

УДК 004.9

ВЫСОКОНАГРУЖЕННОЕ ПРИЛОЖЕНИЕ «СОЦИАЛЬНАЯ СЕТЬ ДЛЯ ПОИСКА РАБОТЫ» С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ БОЛЬШИХ ДАННЫХ

Турок К.А. ¹, студент гр.356241

Белорусский государственный университет информатики и радиоэлектроники¹,
г. Минск, Республика Беларусь

Хмелева А.В. – канд.техн.наук, доцент, доцент кафедры ПОИТ

Аннотация. Поиск работы зачастую является сложным процессом. Помимо получения списка вакансий, соискателям необходимо ознакомиться с предполагаемым местом работы и изучить информацию о возможном работодателе. Удобное представление информации о вакансиях будет полезно многим пользователям данного веб-приложения.

Ключевые слова: микросервисная архитектура, большие данные, социальная сеть.

Введение. С развитием информационных технологий, в частности сети Интернет, пользователи все чаще стали искать требуемую информацию в режиме онлайн. Это связано с удобством и простотой данного подхода. Примером могут являться социальные сети. Они позволяют пользователям с лёгкостью обмениваться информацией и разного рода сообщениями. Социальные сети могут использоваться в том числе и для поиска работы.

В данной статье автором показано, что для создания социальной сети, соответствующей требованиям современной веб-системы, необходимо применение технологий больших данных для успешной обработки огромного числа запросов и хранения большого количества информации.

Основная часть.

Целью создания социальной сети для поиска работы является предоставление эффективного и удобного инструмента для соискателей и работодателей, который позволит им быстрее и проще находить подходящие вакансии и кандидатов соответственно. Это позволит сократить время на поиск работы и найти лучших кандидатов для работодателей. Кроме того, целью создания социальной сети для поиска работы является улучшение качества подбора кадров и расширение возможностей для соискателей и работодателей. Благодаря этой социальной сети, соискатели смогут быстрее находить вакансии, которые соответствуют их специальности и предпочтениям, а работодатели смогут находить кандидатов с нужным опытом и квалификацией.

Объектами исследования для разработки являются социальные сети для поиска работы.

Предметом исследования является структура и функциональность разрабатываемого веб-приложения, а также технологии, используемые для его проектирования и разработки.

Главной задачей программного средства является возможность создания пользователями социальной сети резюме. Добавление информации о себе как о соискателе работы позволяет многим организациям, использующим данное программное средство, с большей вероятностью заметить резюме. Помимо поиска будущих работников, организациям предоставляется возможность размещения в веб-приложении информации о своих вакансиях, что позволяет большему числу пользователей увидеть данную запись.

Веб-приложение построено на клиент-серверной архитектуре. Клиентская часть представляет собой одностраничное приложение, написанное на языке Javascript при помощи библиотеки React. Данный подход в построении клиентской части был выбран в связи с распространённостью и удобством одностраничных приложений: они уменьшают время загрузки страниц и упрощают способ общения с серверной частью. Серверная же часть написана с использованием фреймворка ASP.NET [1] на технологии .NET. Была использована микросервисная архитектура [1]. Данный подход был выбран в связи с необходимостью создания высоконагруженного приложения. Именно микросервисная архитектура позволяет при помощи разделения по доменам создавать независимые части большого веб-приложения. Схема сервисов данного приложения представлена на рисунке 1.

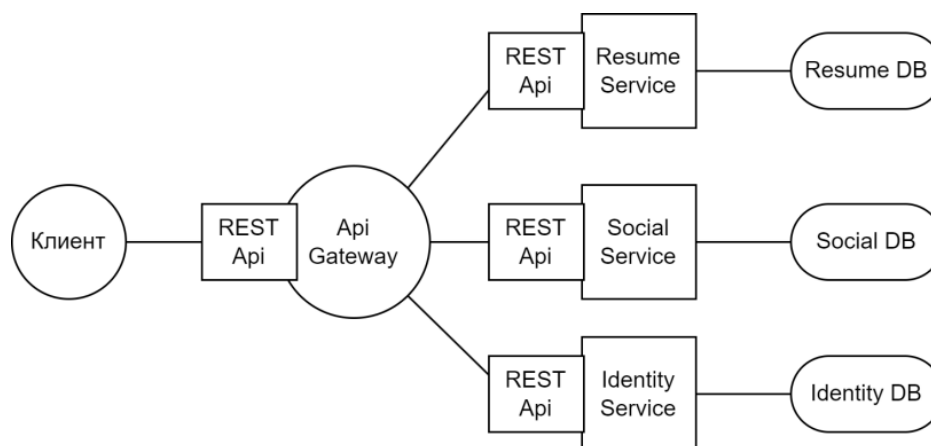


Рисунок 1 – Схема сервисов серверной части приложения

В соответствии с целью исследования необходимо создать приложение, которое будет соответствовать критериям качественного приложения, способного обрабатывать большое количество информации и запросов. К важным факторам также необходимо отнести и работу с данными. Ведь хранение и быстрая обработка являются неотъемлемой частью нормального функционирования любого приложения.

Создание социальной сети с использованием больших данных (Big Data) включает в себя обработку и анализ обширных объемов информации, которая генерируется и потребляется в социальных медиа [2]. Вот несколько аспектов, которые могут быть важными при разработке социальной сети с использованием больших данных.

Хранение и управление данными.

Хранение и управление данными являются ключевыми аспектами при создании социальной сети, особенно если речь идет о больших объемах информации и потоков данных. Необходимо определить типы данных, которые будут храниться в социальной сети. Это может включать в себя профили пользователей, сообщения, изображения, видео, комментарии, лайки и другую информацию.

Выбор подходящей базы данных зависит от требований приложения. В случае больших данных, NoSQL базы данных (MongoDB, Cassandra) могут быть предпочтительными, так как они предоставляют гибкость в хранении неструктурированных данных и способны масштабироваться горизонтально. Но, помимо этого, для рассматриваемого приложения необходимо рассмотреть и графовые базы данных. Ведь именно они позволят не только сохранять сущности, но и отображать связи между ними.

Графовые базы данных представляют собой категорию NoSQL баз данных, специально разработанных для хранения и обработки данных в виде графовых структур [3]. Графовая модель данных использует узлы (вершины) и ребра (связи) для представления отношений между объектами.

Графовые базы данных (ГБД) широко применяются в социальных сетях для эффективного моделирования и анализа социальных графов. Одним из примеров использования графовых баз данных в социальной сети является поиск и предоставление рекомендаций для связей между пользователями.

Рассмотрим использование графовой базы данных в социальной сети для поиска работы. Пусть каждый пользователь представляется вершиной графа, а связи между пользователями (дружба, подписка и т. д.) будут представлены ребрами. С использованием запросов к графовой базе данных можно анализировать дружеские связи пользователя, определять их степень близости и взаимодействия.

У графовых баз данных есть преимущества при использовании в данном контексте. Модель графа естественным образом отражает связи в социальной сети, что упрощает анализ и поиск взаимосвязей. Графовые базы данных предоставляют возможности для применения алгоритмов машинного обучения и аналитики, что дополнительно улучшает качество рекомендаций.

Графовые базы данных обеспечивают эффективные алгоритмы поиска и трансверсинга графов, что делает быстрым анализ связей между пользователями.

Таким образом, графовые базы данных являются мощным инструментом для обработки и анализа данных с учетом сложных отношений, что делает их особенно полезными в контексте социальных сетей и других приложений, где важны связи между объектами.

Масштабируемость

Масштабируемость в социальной сети является критически важным аспектом, поскольку популярные социальные платформы часто сталкиваются с резким ростом количества пользователей и объема данных. Проектирование системы с учетом масштабируемости позволяет обеспечить стабильную и эффективную работу при росте пользовательской базы и активности.

Данное приложение будет разворачиваться с использованием горизонтального масштабирования. Горизонтальное масштабирование представляет собой распределение нагрузки между несколькими серверами. Помимо этого, предполагается и разделение содержимого баз данных.

В соответствии с типом применяемой базы данных будут применяться различные подходы для масштабирования. Используемые нереляционные базы данных будут применять подход шардинга.

Шардинг — это стратегия горизонтального масштабирования базы данных, при которой данные разделяются на части, называемые шардами, которые затем распределяются по различным серверам или узлам. Эта техника используется для обработки роста объемов данных и повышения производительности путем распределения нагрузки.

Данные в базе данных делятся на логические части, называемые шардами, в соответствии с каким-то критерием, например, хэшем от ключа данных.

Каждый шард затем размещается на отдельном сервере или узле в распределенной инфраструктуре. Это может быть физически разные серверы или разные экземпляры базы данных на одном сервере. Когда запрос поступает, система использует ключ для определения, в каком шарде находится нужная информация. Затем запрос направляется только к этому конкретному шарду, минимизируя взаимодействие с другими шардами. После обработки запроса каждым шардом результаты объединяются и возвращаются пользователю, предоставляя полное представление данных.

Для реляционных баз данных применяется подход репликации.

Репликация – это процесс создания и поддержания копий данных (реплик) в разных местах или на различных серверах. Эта стратегия используется для повышения производительности, обеспечения отказоустойчивости и улучшения общей доступности данных.

Одна база данных (или один узел) служит в качестве основной реплики, где происходят изменения данных. Эта реплика известна как "мастер" или "первичная" реплика.

Одна или несколько дополнительных копий данных создаются на других серверах или узлах. Эти копии называются "репликами" или "вторичными" репликами. Изменения, которые происходят на основной реплике (мастере), регулярно передаются на вторичные реплики. Это может осуществляться с использованием различных методов передачи данных, таких как журналы транзакций или копирование данных. Вторичные реплики могут использоваться для чтения данных, а иногда и для записи (если система поддерживает запись на реплики). Это может улучшить производительность при обработке большого объема запросов на чтение.

Особого внимания в разрабатываемом приложении требует обработка медиа-контента. К нему относятся картинки, видео и т.д. При развертывании приложения используется Content Delivery Network.

Content Delivery Network (CDN) — это распределенная сеть серверов, которая используется для эффективной доставки контента, такого как веб-страницы, изображения, видео, аудио и другие файлы, до конечных пользователей на основе их местоположения. Основной целью CDN является улучшение производительности, уменьшение задержек при загрузке контента и обеспечение более высокой доступности ресурсов.

CDN включает в себя множество серверов, которые физически распределены по всему миру. Эти серверы называются узлами CDN или кэширующими серверами. Когда пользователь запрашивает определенный контент, например, изображение с веб-сайта, CDN кэширует (сохраняет) копию этого контента на ближайшем к пользователю узле CDN.

Кэшированные копии контента распределены по всем узлам CDN, что позволяет обеспечить близость контента к конечному пользователю и снижение времени отклика. При следующем запросе на тот же контент CDN автоматически выбирает ближайший узел, чтобы ускорить доставку.

Заключение

Таким образом, созданное приложение удовлетворяет критериям современной социальной сети, для которой нужен особый подход при разработке и развертывании. Рассмотренные выше методы полезны не только для данного типа социальной сети: поиск работы. Данные подходы применимы для многих типов современных веб-приложений.

Список использованных источников:

1. *Микросервисная архитектура [Электронный ресурс] – Режим доступа: microservices.io.*
2. *Высоконагруженные приложения. Программирование, масштабирование, поддержка. – СПб.: Питер, 2019. – 640 с.*
3. *Графовые базы данных [Электронный ресурс] – Режим доступа: <https://aws.amazon.com/ru/nosql/graph/>*

UDC 004.9

HIGHLY LOADED APPLICATION «SOCIAL NETWORK FOR JOB SEARCH» USING BIG DATA TECHNOLOGIES

Turok K. A.¹

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Hmeleva A. V. – PhD, assistant professor, associate professor of the department of SoIT

Annotation. Finding a job is often a difficult process. In addition to receiving a list of vacancies, applicants need to familiarize themselves with the proposed place of work and study information about the possible employer. Convenient presentation of information about vacancies will be useful to many users of this web application.

Keywords: microservice architecture, big data, social network.