUN mkdir /app/build && cd /app/build && \  conan install .. -s compiler.libcxx=libstdc++11 -s build\_type=Debug --build=missing  # Только после этого копируем остальные исходники  COPY server.cpp \  logger.h logger.cpp \  server\_interface.h server\_interface.cpp \  server\_routines.h server\_routines.cpp\  CMakeLists.txt \  /app/  RUN cd /app/build && \  cmake .. -DCMAKE\_BUILD\_TYPE=Debug && \  cmake --build .  # Второй контейнер в том же докерфайле  FROM ubuntu:22.04 as run  # # Создадим пользователя usr  # RUN groupadd -r usr && useradd -r -g usr usr  # USER usr  # Скопируем приложение со сборочного контейнера в директорию /app.  COPY --from=build /app/build/bin/server /app/  # Запускаем сервер  ENTRYPOINT ["/app/server"] |

После сборки образа, необходимо запустить контейнер со следующими параметрами:

|  |
| --- |
| sudo docker build -t server .  sudo docker run --rm -p 5000:5000 server |

Параметры запуска:

* --rm : контейнер удаляется после остановки работы
* -p 5000:5000 : 5000 порт контейнера соединяется с портом 5000 на устройстве

Dockerfile для сборки образа тестировщика приведен частично, т.к его build часть отличается от предыдущего примера лишь файлами сборки:

|  |
| --- |
| # Второй контейнер в том же докерфайле  FROM ubuntu:22.04 as run  # Создадим пользователя usr  RUN groupadd -r usr && useradd -r -g usr usr  USER usr  # Скопируем приложение со сборочного контейнера в директорию /app.  COPY --from=build /app/build/bin/tester /app/  # Запускаем тестировщик  ENTRYPOINT ["/app/tester", "10"] |

Аналогично собираем и запускаем тестировщика:

|  |
| --- |
| sudo docker build -t tester .  sudo docker run ---rm --network=host tester | python3 script.py |

Запуск тестировщика отличается тем, что он идет не в режиме detached и в том, что вывод программы будет передаваться в запущенный python скрипт для отправки логов на Prometheus.

В нем будут созданы 4 метрики, которые после будут обработаны:

количество правильных и неправильных переданных JSON, количество ошибок чтения (будут индикаторами окончания сессии по причине ошибки соединения) и задержка ответа на запрос.

Далее будут приведены части скрипта, отвечающие за запуск сервера, передающего данные логов, задание метрик и передачу метрики

|  |
| --- |
| import prometheus\_client as prom  …  prom.start\_http\_server(9200) |

|  |
| --- |
| # Создаём метрики  good\_lines = prom.Counter('webexporter\_good\_lines', 'Good JSON records')  wrong\_lines = prom.Counter('webexporter\_wrong\_lines', 'Wrong JSON records')  read\_errors = prom.Counter('client\_read\_errors', 'Client read errors')  response\_time = prom.Histogram('webserver\_request\_duration', 'Response time', labelnames=['client\_request\_e', 'server\_response\_e'],  buckets=(.001, .002, .005, .010, .020, .050, .100, .200, .500, float("inf"))) |

|  |
| --- |
| # Регистрируем запрос  if data["message"] == "Read first response from server" or data["message"] == "Wrote first response":  total\_time\_seconds = data["data"]["response time"] / 1000 response\_time.labels(client\_request\_e = data["data"]["response"]["client\_request\_e"],  server\_response\_e =  data["data"]["response"]["server\_response\_e"]).observe(total\_time\_seconds) |

Остается лишь подключить к Grafana источник данных Prometheus и приступить к анализу получаемых данных при ее помощи.

## Анализ результатов

В этом разделе мы проанализируем такие показатели, как использование CPU в процентах, использование RAM в процентах и задержку ответа.

Метрики использования CPU и RAM - ключевые при измерении производительности сервера в силу того, что они влияют на скорость работы и доступность ресурсов для обработки запросов, посылаемых клиентом на сторону сервера. Эти метрики помогают определить, насколько эффективно он использует свои ресурсы и готов ли обрабатывать большое количество запросов без ущерба для качества обслуживания пользователей. Поэтому, для поддержания высокой производительности сервера, важно следить за этими метриками и управлять ресурсами сервера соответственно.

Для тестирования производительности серверов будет использоваться виртуальная машина на ОС Linux Ubuntu со следующими параметрами:

* 4 процессорных ядра Intel Ice Lake
* 12 Гб RAM
* 60 Гб дискового пространства

Такой мощности вполне достаточно для стабильной работы сервера при достаточно высоких нагрузках. Дисковое пространство необходимо для хранения образов запускаемых контейнеров и дальнейшего хранения файлов чатов в уже работающих контейнерах.

С этого устройства при помощи node-exporter будем получать данные на Grafana. Данные использования CPU будут визуализированы классическими графиками по процентному соотношению от 0% до 100%.

Использование RAM лучше не иллюстрировать при помощи численного варианта – график будет показывать процентное значение используемой в данный момент памяти, исходя из ресурсов системы.

Стоит отдельно отметить, что задержка ответа будет измеряться на стороне приложения, генерирующего нагрузку, т.к. запросы в многопоточном сервере опираются на низкоуровневую логику Linux и невозможно будет корректно измерять задержку от запроса до ответа без существенного влияния на производительность: синхронизация, требующаяся для постоянного отслеживания приходящих запросов, вызовет захват ресурсов от их чтения, что конечно же будет отрицательно влиять на адекватность собираемых данных.

Для иллюстративности, мы рассмотрим три случая для каждой реализации сервера:

1. 2 клиента для иллюстрации минимальной нагрузки
2. 10 клиентов для того, чтобы проиллюстрировать как небольшой рост нагрузки на сервер будет влиять на изучаемые параметры
3. 100 клиентов на каждый сервер как максимальная измеряемая нагрузка

Отдельно отметим, что 100 клиентов при стандартных условиях – весьма малое число, однако тестирующее приложение способно нагружать сервер многократно большим количеством запросов от каждого клиента: реальный пользователь никогда не достигнет такого количества отправляемых сообщений и переходов между чатами как минимум в силу того, что генератор нагрузки не пользуется графической частью, а обрабатывает данные напрямую.

### Тестирование многопоточного сервера

**2 клиента**

Отобразим динамику расхода ресурсов:

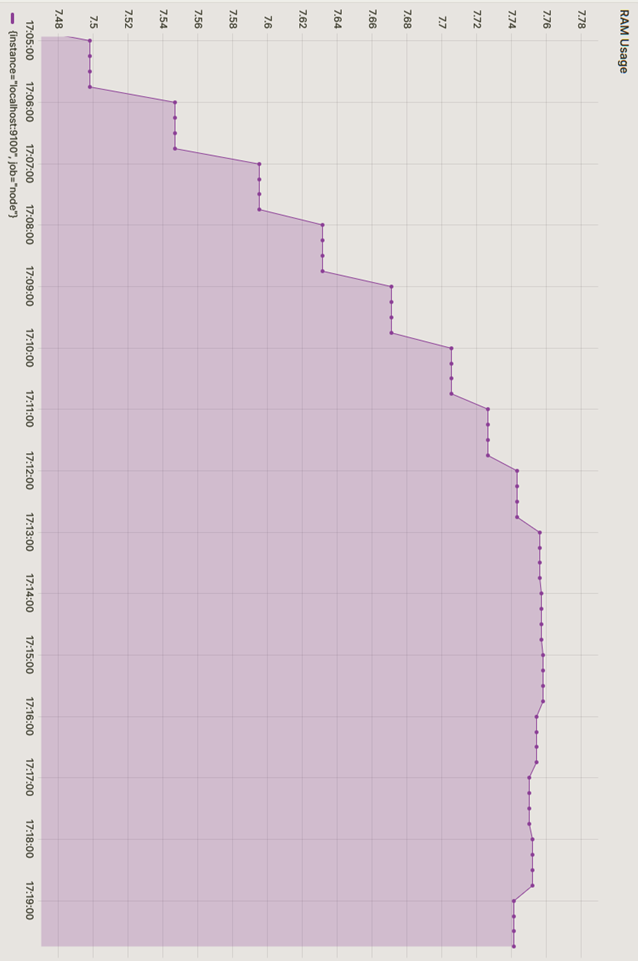


Рисунок 18 – График использования ОП на устройстве сервера при 2 клиентах для многопоточного сервера

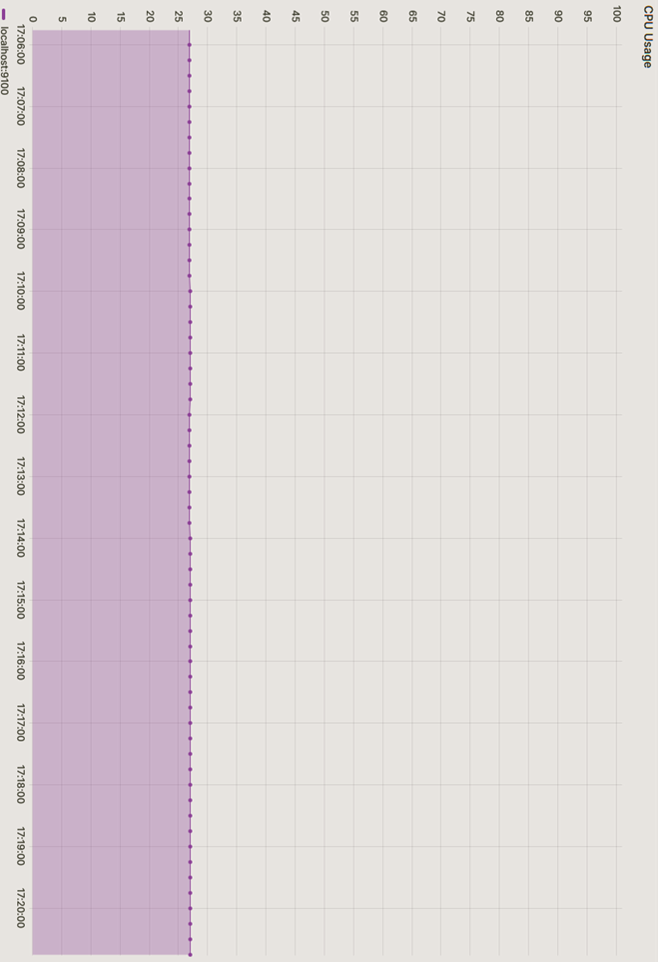


Рисунок 19 – График использования ресурсов процессора на устройстве сервера при 2 клиентах для многопоточного сервера

Как видно на графиках, захватив необходимое количество операционной памяти, приложение нашло баланс в ее использовании.

Колебания в использовании CPU минимальны, это объяснимо с учетом архитектуры сервера: выделяются четыре потока обработки, дополнительные формируются исключительно для использования чатов при необходимости. Т.к. количество клиентов минимально, то не стоит ожидать сильного использования ресурсов.

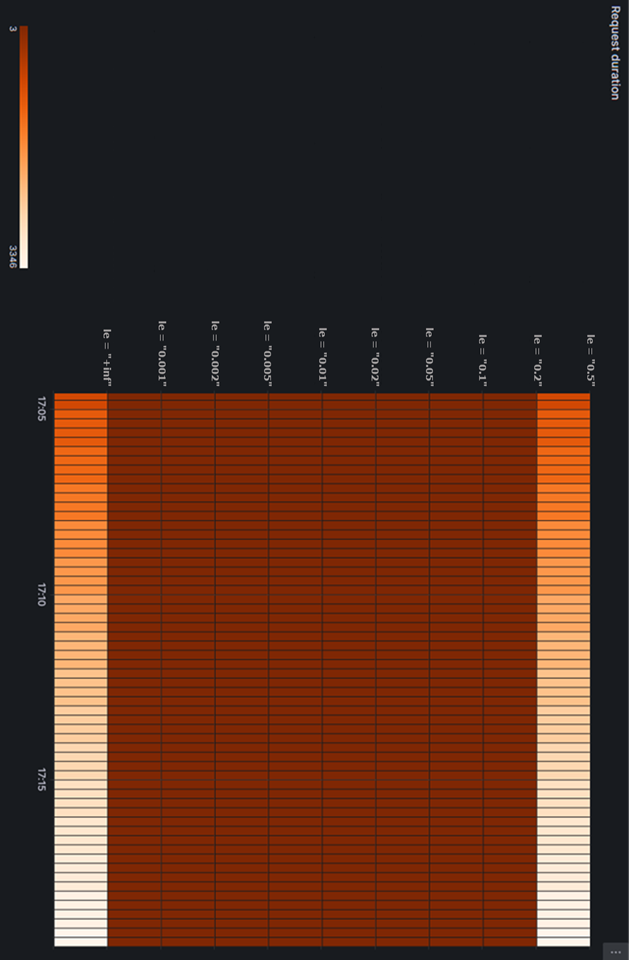


Рисунок 20 – Тепловая карта задержек ответов на запросы клиентов при 2 клиентах для многопоточного сервера

Задержка нарастает со временем использования сервера, т.к. ускоренные тестирующим приложением клиенты имеют возможность перескакивать из чата в чат, каждый раз получая все его содержимое. Отдельно стоит отметить, что распределение основных задержек между 0.5 и +inf объясняется, тем, что т.к. ожидание сообщений замораживает поток, выполняющий связанную задачу.

**10 клиентов**

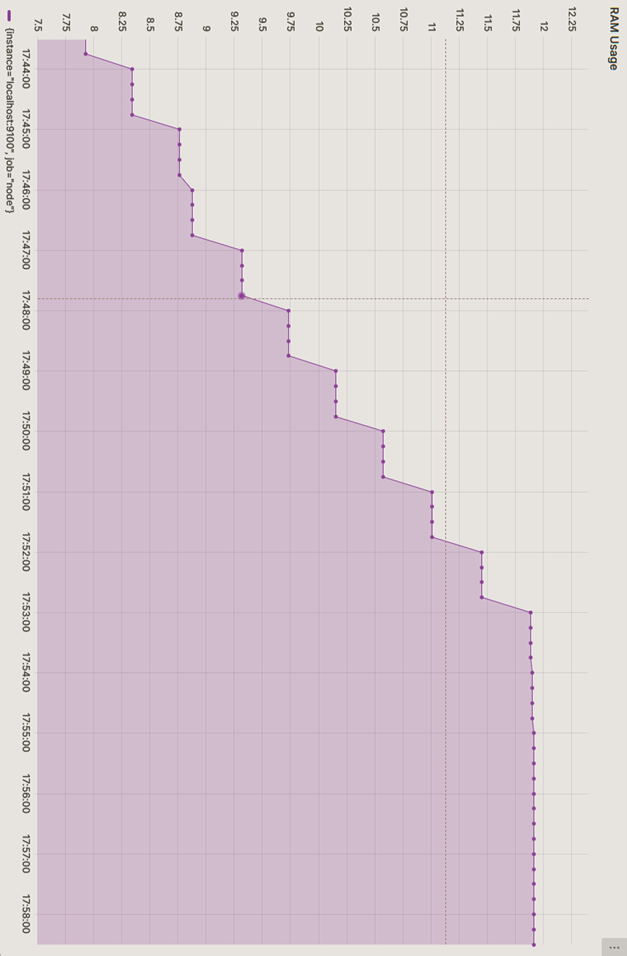


Рисунок 21 – График использования ОП на устройстве сервера при 10 клиентах для многопоточного сервера

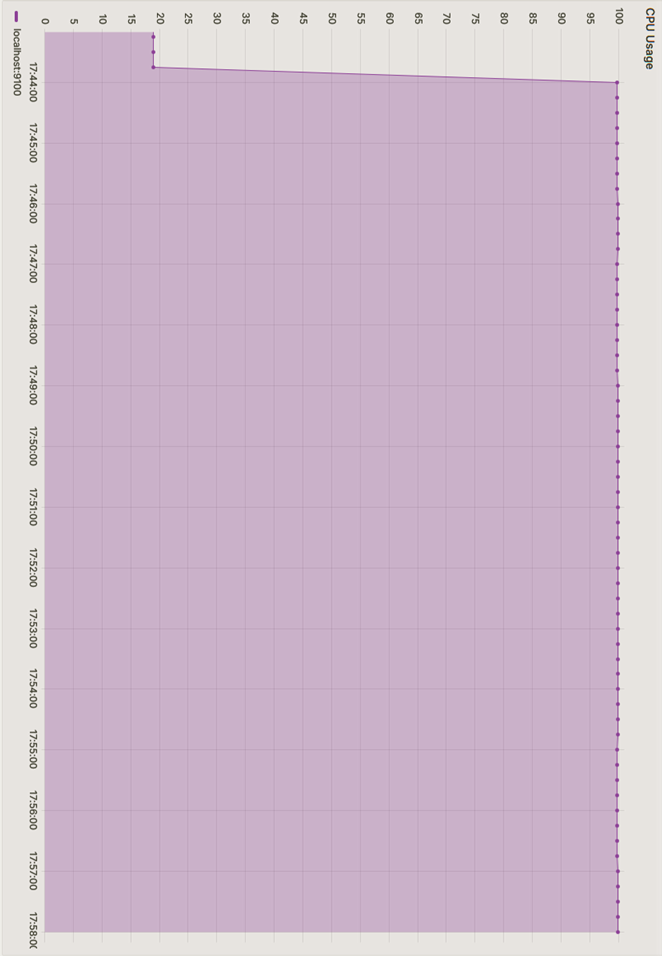


Рисунок 22 - График использования ресурсов процессора на устройстве сервера при 10 клиентах для многопоточного сервера

Новые чаты, опирающиеся на новые потоки, занимают все возможные ресурсы процессора, что, без всякого сомнения, крайне негативно отражается на работе сервера. Использование памяти остановилось в росте, набран необходимый для работы объем.

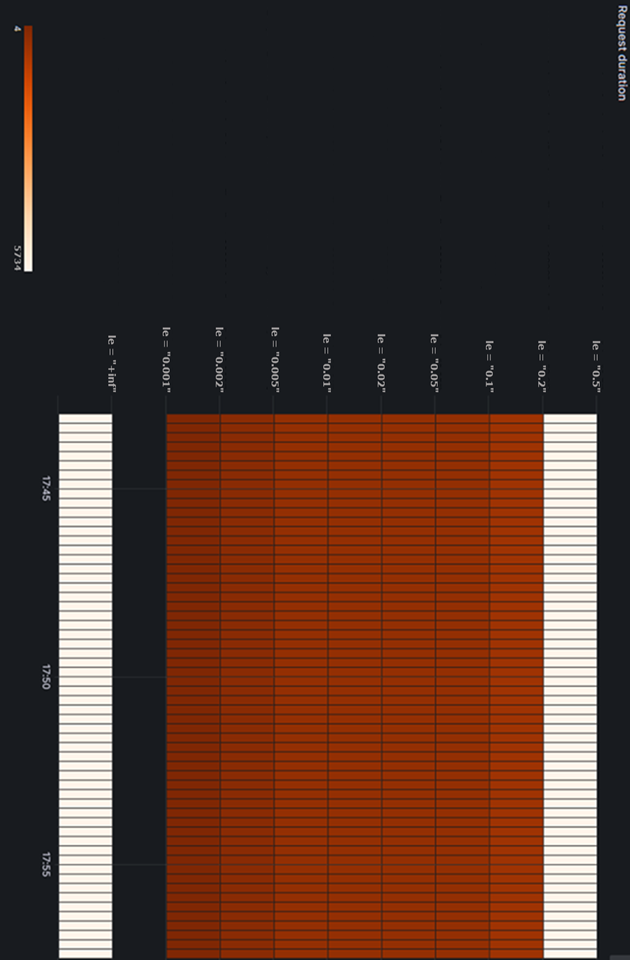


Рисунок 23 - Тепловая карта задержек ответов на запросы клиентов при 2 клиентах для многопоточного сервера

Здесь заметно, что при большем количестве клиентов резко растет задержка на ответ – снова дает знать о себе остановка потока в ожидании сообщения.

**100 клиентов**

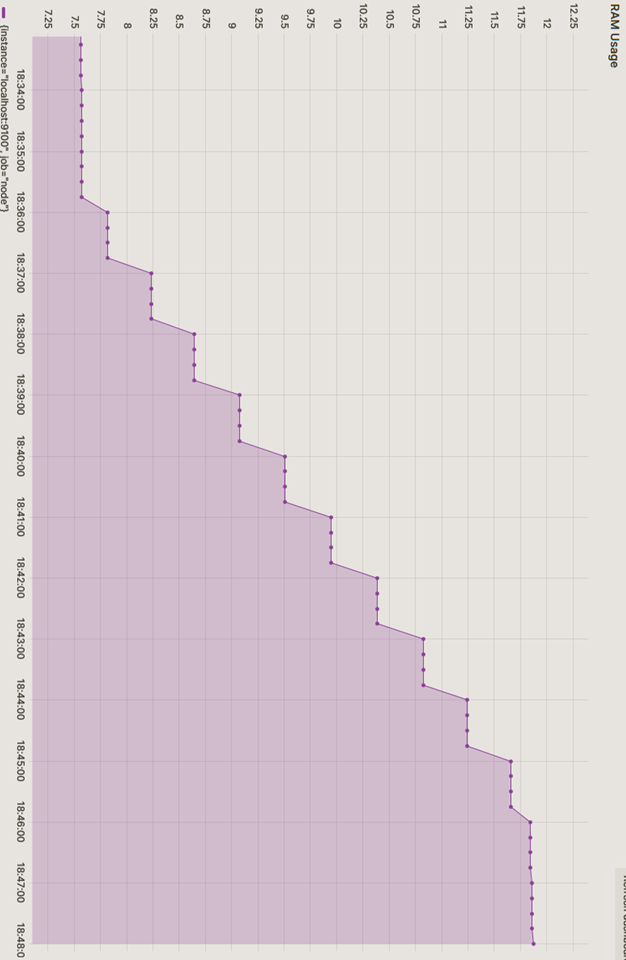


Рисунок 24 - График использования ОП на устройстве сервера при 100 клиентах для многопоточного сервера

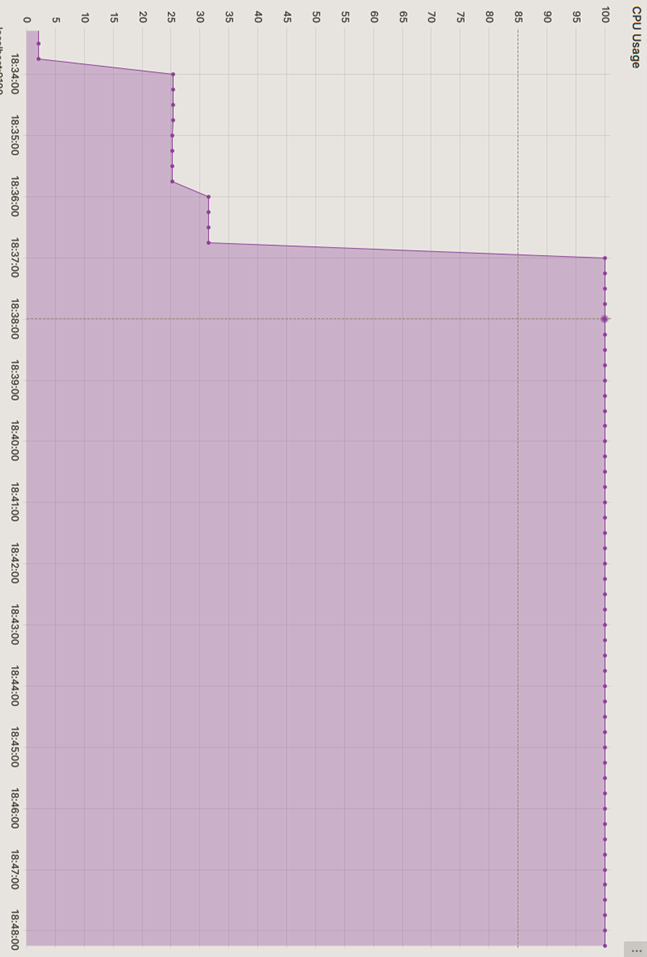


Рисунок 25 - График использования ресурсов процессора на устройстве сервера при 100 клиентах для многопоточного сервера

Как и в предыдущем тесте, захватив необходимое количество оперативной памяти, сервер уперся в потолок в ресурсах процессора и продолжил свою работу, снизив эффективность обработки.

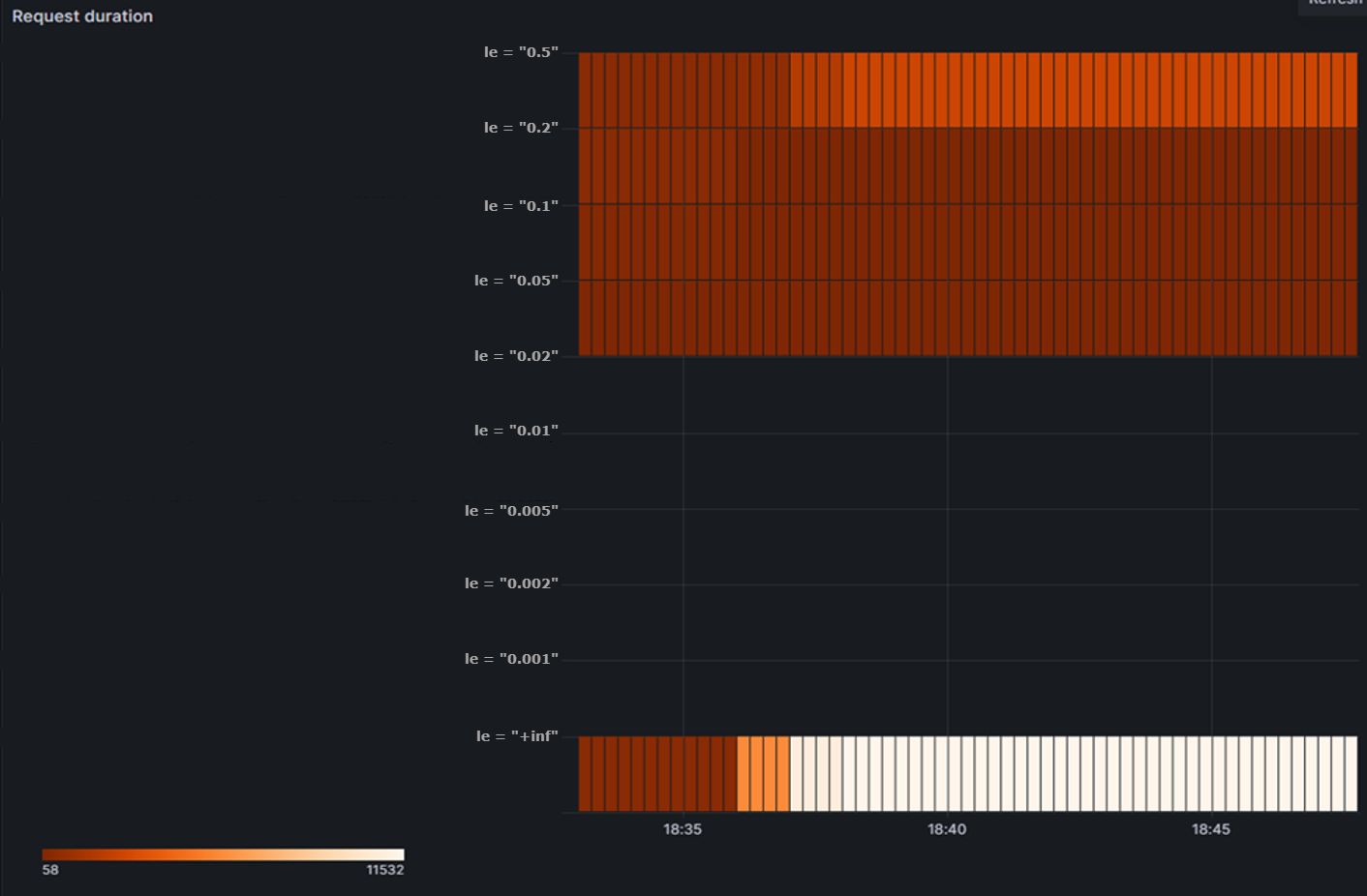


Рисунок 26 - Тепловая карта задержек ответов на запросы клиентов при 100 клиентах для многопоточного сервера

Приведенный график еще ярче показывает, что не асинхронный подход имеет яркие недостатки при работе с веб сервером. Доминирующее число ответов сервера занимали более полсекунды, что неприемлемо для приложения, рассчитанного на высокую нагрузку.

**Итоги тестирования многопоточного сервера**

Полученные данные о работе сервера, основанного на многопоточной конвейерной архитектуре, ярко иллюстрируют сложившуюся для него ситуацию:

* Сервер, базирующийся на выбранной в данном разделе логике, главным своим достоинством имеет малое использование оперативной памяти
* Опора исключительно на дополнительные потоки для каждой рутины сервера приводит к нерациональному расходу ресурсов процессора создаваемыми потоками. Ключевое влияние оказывают потоки чатов.
* Задержка, с которой сталкивается клиент, растет крайне быстро и в таком же темпе достигает неприемлемых значений

### Тестирование многопоточно-асинхронного сервера

**2 клиента**

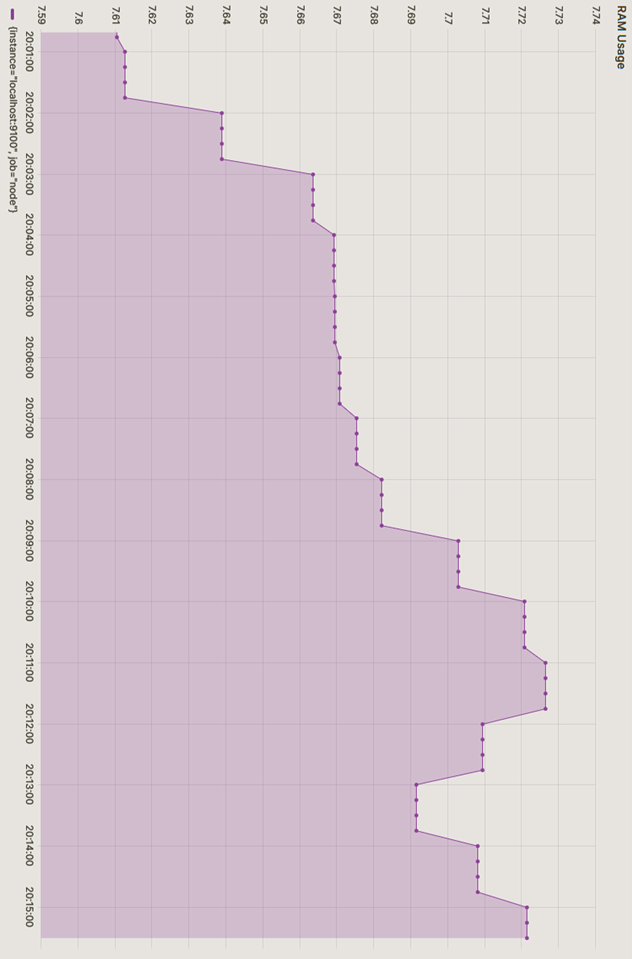


Рисунок 27 - График использования ОП на устройстве сервера при 2 клиентах для многопоточно-асинхронного сервера

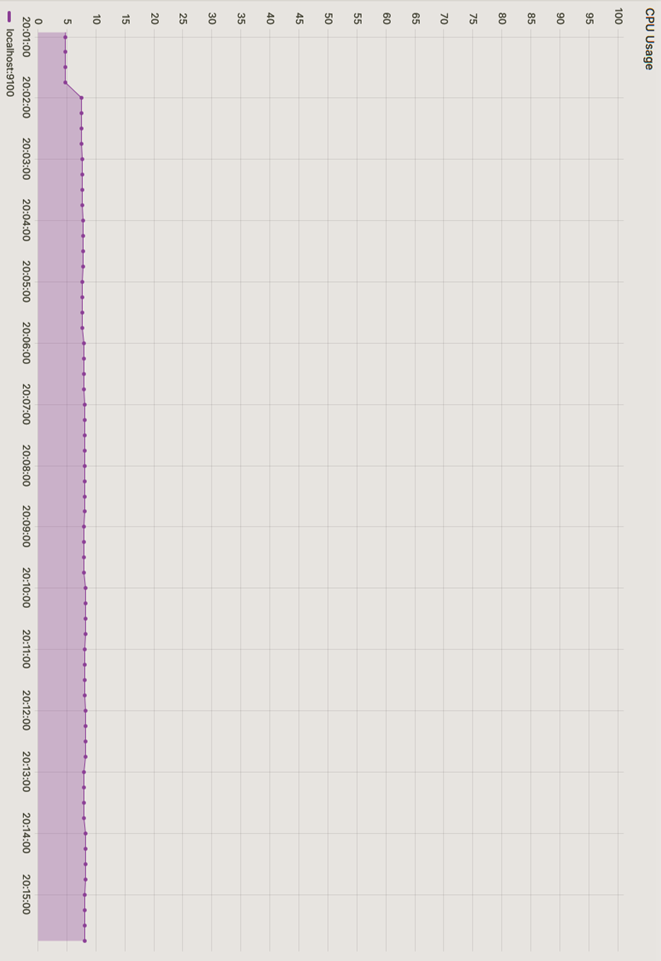


Рисунок 28 - График использования ресурсов процессора на устройстве сервера при 2 клиентах для многопоточно-асинхронного сервера

Скачки в объеме используемой памяти объясняются созданием и деструкцией промежуточных объектов корутин – приложение полностью опирается на ООП. В дальнейшем динамика сохранится.

Уже на этом этапе стоит отметить крайне рациональное использование ресурсов процессора. Эта особенность проявляется в силу использования пула потоков.

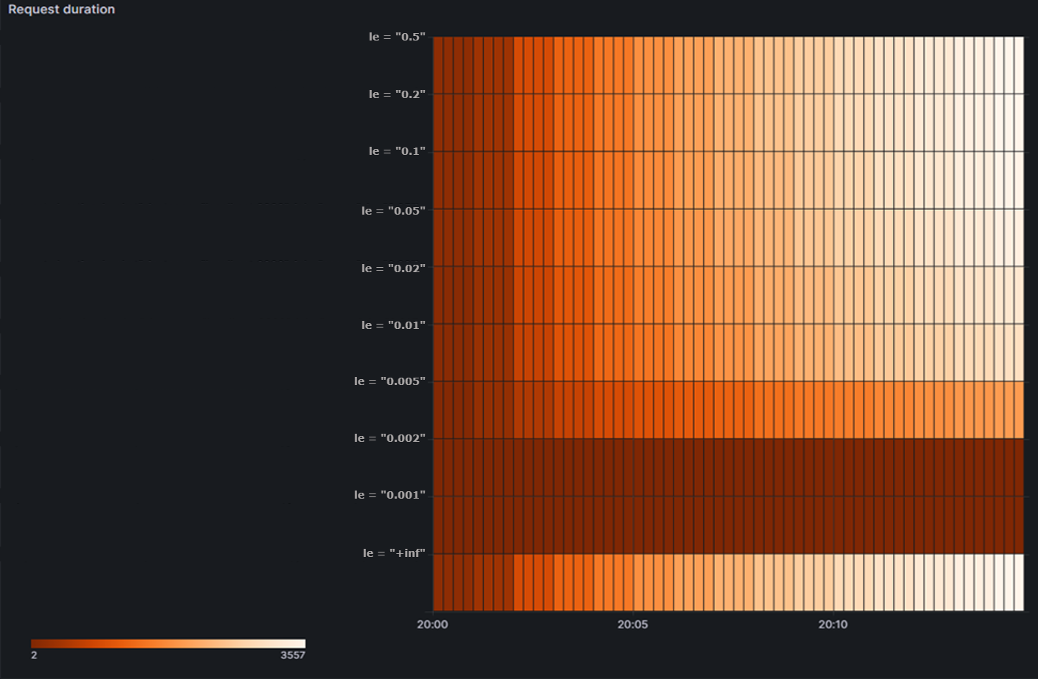


Рисунок 29 - Тепловая карта задержек ответов на запросы клиентов при 2 клиентах для многопоточно-асинхронного сервера

Задержки до получения ответа клиентами имеют куда более распределенный характер. Это вызвано тем, что клиентские запросы вызывают различные процедуры, которые происходят асинхронно, с минимизированным ожиданием друг друга, поэтому куда ярче заметно распределение задержек для запросов в разных контекстах работы сервера.

**10 клиентов**

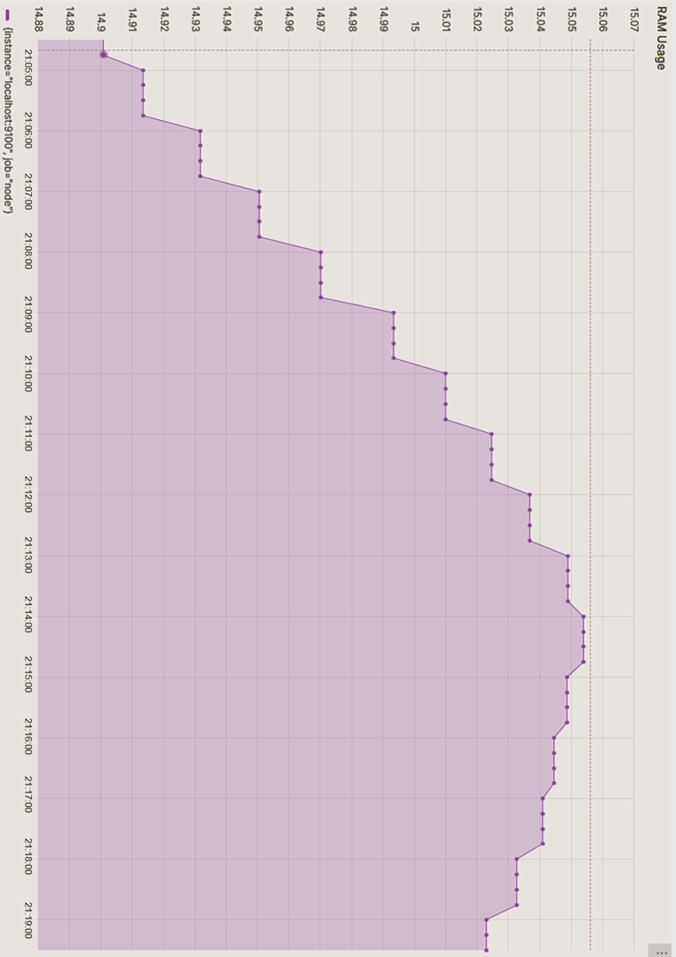


Рисунок 30 - График использования ОП на устройстве сервера при 10 клиентах для многопоточно-асинхронного сервера

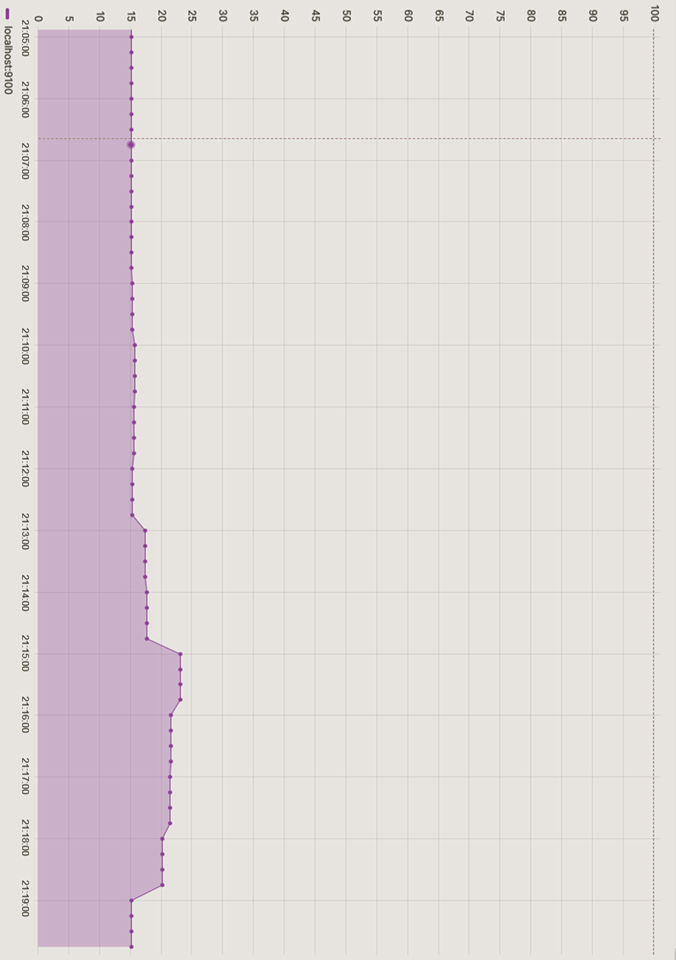


Рисунок 31 – График использования ресурсов процессора на устройстве сервера при 10 клиентах для многопоточно-асинхронного сервера

Динамика, отмеченная в предыдущем тесте, повторяется с учетом обновленных условий. Наблюдаются скачки в использовании ресурсов памяти, а возросшие затраты ресурсов процессора все еще можно назвать крайне эффективными в использовании.

Смена динамики возрастания используемой ОП на уменьшение могла возникнуть в силу постепенного избавления от временных объектов многочисленно вызванных корутин.

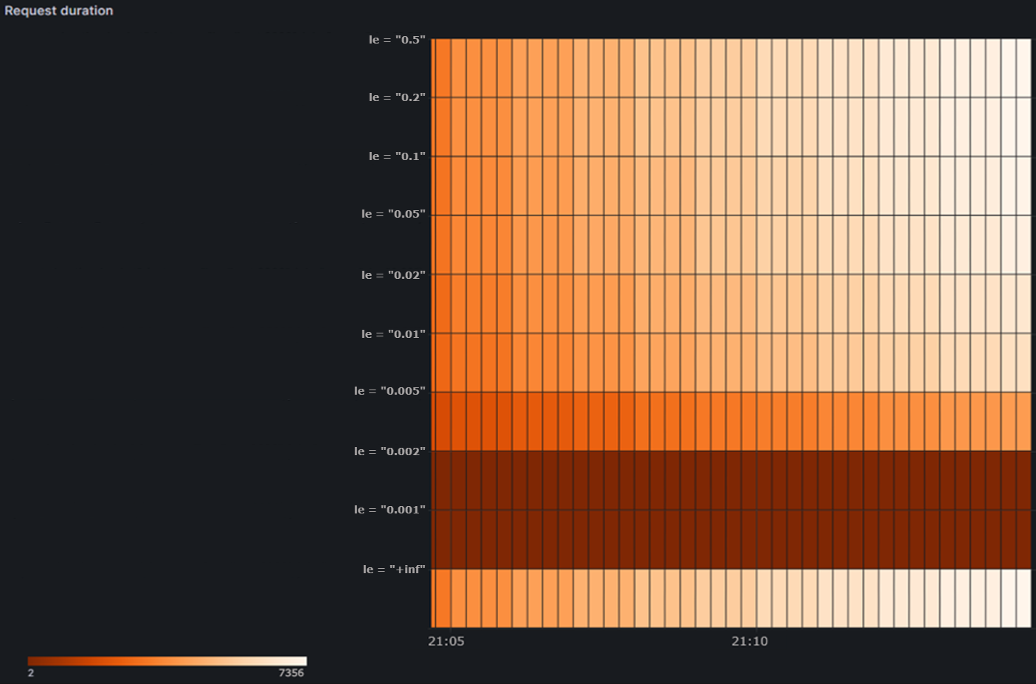
****

Рисунок 32 – Тепловая карта задержек ответов на запросы клиентов при 10 клиентах для многопоточно-асинхронного сервера

Тепловая карта показывает, что прирост задержек увеличил скорость, что объясняется возросшим количеством пользователей и возросшим вследствие количества сообщений в каждом чате.

Несколько ярче стало смещение в сторону более высоких задержек, однако они все еще остаются вполне адекватными.

**100 клиентов**

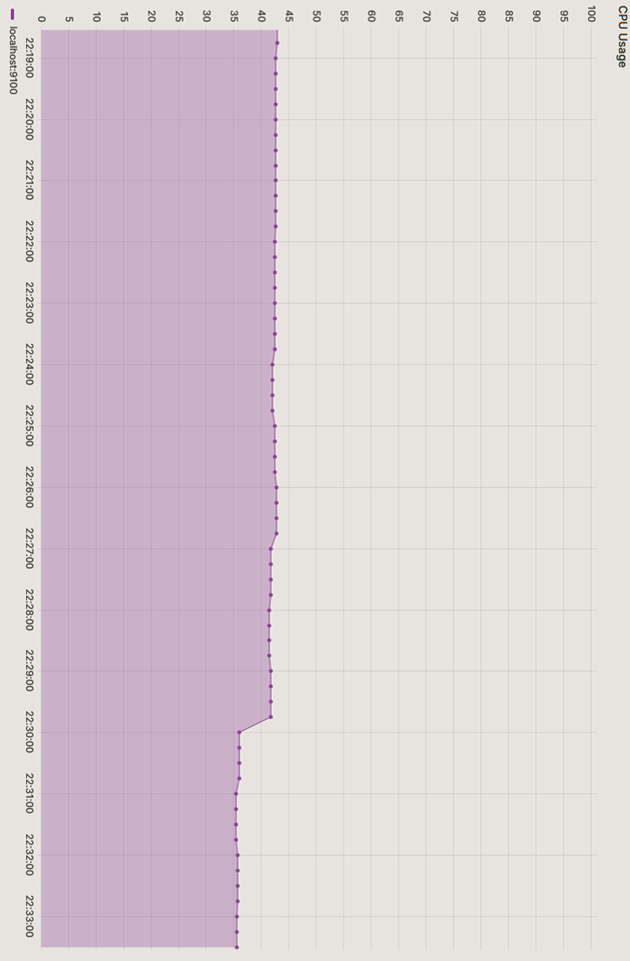


Рисунок 33 - График использования ОП на устройстве сервера при 100 клиентах для многопоточно-асинхронного сервера

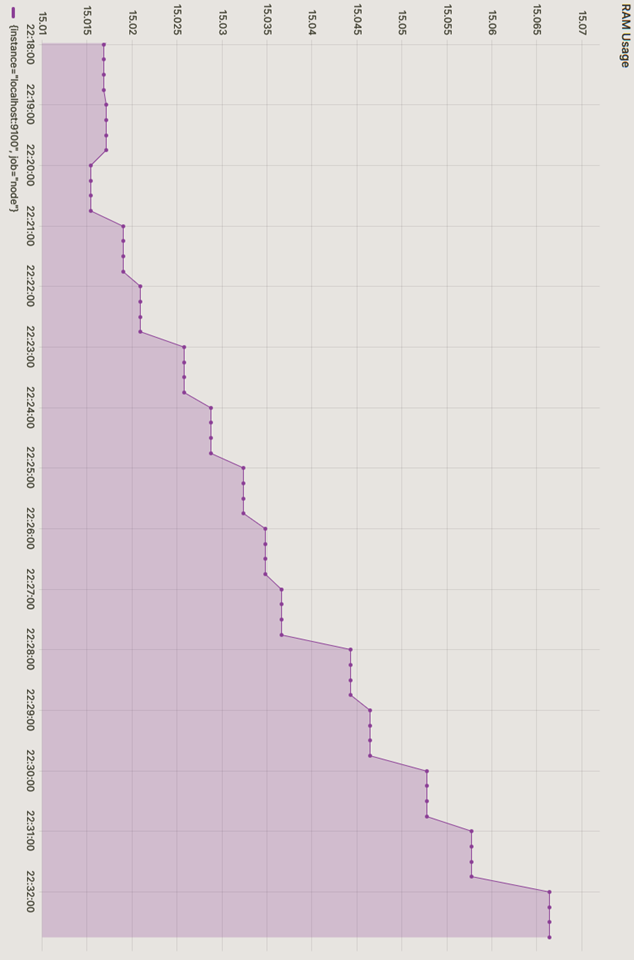


Рисунок 34 - График использования ресурсов процессора на устройстве сервера при 100 клиентах для многопоточно-асинхронного сервера

На самом сложном этапе тестирования подскочило использование памяти – множество запущенных асинхронно корутин скачками захватывали ресурсы памяти, найдя баланс лишь в самом конце периода измерения.

Объем памяти, используемый приложением, все еще можно считать весьма малым.

Ресурсы процессора, будучи безусловно более загруженными относительно предыдущих кейсов тестирования, все еще имеют запас для роста расходов.

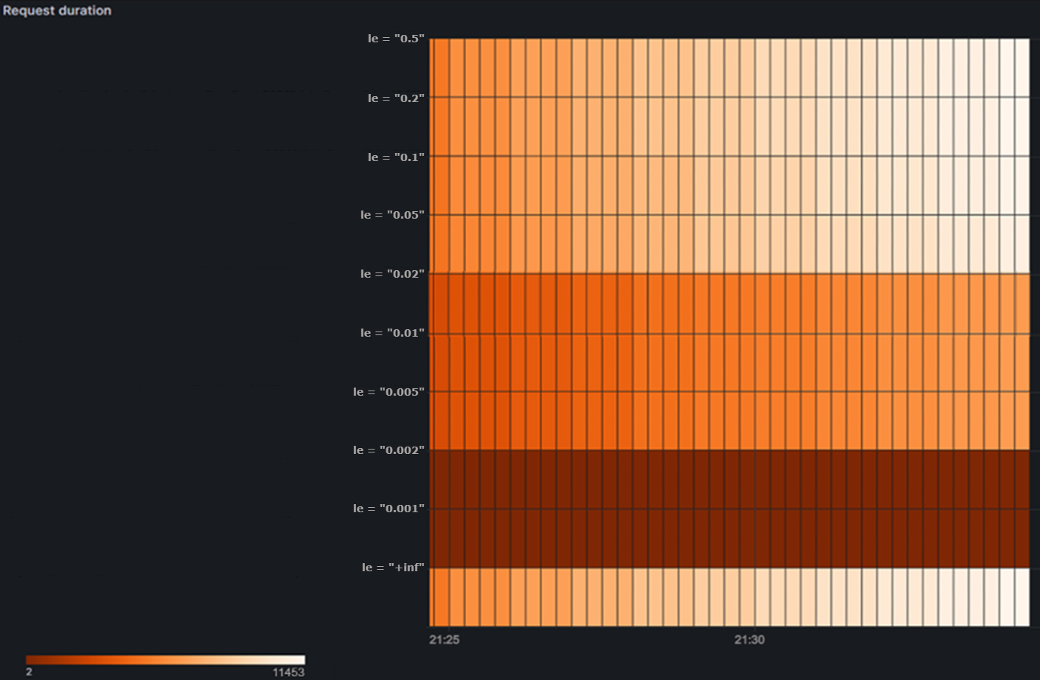
****

Рисунок 35 - Тепловая карта задержек ответов на запросы клиентов при 100 клиентах для многопоточно-асинхронного сервера

В данном случае тестирования максимальное выбранное количество пользователей хорошо проявило способности сервера выдерживать возрастающие нагрузки. Общая доля запросов, попадающих во временной отрезок, необходимый для ответа, больший полсекунды, все еще не является даже близко доминирующей, хотя и есть проседание доли малых задержек.

**Итоги тестирования многопоточно-асинхронного сервера**

Анализ полученных в результате тестов данных показал следующие моменты:

* ООП, используемое в данной реализации сервера в куда более полной форме, имеет яркое влияние на его работу: хотя использование памяти и возросло, приспособление методов классов к асинхронности через превращение их в корутины дало большой скачок в производительности.
* Задержки, с которыми сталкивались клиенты, были вполне удовлетворительными
* Использование асинхронности предотвратило создание гигантских очередей, уничтожающих шансы на эффективную работу системы

### Общие итоги тестирования, сравнение результатов

В сравнении столь отличных реализаций серверов стоит структурировать параметры, относительных которых оно будет происходить.

Использование ресурсов памяти:

Многопоточный сервер имеет более скромные потребности в оперативной памяти. К тому же характер ее использования не имеет склонности к резким перепадам, в отличие от многопоточно-асинхронной архитектуры, чья полная опора на ООП создает необходимость в создании временных объектов классов, которые являются основой логики работы приложения.

Использование ресурсов процессора:

Явное преимущество в эффективности использования вычислительных ресурсов процессора достигается многопоточно-асинхронной реализацией за счет нескольких ключевых факторов.

1. Использование пула потоков вместо динамического создания при необходимости ограничивает рост используемых ресурсов процессора.
2. Опора на асинхронный движок BOOST позволяет создавать отдельные очереди для синхронизации, при этом сохраняя возможность автоматического асинхронного многопоточного выполнения задач.

При готовности и отсутствии более раннего претендента на выделенные ресурсы задача захватит контроль над вычислительными ресурсами потока из зафиксированного пула. Это дает возможность реализации максимально эффективных процессов чтения и обработки, минимизирующих ожидание.

1. Более конкретные и упрощенные сценарии синхронизации в ООП дают возможность лучше оптимизировать процессы системы

Задержки:

В первую очередь, в силу ранее отмеченных преимуществ, многопоточно-асинхронная реализация серверной логики является очевидным победителем по этому параметру.

Однако стоит подметить, что задержки в своем многообразии имеют разные характеры распределения.

Т.к. многопоточный сервер сталкивается с формированием очередей на чтение с полной заморозкой потоков, то его периоды задержки ярко смещены в сторону длинных отрезков времени и чем выше нагрузка, тем все большая доля попадает к запросам, ожидавшимся более чем 0.5 секунды.

В свою очередь, многопоточно-асинхронный сервер имеет куда более красочный и интересный характер распределения периодов ожидания, который заметно проявляется на тепловой карте задержек в тестах. Т.к. фактор заморозки потоков минимизирован до того, что его не обязательно учитывать, куда большее влияние имеет уже состояние системы, на которой основан сервер. За счет этого представлены куда более разнообразные варианты распределений, которые, к тому же, будут незначительно отличаться в зависимости от того, как будут поданы задачи на сервер.

Финализируя проведенный анализ, более современный подход многопоточного сервера обеспечил ему победу в сравнении итоговой производительности, однако внедрение асинхронности потенциально будет тем, что позволит многопоточному серверу, базирующемуся на конвейерной логике, догнать своего оппонента.

Заключение

В данной выпускной квалификационной работе приведены и проиллюстрированы этапы анализа исходных данных и проектирования, разработки программного обеспечения, тестирования двух архитектур серверов и работоспособности тестирующей системы в целом и анализа результатов тестирования в форме графического отображения.

Глава анализа исходных данных и проектирования структурировал и описал этапы конструирования архитектур серверов, особенностей получаемых и отправляемых серверами данных.

Глава разработки реализаций сервера и клиента показала основную часть работы – непосредственное создание программного продукта в форме двух серверов и клиентского приложения. Также было разработано приложение для имитации нагрузки от выбранного числа клиентов, которое будет использовано для тестирования в дальнейшем.

Глава тестирования разделилась на логические три части – тестирование работоспособности системы, нагрузочное тестирование серверов и анализ результатов. Первая часть заключалась в настройке и запуске докером основных контейнеров, отвечающих за обработку передаваемых данных, которые будут использованы для анализа. Вторая отвечает за формирование первичных выводов на основе графической визуализации полученных во время тестирования данных. Последняя – за сравнительный анализ данных, полученных на всех этапах тестирования двух реализаций серверов.

В результате выполненной работы была создана программная платформа для сравнения производительности и анализа положительных и отрицательных характеристик многопоточной и многопоточно-асинхронной серверных архитектур относительно друг друга.

Используемые программные платформы и средства разработки подходят выбранным операционной системе и используемым в разработке языкам программирования.

Разработка серверов на двух различных архитектурах показала недостатки и преимущества подходов со стороны разработки:

**Для многопоточного сервера:**

**Достоинства**: простота проектирования, малое использование ОП, относительно простая отладка

**Недостатки**: низкая читаемость низкоуровневого кода, высокий расход ресурсов процессора, заморозка потоков при чтении запроса

**Возможности для улучшения**: внедрение асинхронности для каждого из подсервисов, использование корутин, внедрение кеширования сообщений на стороне клиента

**Для многопоточно-асинхронного сервера:**

**Достоинства**: хорошая читаемость опирающегося на ООП кода, низкий расход ресурсов процессора, простота настройки синхронизации за счет использования современной библиотеки BOOST

**Недостатки**: сложность отладки основанного в большой мере на библиотеке приложения, высокое использование ОП

**Возможности для улучшения**: внедрение кеширования на стороне клиента, оптимизация исключения клиентов из списка рассылки для чатов

# Список использованных источников

1. Документация библиотеки BOOST: <https://www.boost.org/doc/>
2. Документация QT: <https://doc.qt.io/>
3. John Torjo; Wisnu Anggoro Boost.Asio C++ Network Programming - Second Edition - Packt Publishing, 2015 – 200 с.
4. Разработка приложений на C++ с использованием Boost. Рецепты, упрощающие разработку вашего приложения / пер. с анг. Д. А. Беликова. – М.: ДМК Пресс, 2020. – 346 с.: ил.
5. Расширенное описание библиотек, входящих в BOOST: https://theboostcpplibraries.com/
6. Шлее М. Qt 5.10. Профессиональное программирование на С++. -СПб.: БХВ-Петербург, 2018. -1072 с.: ил.
7. Моуэт Э. Использование Docker / пер. с англ. А.В. Снастина; науч. Ред. А.А. Маркелов. – М.: ДМК Пресс, 2017. – 354 с.: ил.

# Приложение A. Файл main.cpp генератора нагрузки

|  |
| --- |
| #include <iostream>  #include <coroutine>  #include <memory>  #include <thread>  #include <sstream>  #include <iomanip>  #include <random>  #include <vector>  #include <chrono>  #include <boost/array.hpp>  #include <boost/scoped\_ptr.hpp>  #include <boost/asio/read.hpp>  #include <boost/asio/write.hpp>  #include <boost/asio/use\_awaitable.hpp>  #include <boost/asio/awaitable.hpp>  #include <boost/asio/experimental/as\_tuple.hpp>  #include <boost/asio/ip/tcp.hpp>  #include <boost/asio/signal\_set.hpp>  #include <boost/asio/io\_context.hpp>  #include <boost/asio/detached.hpp>  #include <boost/asio/co\_spawn.hpp>  #include <boost/asio/connect.hpp>  // #include <boost/json.hpp>  #include "logger.h"  namespace net = boost::asio;  using tcp = net::ip::tcp;  namespace sys = boost::system;  using namespace std::literals;  enum client\_request\_e {  c\_set\_name = 0,  c\_create\_chat,  c\_connect\_chat,  c\_send\_message,  c\_leave\_chat,  c\_disconnect,  c\_receive\_message,  c\_wrong\_request,  c\_get\_chats  };  enum server\_response\_e{  s\_success = 0,  s\_failure,  };  struct client\_data\_t{  client\_request\_e request;  std::string message\_text;  };  struct server\_data\_t{  client\_request\_e request;  server\_response\_e response;  std::string message\_text;  };  //get random element from container  template<typename Iter, typename RandomGenerator>  Iter select\_randomly(Iter start, Iter end, RandomGenerator& g) {  std::uniform\_int\_distribution<> dis(0, std::distance(start, end) - 1);  std::advance(start, dis(g));  return start;  }  template<typename Iter>  Iter select\_randomly(Iter start, Iter end) {  static std::random\_device rd;  static std::mt19937 gen(rd());  return select\_randomly(start, end, gen);  }  // Запускает функцию fn на n потоках, включая текущий  template <typename Fn>  void RunThreads(unsigned n, const Fn& fn) {  n = std::max(1u, n);  std::vector<std::jthread> workers;  workers.reserve(n - 1);  // Запускаем n-1 рабочих потоков, выполняющих функцию fn  while (--n) {  workers.emplace\_back(fn);  }  fn();  }  std::string StringifyRequest(client\_request\_e r){  switch(r){  case 0: return "c\_set\_name"s;  case 1: return "c\_create\_chat"s;  case 2: return "c\_connect\_chat"s;  case 3: return "c\_send\_message"s;  case 4: return "c\_leave\_chat"s;  case 5: return "c\_disconnect"s;  case 6: return "c\_receive\_message"s;  case 7: return "c\_wrong\_request"s;  case 8: return "c\_get\_chats"s;  default: return "wrong input"s;  }  }  std::string StringifyResponse(server\_response\_e r){  switch(r){  case 0: return "s\_success"s;  case 1: return "s\_failure"s;  default: return "wrong input"s;  }  }  net::awaitable<void> Write(std::shared\_ptr<tcp::socket> socket, client\_data\_t&& request){  size\_t message\_length = request.message\_text.size();  boost::scoped\_ptr<client\_data\_t> safe\_request(new client\_data\_t(request));  // std::cout << "Sending message: {request: " << response.request  // << ", response: " << response.response << ", message: \"" << response.message\_text << "\"}" << std::endl;  auto [ec, bytes] = co\_await net::async\_write(  \*socket,  boost::array<boost::asio::mutable\_buffer, 4>({  net::buffer((void\*)&safe\_request->request, sizeof(client\_request\_e)),  net::buffer((void\*)&message\_length, sizeof(size\_t)),  net::buffer(safe\_request->message\_text)  }),  // boost::asio::transfer\_all(), // maybe needed  net::experimental::as\_tuple(net::use\_awaitable)  );  if(ec) {  //log here  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Client write error"sv, data);  throw ec;  }  boost::json::value data = {  {"request",  {  {"client\_request\_e", StringifyRequest(request.request)},  {"message", request.message\_text}  }  }  };  logger::Logger::GetInstance().Log("Wrote request to server"sv, data);  }  net::awaitable<server\_data\_t> Read(std::shared\_ptr<tcp::socket> socket){  server\_data\_t response;  size\_t message\_length;  auto [ec, bytes] = co\_await  boost::asio::async\_read(  \*socket,  boost::array<boost::asio::mutable\_buffer, 3>({  net::buffer((void\*)&response.request, sizeof(client\_request\_e)),  net::buffer((void\*)&response.response, sizeof(server\_response\_e)),  net::buffer((void\*)&message\_length, sizeof(size\_t)),  }),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if (!ec) {  response.message\_text.resize(message\_length);  auto [ec2, bytes2] = co\_await boost::asio::async\_read(  \*socket,  boost::asio::buffer(response.message\_text),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if(ec2) {  //log here  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Client read error"sv, data);  throw ec2;  }  }  else {  //log here  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Client read error"sv, data);  throw ec;  }  boost::json::value data = {  {"response",  {  {"client\_request\_e", StringifyRequest(response.request)},  {"server\_response\_e", StringifyResponse(response.response)},  {"message", response.message\_text}  }  }  };  logger::Logger::GetInstance().Log("Read response from server"sv, data);  co\_return response;  }  net::awaitable<server\_data\_t> Read(std::shared\_ptr<tcp::socket> socket, std::chrono::system\_clock::time\_point start\_ts){  server\_data\_t response;  size\_t message\_length;  auto [ec, bytes] = co\_await  boost::asio::async\_read(  \*socket,  boost::array<boost::asio::mutable\_buffer, 3>({  net::buffer((void\*)&response.request, sizeof(client\_request\_e)),  net::buffer((void\*)&response.response, sizeof(server\_response\_e)),  net::buffer((void\*)&message\_length, sizeof(size\_t)),  }),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if (!ec) {  response.message\_text.resize(message\_length);  auto [ec2, bytes2] = co\_await boost::asio::async\_read(  \*socket,  boost::asio::buffer(response.message\_text),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if(ec2) {  //log here  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Client read error"sv, data);  throw ec2;  }  }  else {  //log here  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Client read error"sv, data);  throw ec;  }  std::chrono::system\_clock::time\_point end\_ts = std::chrono::system\_clock::now();  boost::json::value data = {  {"response",  {  {"client\_request\_e", StringifyRequest(response.request)},  {"server\_response\_e", StringifyResponse(response.response)},  {"message", response.message\_text}  }  },  {"response time", (end\_ts - start\_ts).count()}  };  logger::Logger::GetInstance().Log("Read first response from server"sv, data);  co\_return response;  }  std::vector<std::string> SplitString(std::string str, const std::string& delimiter){  size\_t pos;  std::vector<std::string> split;  while ((pos = str.find(delimiter)) != std::string::npos) {  split.push\_back(str.substr(0, pos));  str.erase(0, pos + delimiter.length());  }  return split;  }  boost::asio::awaitable<void> User(tcp::socket socket\_){  try{  std::shared\_ptr<tcp::socket> socket = std::make\_shared<tcp::socket>(std::move(socket\_));  client\_data\_t request;  server\_data\_t response;  ////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////  //set name  std::random\_device random\_device;  std::mt19937\_64 generator{[&] {  std::uniform\_int\_distribution<std::mt19937\_64::result\_type> dist;  return dist(random\_device);  }()};  do{ // try until the name will be set  //generating random name  std::stringstream ss;  ss << "Bot" << std::setw(8) << std::setfill('0') << std::hex << generator();  request.message\_text = ss.str();  request.request = c\_set\_name;  co\_await Write(socket, std::move(request));  response = co\_await Read(socket);  }while(response.response == s\_failure);  while(true){ //loop to switch between being in and out of chat  ////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////  //list of chats  //getting rid of extra messages  do{  response = co\_await Read(socket); // receiving list of chats  }while(response.request != c\_get\_chats); // other messages are just ignored  //split here  auto chats = SplitString(response.message\_text, "\n"s);  while(true){ // list of chat loop  if(chats.size() < 4 || generator() % 1000 < 2){  // 0.2% possibility to create chat & if no chats -> it will be created by default  request.request = c\_create\_chat;  // generating random chat name  std::stringstream ss;  ss << "Chat" << std::setw(8) << std::setfill('0') << std::hex << generator();  request.message\_text = ss.str();  co\_await Write(socket, std::move(request));  while(true){  response = co\_await Read(socket);  if(response.request == c\_get\_chats) {  chats = SplitString(response.message\_text, "\n"s);  break;  }  }  }  else{ // connect to chat  request.request = c\_connect\_chat;  request.message\_text = \*select\_randomly(chats.begin(), chats.end());  co\_await Write(socket, std::move(request));  while(true){  response = co\_await Read(socket);  if(response.request == c\_get\_chats) chats = SplitString(response.message\_text, "\n"s);  else break;  }  if(response.response == s\_success) break; // after connecting  }  }  ////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////  //in\_chat  while(true){ // chat loop  //getting rid of extra messages  while(socket->available() > 0){  response = co\_await Read(socket);  if(response.request != c\_get\_chats && response.request != c\_receive\_message) break;  }  auto i = generator();  if(i % 100 < 5){ // 5% probability that user leaves chat  request.message\_text.clear();  request.request = c\_leave\_chat;  co\_await Write(socket, std::move(request));  break;  }  else{  // random message  std::stringstream ss;  ss << std::setw(16) << std::setfill('0') << std::hex << generator() << generator();  request.message\_text = ss.str();  request.request = c\_send\_message;  co\_await Write(socket, std::move(request));  }  }  ////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////  //after exiting the chat loop user returns to out-of-chat loop  }  }  catch(sys::error\_code& ec){}  catch(std::exception& ex){  boost::json::value data = {  {"exception", ex.what()}  };  logger::Logger::GetInstance().Log("Client caught exception"sv, data);  }  catch(...){  boost::json::value data = {};  logger::Logger::GetInstance().Log("Client unknown error"sv, data);  }  boost::json::value data = {};  logger::Logger::GetInstance().Log("Client finished it's work"sv, data);  co\_return;  }  int main(int argc, const char\* argv[]) {  if(argc != 2){  std::cout << "The only one correct parameter - number of user (currently " << argc - 1 << "arguments used)" << std::endl;  return EXIT\_FAILURE;  }  int N = std::stoi(argv[1]);  try {  LOG\_INIT()  // 1. Инициализируем io\_context  const unsigned num\_threads = std::thread::hardware\_concurrency();  net::io\_context ioc(num\_threads);  // 2. Добавляем асинхронный обработчик сигналов SIGINT и SIGTERM  net::signal\_set signals(ioc, SIGINT, SIGTERM);  signals.async\_wait([&ioc](const sys::error\_code& ec, [[maybe\_unused]] int signal\_number) {  if (!ec) {  ioc.stop();  }  });  // 3. Настройка сетевых параметров  constexpr boost::asio::ip::port\_type port = 5000;  // пока локальный адрес, но дальше надо вписать другой  net::ip::tcp::endpoint endpoint = net::ip::tcp::endpoint(net::ip::address\_v4::from\_string("127.0.0.1"), port);  //4. Запускаем N клиентов внутри контекста  for(auto i = 0; i < N; ++i){  tcp::socket socket(ioc);  tcp::resolver resolver(ioc);  net::connect(socket, resolver.resolve(endpoint));  boost::asio::co\_spawn(ioc, User(std::move(socket)), net::detached);  }  // Эта надпись сообщает тестам о том, что тестер запущен и готов генерировать нагрузку  boost::json::value data = {  {"port", port},  {"endpoint address", endpoint.address().to\_string()}  };  logger::Logger::GetInstance().Log("Tester started"sv, data);  // 5. Запускаем обработку асинхронных операций  RunThreads(std::max(1u, num\_threads), [&ioc] {  ioc.run();  });  }  catch (const std::exception& ex) {  boost::json::value data = {  {"exception", ex.what()}  };  logger::Logger::GetInstance().Log("Tester caught exception"sv, data);  return EXIT\_FAILURE;  }  } |

Приложение Б. Файл server.cpp многопоточного сервера

|  |
| --- |
| #include "server\_interface.h"  #include "server\_routines.h"  #include "logger.h"  extern std::unordered\_map<std::string, chat\_info> chats;  extern std::unordered\_set<std::string> used\_usernames;  extern std::unordered\_map<int, std::string> user\_data;  extern int disco\_pipe[2];  extern int list\_chats\_pipe[2];  extern int username\_enter\_pipe[2];  int main(int argc, char \*argv[])  {  LOG\_INIT()  using namespace std::chrono\_literals;  if (!server\_starting())  exit(EXIT\_FAILURE);  int listen\_socket;  struct sockaddr\_in address;  // fd\_set readfds;  /\* creates an UN-named socket inside the kernel and returns  \* an integer known as socket descriptor  \* This function takes domain/family as its first argument.  \* For Internet family of IPv4 addresses we use AF\_INET  \*/  listen\_socket = socket(AF\_INET, SOCK\_STREAM /\*| SOCK\_NONBLOCK\*/, IPPROTO\_TCP);  bzero(&address, sizeof(address));  address.sin\_port = htons(5000);  address.sin\_family = AF\_INET;  address.sin\_addr.s\_addr = inet\_addr("0.0.0.0");  /\*  \* The call to the function "bind()" assigns the details specified  \* in the structure address to the socket created in the step above  \*/  if (bind(listen\_socket, (struct sockaddr \*)&address, sizeof(address)) == -1)  {  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("Bind error"sv, data);  close(listen\_socket);  return EXIT\_FAILURE;  }  if (listen(listen\_socket, 1000) == -1)  {  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("Listen error"sv, data);  shutdown(listen\_socket, SHUT\_RDWR);  return EXIT\_FAILURE;  }  if(pipe(list\_chats\_pipe) < 0){  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("List of chats pipe opening error"sv, data);  shutdown(listen\_socket, SHUT\_RDWR);  return EXIT\_FAILURE;  }  if(pipe(disco\_pipe) < 0){  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("Disconnection pipe opening error"sv, data);  shutdown(listen\_socket, SHUT\_RDWR);  close(list\_chats\_pipe[0]);  close(list\_chats\_pipe[1]);  return EXIT\_FAILURE;  }  if(pipe(username\_enter\_pipe) < 0){  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("Username enter pipe opening error"sv, data);  shutdown(listen\_socket, SHUT\_RDWR);  close(list\_chats\_pipe[0]);  close(list\_chats\_pipe[1]);  close(disco\_pipe[0]);  close(disco\_pipe[1]);  return EXIT\_FAILURE;  }  fcntl(list\_chats\_pipe[0], F\_SETFL, fcntl(list\_chats\_pipe[0], F\_GETFL) | O\_NONBLOCK);  fcntl(disco\_pipe[0], F\_SETFL, fcntl(disco\_pipe[0], F\_GETFL) | O\_NONBLOCK);  fcntl(username\_enter\_pipe[0], F\_SETFL, fcntl(username\_enter\_pipe[0], F\_GETFL) | O\_NONBLOCK);  std::thread(user\_name\_enter, username\_enter\_pipe[0], list\_chats\_pipe[1]).detach();  std::thread(list\_of\_chats, list\_chats\_pipe[0]).detach();  std::thread(disconnect).detach();  struct sockaddr\_in client\_address;  socklen\_t address\_length = sizeof(sockaddr\_in);  int connection\_socket;  while ((connection\_socket = accept(listen\_socket, (struct sockaddr \*)&client\_address, &address\_length)) != -1)  {  struct timeval tv;  tv.tv\_sec = 0;  tv.tv\_usec = TIMEOUT\_MICRO;  setsockopt(connection\_socket, SOL\_SOCKET, SO\_RCVTIMEO, (const char\*)&tv, sizeof tv);  write(username\_enter\_pipe[1], &connection\_socket, sizeof(int));  }  if (!(errno & (EAGAIN | EWOULDBLOCK)))  {  boost::json::value data = {{"error code", std::error\_code{errno, std::generic\_category()}.message()}};  logger::Logger::GetInstance().Log("Accept failure"sv, data);  return EXIT\_FAILURE;  }  // perror("Server died with status");  server\_ending(); // because why not  exit(EXIT\_SUCCESS);  } |

Приложение В. Файл main.cpp многопоточно-асинхронного сервера

|  |
| --- |
| #include <boost/asio/io\_context.hpp>  #include <boost/asio/signal\_set.hpp>  #include <iostream>  #include <thread>  #include "server.h"  #include "logger.h"  using namespace std::literals;  namespace net = boost::asio;  namespace sys = boost::system;  namespace {  // Запускает функцию fn на n потоках, включая текущий  template <typename Fn>  void RunThreads(unsigned n, const Fn& fn) {  n = std::max(1u, n);  std::vector<std::jthread> workers;  workers.reserve(n - 1);  // Запускаем n-1 рабочих потоков, выполняющих функцию fn  while (--n) {  workers.emplace\_back(fn);  }  fn();  }  } // namespace  int main(int argc, const char\* argv[]) {  try {  LOG\_INIT()  // 1. Инициализируем io\_context  const unsigned num\_threads = std::thread::hardware\_concurrency();  net::io\_context ioc(num\_threads);  // 2. Добавляем асинхронный обработчик сигналов SIGINT и SIGTERM  net::signal\_set signals(ioc, SIGINT, SIGTERM);  signals.async\_wait([&ioc](const sys::error\_code& ec, [[maybe\_unused]] int signal\_number) {  if (!ec) {  boost::json::value data = {  {"signal", signal\_number}  };  logger::Logger::GetInstance().Log("Signal received"sv, data);  ioc.stop();  }  });  // 3. Настройка сетевых параметров  constexpr boost::asio::ip::port\_type port = 5000;  const auto address = net::ip::tcp::endpoint(net::ip::tcp::v4(), port).address();  net::ip::tcp::endpoint endpoint = {address, port};  //4. Запускаем серверную обработку запросов  std::make\_shared<server::Server>(ioc, endpoint, "./chats"s)->Run();  // Эта надпись сообщает тестам о том, что сервер запущен и готов обрабатывать запросы  boost::json::value data = {  {"address", address.to\_string()},  {"port", port}  };  logger::Logger::GetInstance().Log("Server started"sv, data);  // 5. Запускаем обработку асинхронных операций  RunThreads(std::max(1u, num\_threads), [&ioc] {  ioc.run();  });  }  catch (const std::exception& ex) {  boost::json::value data = {  {"exception", ex.what()}  };  logger::Logger::GetInstance().Log("Exception occured"sv, data);  return EXIT\_FAILURE;  }  } |

Приложение Г. Файл server.cpp многопоточно-асинхронного сервера

|  |
| --- |
| #include "server.h"  #include "logger.h"  #include <iostream>  #include <filesystem>  #include <fstream>  #include <ctime>  #include <boost/asio/detached.hpp>  #include <boost/asio/co\_spawn.hpp>  #include <boost/scoped\_ptr.hpp>  namespace server {  std::string StringifyRequest(client\_request\_e r){  switch(r){  case 0: return "c\_set\_name"s;  case 1: return "c\_create\_chat"s;  case 2: return "c\_connect\_chat"s;  case 3: return "c\_send\_message"s;  case 4: return "c\_leave\_chat"s;  case 5: return "c\_disconnect"s;  case 6: return "c\_receive\_message"s;  case 7: return "c\_wrong\_request"s;  case 8: return "c\_get\_chats"s;  default: return "wrong input"s;  }  }  std::string StringifyResponse(server\_response\_e r){  switch(r){  case 0: return "s\_success"s;  case 1: return "s\_failure"s;  default: return "wrong input"s;  }  }  net::awaitable<void> Write(std::shared\_ptr<tcp::socket> socket, server\_data\_t&& response){  size\_t message\_length = response.message\_text.size();  boost::scoped\_ptr<server\_data\_t> safe\_response(new server\_data\_t(response));  boost::json::value data = {  {"response",  {  {"client\_request\_e", StringifyRequest(safe\_response->request)},  {"server\_response\_e", StringifyResponse(safe\_response->response)},  {"message", safe\_response->message\_text}  }  },  {"socket", int(socket->native\_handle())}  };  logger::Logger::GetInstance().Log("Wrote response"sv, data);  auto [ec, bytes] = co\_await net::async\_write(  \*socket,  boost::array<boost::asio::mutable\_buffer, 4>({  net::buffer((void\*)&safe\_response->request, sizeof(client\_request\_e)),  net::buffer((void\*)&safe\_response->response, sizeof(server\_response\_e)),  net::buffer((void\*)&message\_length, sizeof(size\_t)),  net::buffer(safe\_response->message\_text)  }),  // boost::asio::transfer\_all(), // maybe needed  net::experimental::as\_tuple(net::use\_awaitable)  );  if(ec) {  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Write error"sv, data);  throw ec;  }  }  // Chat methods  boost::asio::awaitable<void> Chat::AddUser(std::shared\_ptr<tcp::socket> socket, std::string username){  co\_await net::dispatch(strand, boost::asio::use\_awaitable);  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", username},  {"chat\_name", path.stem().c\_str()}  };  logger::Logger::GetInstance().Log("User added to chat"sv, data);  messages\_offset.insert({socket, 0});  std::fstream chat\_file(path);  chat\_file.seekp(0, chat\_file.end);  auto file\_size = chat\_file.tellp();  server\_data\_t resp;  std::streamoff off = 0;  chat\_file.seekg(off, chat\_file.beg);  while (off != file\_size) {  resp.response = server\_response\_e::s\_success;  resp.request = client\_request\_e::c\_receive\_message; // now we change the request that we are serving because we did connect to chat  std::getline(chat\_file, resp.message\_text, '\r');  off = chat\_file.tellp();  co\_await Write(socket, std::move(resp));  // invoking saved method to send all unread chat messages to connected user  // std::cout << "[off = " << off << ", file\_size = " << file\_size << "]\n";  }  messages\_offset[socket] = off;  usernames[socket] = username;  }  boost::asio::awaitable<void> Chat::DeleteUser(std::shared\_ptr<tcp::socket> socket){  co\_await net::dispatch(strand, boost::asio::use\_awaitable);  std::string username;  if(usernames.contains(socket)) username = usernames.at(socket);  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", username},  {"chat\_name", path.stem().c\_str()}  };  logger::Logger::GetInstance().Log("User deleted from chat"sv, data);    messages\_offset.erase(socket);  usernames.erase(socket);  }  boost::asio::awaitable<void> Chat::SendMessage(std::shared\_ptr<tcp::socket> socket, std::string message){  co\_await net::dispatch(strand, boost::asio::use\_awaitable);  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", usernames.at(socket)},  {"chat\_name", path.stem().c\_str()},  {"message", message}  };  logger::Logger::GetInstance().Log("User sent message to chat"sv, data);  if(messages\_offset.contains(socket)){  std::fstream chat\_file;    chat\_file.open(path, std::fstream::out | std::fstream::app);  chat\_file << usernames[socket] << "\t";    time\_t curr\_time;  tm \* curr\_tm;  std::string timedate;  timedate.reserve(100);  time(&curr\_time);  curr\_tm = localtime(&curr\_time);  strftime(timedate.data(), 99, "%T %D", curr\_tm);  chat\_file << timedate.c\_str() << std::endl << message << "\r";    auto file\_size = chat\_file.tellp();  chat\_file.close();  chat\_file.open(path, std::fstream::in | std::fstream::app);  for(auto& [usr\_socket, off] : messages\_offset){  chat\_file.seekg(off, chat\_file.beg);  while(off != file\_size){  server\_data\_t resp;  resp.request = client\_request\_e::c\_receive\_message;  resp.response = server\_response\_e::s\_success;  std::getline(chat\_file, resp.message\_text, '\r');  off = chat\_file.tellp();  try{  co\_await Write(usr\_socket, std::move(resp));  } catch(sys::error\_code& ec){  if(usr\_socket == socket) throw ec;  // if our socket is broken - that's a problem  // else - let it be destroyed afterwards by it's session  }  }  }  chat\_file.close();  }  }  boost::asio::awaitable<bool> Chat::ContainsUser(std::shared\_ptr<tcp::socket> socket){  co\_await net::dispatch(strand, boost::asio::use\_awaitable);  std::string username;  if(usernames.contains(socket)) username = usernames.at(socket);  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", username},  {"chat\_name", path.stem().c\_str()}  };  logger::Logger::GetInstance().Log("Checked if user is in the chat"sv, data);  co\_return messages\_offset.contains(socket);  }  // // ChatManager methods  boost::asio::awaitable<bool> ChatManager::SetName(std::string name, std::shared\_ptr<tcp::socket> socket){  bool result = true;  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  for(auto [socket\_ptr, user] : users\_){  if(user.name == name){  result = false;  break;  }  }  if(result) {  users\_.insert({socket, {name, ""}});  }  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", name},  {"result", int(result)} // maybe change to another type  };  logger::Logger::GetInstance().Log("Name requested by a user"sv, data);  co\_return result;  }  boost::asio::awaitable<std::string> ChatManager::ChatList(){  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);  std::string list;  for(auto chat : chats\_) list += chat.first + "\n";  boost::json::value data = {  {"chat list", list}  };  logger::Logger::GetInstance().Log("Formed a chatlist"sv, data);  co\_return list;  }  boost::asio::awaitable<void> ChatManager::UpdateChatList(std::shared\_ptr<tcp::socket> socket){  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  server\_data\_t resp;  std::string list = co\_await ChatList();  boost::json::value data = {  {"chat list", list}  };  logger::Logger::GetInstance().Log("Chat list updated"sv, data);  for(auto [user\_socket, user] : users\_){  if(user.chat\_name.empty()){  resp.message\_text = list;  resp.request = c\_get\_chats;  resp.response = s\_success;  try{  co\_await Write(user\_socket, std::move(resp));  } catch(sys::error\_code& ec){  if(user\_socket == socket) throw ec;  // our socket - our problem, else - nah  }  }  }  }  boost::asio::awaitable<void> ChatManager::Disconnect(std::shared\_ptr<tcp::socket> socket){  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);    std::string name;  if(users\_.contains(socket)) name = users\_.at(socket).name;  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", name}  };  logger::Logger::GetInstance().Log("User disconnected"sv, data);  if(users\_.contains(socket)) {  User usr = users\_.at(socket);  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);  //synchronized deletion of the user listing in chat through chat's method  if(chats\_.contains(usr.chat\_name)) {  co\_await chats\_.at(usr.chat\_name).DeleteUser(socket); // throws exception because ???  }  //synchronized deletion of the user from the list of users in the chat manager  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  users\_.erase(socket);  }  }  boost::asio::awaitable<void> ChatManager::ConnectChat(std::string chat\_name, std::shared\_ptr<tcp::socket> socket){ //Sends everything the client needs, including success or failure notification  using namespace std::literals;  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  std::string name;  if(users\_.contains(socket)) name = users\_.at(socket).name;  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", name},  {"chat name", chat\_name}  };  logger::Logger::GetInstance().Log("User connected to chat"sv, data);  if(users\_.contains(socket)) {  User usr = users\_.at(socket);  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);    server\_data\_t resp;  resp.message\_text = chat\_name;  resp.request = client\_request\_e::c\_connect\_chat;  if(!chats\_.contains(chat\_name)) {  resp.response = server\_response\_e::s\_failure;  co\_await Write(socket, std::move(resp)); // invoking saved method to inform client that the chat failed to connect  co\_return;  }    resp.response = server\_response\_e::s\_success;  co\_await Write(socket, std::move(resp)); // invoking saved method to inform client that the chat is connected RN  //keep the state because AddUser method will be the last one to send the message  co\_await chats\_.at(chat\_name).AddUser(socket, usr.name);    //synchronized update of user status  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);    if(users\_.contains(socket)) users\_.at(socket).chat\_name = chat\_name;  }  }  boost::asio::awaitable<bool> ChatManager::CreateChat(std::string chat\_name){  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);  bool result = true;  if(chats\_.contains(chat\_name) || chat\_name == ""){  result = false;  }  if(result) chats\_.insert({chat\_name, {io\_, path\_of\_chats / (chat\_name + ".chat")}});  boost::json::value data = {  {"chat name", chat\_name},  {"result", int(result)}  };  logger::Logger::GetInstance().Log("Chat creation request"sv, data);  co\_return result;  }  boost::asio::awaitable<bool> ChatManager::LeaveChat(std::shared\_ptr<tcp::socket> socket){  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  bool result;  if(users\_.contains(socket)){  User usr = users\_.at(socket);  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);  if(!chats\_.contains(usr.chat\_name) ||  !co\_await chats\_.at(usr.chat\_name).ContainsUser(socket)){  result = false;  }  else{  co\_await chats\_.at(usr.chat\_name).DeleteUser(socket);  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  if(users\_.contains(socket)) users\_.at(socket).chat\_name.clear();  result = true;  }  }  std::string name;  if(users\_.contains(socket)) name = users\_.at(socket).name;  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", name},  {"result", int(result)}  };  logger::Logger::GetInstance().Log("Leave chat request"sv, data);  co\_return result;  }  boost::asio::awaitable<void> ChatManager::SendMessage(std::string message, std::shared\_ptr<tcp::socket> socket){ // this one actually sends messages by itself  co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  std::string name;  if(users\_.contains(socket)) name = users\_.at(socket).name;  boost::json::value data = {  {"socket", int(socket->native\_handle())},  {"username", name},  {"message", message}  };  logger::Logger::GetInstance().Log("Send message request"sv, data);  if(users\_.contains(socket)){  User usr = users\_.at(socket);  co\_await net::dispatch(chats\_strand\_, boost::asio::use\_awaitable);  if(chats\_.contains(usr.chat\_name)) co\_await chats\_.at(usr.chat\_name).SendMessage(socket, message);  }  }  // boost::asio::awaitable<User\*> ChatManager::IdentifyUser(tcp::socket\* socket){  // co\_await net::dispatch(usernames\_strand\_, boost::asio::use\_awaitable);  // boost::json::value data = {  // {"socket", int(socket->native\_handle())},  // };  // logger::Logger::GetInstance().Log("Identify user request"sv, data);  // User\* usr;  // if(!users\_.contains(socket)) usr = nullptr;  // else usr = &users\_.at(socket);  // co\_return usr;  // }  // Server methods  void Server::Run(){  DoAccept();  }  std::shared\_ptr<Server> Server::GetSharedThis(){  return this->shared\_from\_this();  }  void Server::DoAccept() {  // Метод socket::async\_accept создаст сокет и передаст его передан в OnAccept  acceptor\_.async\_accept(  // Передаём последовательный исполнитель, в котором будут вызываться обработчики  // асинхронных операций сокета  net::make\_strand(ioc\_),  // При помощи bind\_front\_handler создаём обработчик, привязанный к методу OnAccept  // текущего объекта.  beast::bind\_front\_handler(&Server::OnAccept, GetSharedThis()));  }  void Server::OnAccept(sys::error\_code ec, tcp::socket socket) {  using namespace std::literals;  if (ec) {  boost::json::value data = {  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Accept error"sv, data);  }  // Для того, чтобы подключение произошло  // нужно прочитать сначала данные для подтверждения подключения    // Асинхронно обрабатываем сессию  AsyncRunSession(std::move(socket));  // Принимаем новое соединение  DoAccept();  }  void Server::AsyncRunSession(tcp::socket&& socket) {  boost::asio::co\_spawn(ioc\_, Session(std::move(socket), &chat\_manager\_), boost::asio::detached);  }  // Session  boost::asio::awaitable<void> Session(tcp::socket&& socket, ChatManager\* chatManager){  bool running = true;  session\_state sessionState = session\_state::reading;  client\_state clientState = client\_state::no\_name;  client\_data\_t request\_;  server\_data\_t response\_;  size\_t message\_length;  std::shared\_ptr<tcp::socket> socket\_ = std::make\_shared<tcp::socket>(std::move(socket));  boost::json::value data = {  {"socket", int(socket\_->native\_handle())}  };  logger::Logger::GetInstance().Log("Session started"sv, data);  try{  while(true){  auto [ec, bytes] = co\_await  boost::asio::async\_read(  \*socket\_,  boost::array<boost::asio::mutable\_buffer, 2>({  boost::asio::buffer((void\*)&request\_.request, sizeof(request\_.request)),  boost::asio::buffer((void\*)&message\_length, sizeof(message\_length))  }),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if (!ec) {  request\_.message\_text.resize(message\_length);  auto [ec2, bytes2] = co\_await  boost::asio::async\_read(  \*socket\_,  boost::asio::buffer(request\_.message\_text),  // boost::asio::transfer\_all(), // maybe needed  boost::asio::experimental::as\_tuple(boost::asio::use\_awaitable)  );  if(!ec2) sessionState = session\_state::handling;  else {  //log here  boost::json::value data = {  {"socket", int(socket\_->native\_handle())},  {"error", ec2.what()}  };  logger::Logger::GetInstance().Log("Read error"sv, data);  throw ec2;  }  }  else {  //log here  boost::json::value data = {  {"socket", int(socket\_->native\_handle())},  {"error", ec.what()}  };  logger::Logger::GetInstance().Log("Read error"sv, data);  throw ec;  }  // std::cout << "Received message: {request: " << request\_.request << ", message: \"" << request\_.message\_text << "\"}" << std::endl;  response\_.request = request\_.request;  switch(request\_.request){  case client\_request\_e::c\_set\_name:  if(clientState != client\_state::no\_name){  response\_.message\_text = "Cannot set name twice!";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  } else {  if(co\_await chatManager->SetName(request\_.message\_text, socket\_)){  clientState = client\_state::list\_chats;  response\_.message\_text = "";  response\_.response = server\_response\_e::s\_success;  co\_await Write(socket\_, std::move(response\_));  response\_.request = client\_request\_e::c\_get\_chats;  response\_.response = server\_response\_e::s\_success;  response\_.message\_text = co\_await chatManager->ChatList();    co\_await Write(socket\_, std::move(response\_));  } else {  response\_.message\_text = "Username is already in use";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  }  break;  case client\_request\_e::c\_create\_chat:  if(clientState != client\_state::list\_chats){  response\_.message\_text = "Cannot create chats, wrong state";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  else{  if(!co\_await chatManager->CreateChat(request\_.message\_text)){  response\_.message\_text = "Chat with such name already exists";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  else{  // send everyone out of chat new list  co\_await chatManager->UpdateChatList(socket\_);  }  }  break;  case client\_request\_e::c\_connect\_chat:  if(clientState != client\_state::list\_chats){  response\_.message\_text = "Cannot connect chat, wrong state!";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  else {  co\_await chatManager->ConnectChat(request\_.message\_text, socket\_);  clientState = client\_state::in\_chat;  }  break;  case client\_request\_e::c\_send\_message:  if(clientState != client\_state::in\_chat){  response\_.message\_text = "";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  else  co\_await chatManager->SendMessage(request\_.message\_text, socket\_);  break;  case client\_request\_e::c\_leave\_chat:  if(clientState == client\_state::in\_chat && co\_await chatManager->LeaveChat(socket\_)){  // Can leave chat, ok state && could to leave the chat  clientState = client\_state::list\_chats;  response\_.request = client\_request\_e::c\_get\_chats;  response\_.response = server\_response\_e::s\_success;  response\_.message\_text = co\_await chatManager->ChatList();  co\_await Write(socket\_, std::move(response\_));  }  else{  response\_.message\_text = "";  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  //no messages about leaving  break;  default:  response\_.message\_text = "Wrong request";  response\_.request = client\_request\_e::c\_wrong\_request;  response\_.response = server\_response\_e::s\_failure;  co\_await Write(socket\_, std::move(response\_));  }  }  }  catch(std::exception ex){  boost::json::value data = {  {"exception", ex.what()}  };  logger::Logger::GetInstance().Log("Exception occured"sv, data);  }  catch(...){}  // catch(beast::error\_code ec){ //read\write errors  // std::cout << "Boost error occured: " << ec.what() << std::endl;  // }  //disconnection  // std::cout << "Session ended!" << std::endl;  co\_await chatManager->Disconnect(socket\_);  data = {  {"socket", int(socket\_->native\_handle())}  };  logger::Logger::GetInstance().Log("Session ended"sv, data);    socket\_->close();  }  } // namespace server |

Приложение Д. Файл mainwindow.cpp клиентского приложения

|  |
| --- |
| #include "mainwindow.h"  #include "./ui\_mainwindow.h"  #include <string.h>  #include <QTableView>  #include <QListWidget>  #include <QPushButton>  #include <QDebug>  #include <QHeaderView>  #include <QStandardItem>  #include "client\_interface.h"  MainWindow::MainWindow(QWidget \*parent)  : QMainWindow(parent)  , ui(new Ui::MainWindow)  {  ui->setupUi(this);  MainWindow::setWindowTitle("Chat");  socket = new QTcpSocket(this);  bool start\_ok = StartConnection();  connect(ui->NameLine, &QLineEdit::returnPressed, this, &MainWindow::on\_nameLine\_returnPressed1st);  CRH = new ReadRoutineHandler(this, socket);  if(start\_ok){  CRH->startThread();  }  else {  QMessageBox::warning(this, "ACHTUNG!", "Server error: could not connect.\nTry reopening app once we reanimate the server!");  this->close();  }  }  void MainWindow::on\_serverDown(){  QMessageBox::warning(this, "ACHTUNG!", "Server is down, try reconnecting later");  this->close();  }  bool MainWindow::StartConnection(){  socket->connectToHost("127.0.0.1", 5000);  socket->setReadBufferSize(10000);  connect(socket, &QTcpSocket::disconnected, socket, &QTcpSocket::deleteLater);  if(socket->waitForConnected(3000)){  qDebug() << "Connected!";  connect(socket, &QTcpSocket::errorOccurred, this, &MainWindow::on\_serverDown);  return true;  }  else{  qDebug() << "Not connected!";  return false;  }  }  void MainWindow::show\_chats(const std::string& chat\_list){  //deleting start ui  MainWindow::setWindowTitle("Chat");  SafeCleaning(ui->verticalLayout->layout());  //creating mid ui  QLineEdit\* NewChatLine = new QLineEdit();  NewChatLine->setPlaceholderText("Enter new chat's name:");  connect(NewChatLine, &QLineEdit::returnPressed, this, &MainWindow::on\_createChat1st);  ui->verticalLayout->insertWidget(0, NewChatLine);  QTableView\* ChatsTable = new QTableView;  QStringList chats = QString::fromStdString(chat\_list).split('\n');  model = new QStandardItemModel(chats.length() - 1, 1);  ChatsTable->setModel(model);  for(size\_t row = 0; row < chats.length() - 1; ++row){  auto chat = new QPushButton();  chat->setText(chats.at(row));  ChatsTable->setIndexWidget(model->index(row, 0), chat);  connect(chat, &QPushButton::clicked, this, &MainWindow::on\_chatButton\_Clicked1st);  connect(ChatsTable, &QTableView::destroyed, chat, &QPushButton::deleteLater);  }  ChatsTable->horizontalHeader()->hide();  ChatsTable->verticalHeader()->hide();  ChatsTable->horizontalHeader()->setStretchLastSection(true); //try to fill the whole widget  ui->verticalLayout->insertWidget(1, ChatsTable);  ChatsTable->show();  }  MainWindow::~MainWindow()  {  delete ui;  CRH->stopThread();  delete CRH;  }  void MainWindow::on\_nameLine\_returnPressed1st()  {  if(ui->NameLine->text().size() > 32){  QMessageBox::warning(this, "ACHTUNG!","This name is too big!");  ui->NameLine->clear();  }  else{  c\_message.request = c\_set\_name;  c\_message.message\_text = ui->NameLine->text().toStdString();  SendToServer(socket, &c\_message);  }  }  void MainWindow::on\_nameLine\_returnPressed2nd(server\_data\_t s\_message){  if(s\_message.responce == s\_failure){  QMessageBox::warning(this, "ACHTUNG!","This username is already used!");  }  }  void MainWindow::list\_of\_chats(server\_data\_t s\_message){  if(s\_message.responce == s\_failure){  QMessageBox::warning(this, "ACHTUNG!", "Could not get list of chats");  }  else{  show\_chats(s\_message.message\_text);  }  }  void MainWindow::on\_chatButton\_Clicked1st(){  QPushButton\* chat = qobject\_cast<QPushButton\*>(sender()); // get QObject that emitted the signal  c\_message.request = c\_connect\_chat;  c\_message.message\_text = chat->text().toStdString();  SendToServer(socket, &c\_message);  }  void MainWindow::on\_chatButton\_Clicked2nd(server\_data\_t s\_message){  if(s\_message.responce == s\_success){  SafeCleaning(ui->verticalLayout->layout());  MainWindow::setWindowTitle("Chat: " + QString::fromStdString(s\_message.message\_text));  QPushButton\* LeaveButton = new QPushButton("Leave chat");  connect(LeaveButton, &QPushButton::clicked, this, &MainWindow::on\_leaveChat1st);  QListWidget\* QLW = new QListWidget();  QLineEdit\* MessageLine = new QLineEdit();  MessageLine->setPlaceholderText("Enter your message:");  connect(MessageLine, &QLineEdit::returnPressed, this, &MainWindow::on\_sendMessage);  ui->verticalLayout->insertWidget(0, LeaveButton);  ui->verticalLayout->insertWidget(1, QLW);  ui->verticalLayout->insertWidget(2, MessageLine);  }  else{  QMessageBox::warning(this, "ACHTUNG!", "Could not connect to this chat!");  }  }  void MainWindow::on\_newMessage(std::string& message\_text){  static size\_t row = 0;  auto QLW = qobject\_cast<QListWidget\*>(ui->verticalLayout->itemAt(1)->widget());  QLW->addItem(QString::fromStdString(message\_text));  }  void MainWindow::on\_createChat1st(){  auto line = qobject\_cast<QLineEdit\*>(sender());  if(line->text().contains(' ')){  QMessageBox::warning(this, "ACHTUNG!", "Chat name cannot contain with a space");  line->clear();  }  else{  c\_message.request = c\_create\_chat;  c\_message.message\_text = line->text().toStdString();  SendToServer(socket, &c\_message);  }  line->clear();  }  void MainWindow::on\_createChat2nd(server\_data\_t s\_message){  if(s\_message.responce == s\_failure){  QMessageBox::warning(this, "ACHTUNG!", "Could not create this chat");  }  }  void MainWindow::on\_leaveChat1st(){  client\_data\_t c\_message;  c\_message.request = c\_leave\_chat;  c\_message.message\_text = "";  SendToServer(socket, &c\_message);  }  void MainWindow::on\_sendMessage(){  client\_data\_t c\_message;  c\_message.message\_text = qobject\_cast<QLineEdit\*>(sender())->text().toStdString();  c\_message.request = c\_send\_message;  SendToServer(socket, &c\_message);  qobject\_cast<QLineEdit\*>(sender())->clear();  } |

Приложение Е. Файлы потока чтения клиентского приложения

readroutine.h – определение класса для неблокирующего чтения в клиентском приложении:

|  |
| --- |
| #pragma once  #include <QThread>  #include <QTcpSocket>  #include <QDebug>  #include <QObject>  #include <memory>  #include "client\_interface.h"  enum client\_state\_t{  no\_name = 0,  no\_chat,  in\_chat  };  class ReadRoutine : public QObject {  Q\_OBJECT  enum unblocked\_read\_state {  REQ\_TYPE,  RESP\_TYPE,  LENGTH,  MESSAGE  };  server\_data\_t s\_d;  QTcpSocket\* tcp\_socket;  size\_t off = 0;  unblocked\_read\_state state = REQ\_TYPE;  size\_t length;  bool readed = false;  public:  std::atomic\_bool go\_on;  std::atomic\_int client\_state;  ReadRoutine(QTcpSocket\*);  public slots:  bool UnblockedReadFromServer(); /\* создает и запускает \*/  void stop(); /\* останавливает \*/  signals:  void finished(); /\* сигнал о завершении работы \*/  void new\_message\_come(std::string&);  void get\_chats(server\_data\_t);  void connect\_chat(server\_data\_t);  void create\_chat(server\_data\_t);  void send\_name(server\_data\_t);  void leave\_chat(server\_data\_t);  };  class ReadRoutineHandler : public QObject {  Q\_OBJECT  QTcpSocket\* tcp\_socket;  public:  ReadRoutineHandler(QObject \*parent, QTcpSocket\*);  ~ReadRoutineHandler();  void startThread();  ReadRoutine\* chat\_proc;  public slots:  void stopThread();  signals:  void stop(); //остановка потока  }; |

readroutine.cpp – определение методов класса из readroutine.h в клиентском приложении:

|  |
| --- |
| #include "readroutine.h"  #include "client\_interface.h"  #include "mainwindow.h"  #include <QDebug>  ReadRoutine::ReadRoutine(QTcpSocket\* socket) {  tcp\_socket = socket;  go\_on = true;  length = sizeof(server\_responce\_e);  }  bool ReadRoutine::UnblockedReadFromServer(){  int received = 1;  static size\_t help\_length;  while (received > 0) {  switch (state) {  case REQ\_TYPE:  received = tcp\_socket->read((char\*)&s\_d.request, length);  if (received < 0) return false;  off += received;  if (off == length) {  state = RESP\_TYPE;  length = sizeof(server\_responce\_e);  off = 0;  }  break;  case RESP\_TYPE:  received = tcp\_socket->read((char\*)&s\_d.responce, length);  if (received < 0) return false;  off += received;  if (off == length) {  state = LENGTH;  length = sizeof(size\_t);  off = 0;  }  break;  case LENGTH:  received = tcp\_socket->read((char\*)&help\_length, length);  if (received < 0) return false;  off += received;  if (off == length) {  length = help\_length;  state = MESSAGE;  s\_d.message\_text.resize(length);  off = 0;  }  break;  case MESSAGE:  received = tcp\_socket->read((char\*)s\_d.message\_text.data(), length);  if (received < 0) return false;  off += received;  if (off == length) {  state = REQ\_TYPE;  length = sizeof(server\_responce\_e);  off = 0;  readed = true;  qDebug() << "Client state: " << client\_state;  switch(client\_state){  case no\_name:  switch(s\_d.request){  case c\_set\_name:  client\_state = no\_chat;  emit send\_name(s\_d);  break;  default:  qDebug() << "Chat routine received wrong data";  }  break;  case no\_chat:  switch(s\_d.request){  case c\_get\_chats:  emit get\_chats(s\_d);  break;  case c\_create\_chat:  emit create\_chat(s\_d);  break;  case c\_connect\_chat:  if(s\_d.responce == s\_success) client\_state = in\_chat;  emit connect\_chat(s\_d);  break;  default:  qDebug() << "Chat routine received wrong data";  }  break;  case in\_chat:  switch(s\_d.request){  case c\_receive\_message:  emit new\_message\_come(s\_d.message\_text);  break;  case c\_get\_chats:  client\_state = no\_chat;  emit get\_chats(s\_d);  break;  default:  qDebug() << "Chat routine received wrong data";  }  }  }  }  }  return true;  }  void ReadRoutine::stop() {  emit finished();  }  ReadRoutineHandler::ReadRoutineHandler(QObject \*parent, QTcpSocket\* tcp\_socket) {  setParent(parent);  this->tcp\_socket = tcp\_socket;  }  void ReadRoutineHandler::startThread(){  chat\_proc = new ReadRoutine(tcp\_socket);  chat\_proc->client\_state = no\_name;  QThread\* thread = new QThread;  chat\_proc->moveToThread(thread);  //signals to connect two threads  connect(chat\_proc, &ReadRoutine::new\_message\_come, qobject\_cast<MainWindow\*>(parent()), &MainWindow::on\_newMessage);  connect(chat\_proc, &ReadRoutine::get\_chats, qobject\_cast<MainWindow\*>(parent()), &MainWindow::list\_of\_chats);  connect(chat\_proc, &ReadRoutine::create\_chat, qobject\_cast<MainWindow\*>(parent()), &MainWindow::on\_createChat2nd);  connect(chat\_proc, &ReadRoutine::connect\_chat, qobject\_cast<MainWindow\*>(parent()), &MainWindow::on\_chatButton\_Clicked2nd);  connect(chat\_proc, &ReadRoutine::send\_name, qobject\_cast<MainWindow\*>(parent()), &MainWindow::on\_nameLine\_returnPressed2nd);  //correct memory management if thread is stopped and so on  connect(chat\_proc, &ReadRoutine::finished, thread, &QThread::quit);  connect(qobject\_cast<QIODevice\*>(tcp\_socket), &QIODevice::readyRead, chat\_proc, &ReadRoutine::UnblockedReadFromServer);  connect(this, &ReadRoutineHandler::stop, chat\_proc, &ReadRoutine::stop);  connect(chat\_proc, &ReadRoutine::finished, chat\_proc, &ReadRoutine::deleteLater);  connect(thread, &QThread::finished, thread, &QThread::deleteLater);  thread->start();  }  void ReadRoutineHandler::stopThread() {  emit stop();  }  ReadRoutineHandler::~ReadRoutineHandler(){} |