DEMEDIA – DECENTRALIZED SOCIAL MEDIA PROTOCOL

Bandara A. M. C. A.

IT120159726

B.Sc. Special (Honors) Degree in Information Technology Specialization in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

September 2023

USER DATA DECENTRALIZATION PROTOCOL

Bandara A. M. C. A.

IT120159726

The dissertation was submitted in partial fulfillment of the requirements for the B.Sc. Special Honors degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

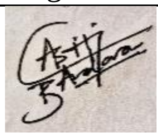
Sri Lanka

September 2023

DECLARATION

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
Bandara A. M. C. A.	IT20159726	

The above candidate has carried out research for the bachelor's degree Dissertation under my supervision.

…………………………..   …………………………

Signature of the supervisor: Date

(Mr. Kavinga Yapa Abeywardena)

ACKNOWLEDGEMENT

The success of this research is due to the efforts of my supervisor, Mr. Kavinga Yapa Abeywardena, as well as the rest of the team. Additionally, I'd like to express my gratitude to the Department of Information Technology at the Sri Lanka Institute of Information Technology, as well as to CDAP academics and staff, for enabling this research.

ABSTRACT

Social networks have become an essential part in the modern world, with billions of people using them as an integral part of their daily lives. Globalization and other reasons have driven the current trend, resulting in a rapid increase in the usage of these platforms. Despite their huge popularity, these networks are centralized, allowing the companies that own them to monitor and track their users' activities. This has given rise to serious concerns regarding the privacy and security of user data, which is often sold to third parties for profit. Given the massive value of user data in today's and tomorrow's world, many people are concerned about this issue. While decentralized, community-driven applications have emerged as a potential solution to this problem, none have yet been able to compete with centralized social network platforms. Consequently, the purpose of this research was to create a decentralized social network architecture that would incorporate the basic functionalities of a social media platform while ensuring the privacy and security of users' data, which would be stored locally on their machines. The implementation of a user data decentralization protocol is required for this research component, as well as the design and development of a client application that can successfully communicate with other network peers via the decentralized protocol provided.

Keywords: decentralization, social media, data privacy, data security, peer-to-peer

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

IPFS	Interplanetary File System
SDLC	Software Development Life Cycle
DApps	Decentralized Applications
AWS S3	Amazon Simple Storage Service
DeSo	Decentralized Social Network

1.0 INTRODUCTION

A large number of people currently use social media, and these platforms have integrated themselves into modern lifestyles. Since "Six Degrees," the first recognized social network, opened in 1997 and allowed users to create profiles and connect with one another, social networks have gained popularity on a worldwide level. There are many social networks available today, and their usage is expanding [1]. Examples include Facebook, WhatsApp, Instagram, Twitter, and more. These social networks are growing in popularity because of their different uses and ways in which they compete with one another.

The reasons for the increasing popularity of social networks must be examined. The most common explanation is that users want to keep connected with others and informed about global events. Physical interaction with others can become difficult due to tight schedules and workloads, so users communicate via the Internet, which proves more practical. Furthermore, social media is used to share photographs and videos for entertainment as well as to express thoughts and ideas. In addition, social media platforms are used for product research, forcing businesses to focus on online marketing and target audience engagement. However, the centralization of these social networks has resulted in several issues that have recently come to attention.

According to an article [2] MIT Sloan professor Sinan Aral said, "Social media is rewiring humanity's central nervous system in real time" at the social media Summit@MIT, which brought professionals together to explore these challenges and focus on solutions that ranged from new oversight committees to breaking up major businesses. Furthermore, key concerns in today's social media networks were mentioned. Some of the current issues impacting social media include the spread of false news and misinformation, the challenging balance between user privacy and platform transparency, and the absence of regulation for social media businesses.

Web3, often talked about as the future of the Internet, is a more representative and decentralized version of what we have now. It's a word that's been floating around the internet, but many people are still confused about what it actually means. Gavin Wood, one of the creative professionals of Ethereum and the founder of Parity Technologies, introduced the term Web3 in 2014 when he expressed his vision for the future of the Internet. To understand Web3, you must first understand how it differs from Web 2.0. Developers in Web 2.0 design apps that rely on a single central server or database, which is generally handled by a single cloud provider. However, with Web3, developers can develop apps that run on a decentralized network known as the blockchain, which consists of multiple peer-to-peer node servers. DApps, or decentralized apps, are at the heart of Web3. [3]

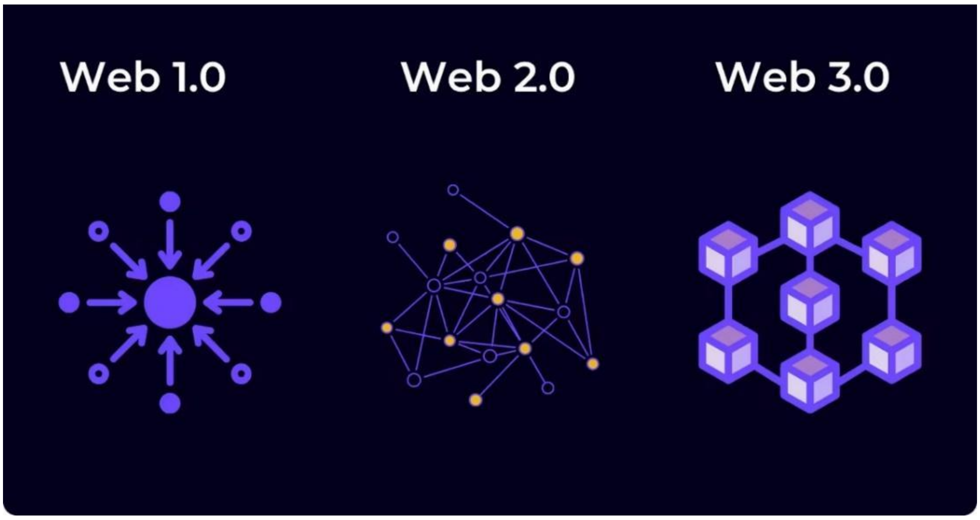


Figure 1.0.1: What is Web 3.0?

Blockchain technology is essential for Web3 because it is the same technology that enables cryptocurrencies such as Bitcoin. The idea of decentralization took hold when developers saw how the blockchain was transforming Bitcoin. Unlike traditional currencies, Bitcoin isn't controlled by a single company or a central authority like a bank; it's decentralized, thanks to blockchain. This trustless approach is the foundation of Web3. We currently use large tech companies to deliver on their promises, while Web3 products and services are decentralized and based on blockchains. This implies that you must trust the underlying algorithm rather than a single person or thing. As a result, Web3 is all about decentralization, and it represents a significant change in how we think about and use the Internet. [4]

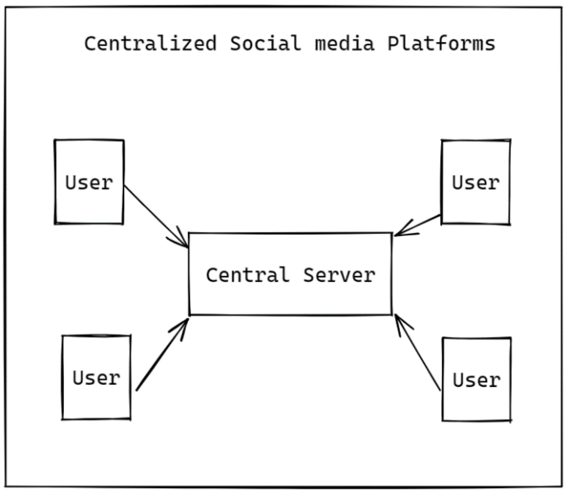


Figure 1.0.2: High-level architecture of centralized social media network

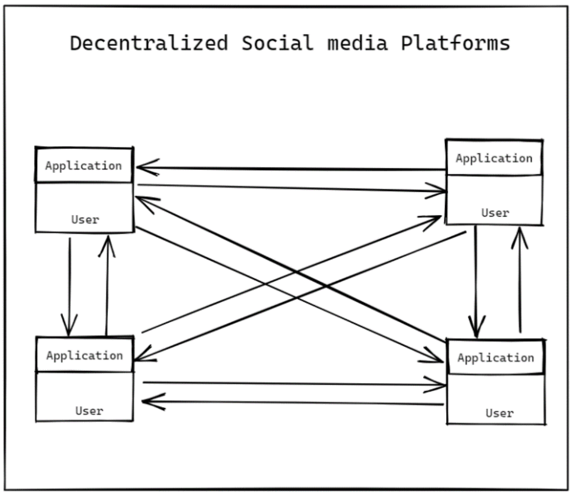


Figure 1.0.3: High-Level architecture of decentralized social media network

As mentioned in [5] the growth of Web 3.0 is expected to continue as users increasingly recognize the advantages that provides over traditional social networking websites. While looking for the best web3 social media platforms, keep in mind a variety of standards such as traditional social media platforms, social media users, decentralized social networks, internet freedom, free expression, and online networks, among others. Mastodon [6] and DeSo [7] are two of the most significant decentralized social networking platforms currently in development; according to the article [8].

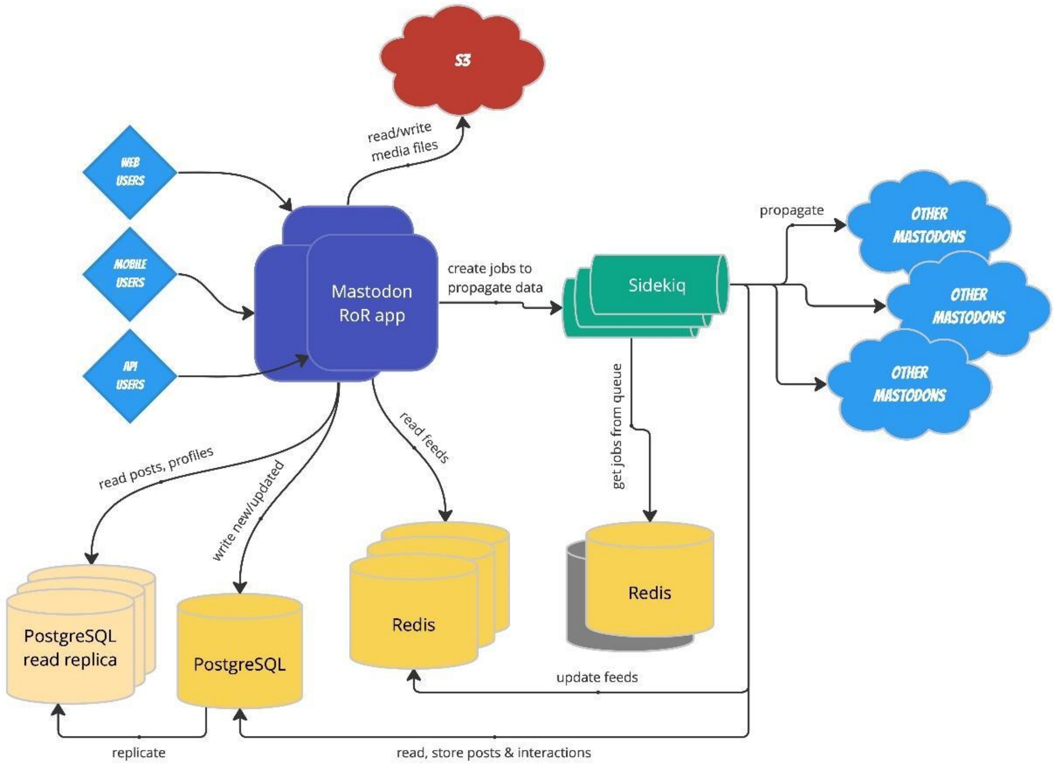


Figure 1.4: Mastodon Architecture

The main goal of these social networks is to overcome the problems associated with existing social media platforms by allowing monopolistic control exercised by important social media corporations and granting authority to the platform's actual users. Despite the current increase of decentralized social media platforms, several restrictions and obstacles must be addressed. Most of these systems use blockchain technology to create decentralization, which, as mentioned in [9], could result in increased processing speed charges.

The primary goal of this research is to develop a decentralized social network that uses the user's device as storage while ensuring privacy and authenticity, with the user having complete control over their data. This research component focuses on the development of a user data decentralization protocol, as well as the development and implementation of a client application capable of smooth connection with other network peers via the decentralized protocol.

Numerous researchers are currently engaged in discussions regarding the implementation of decentralized social media platforms. Because of the overwhelming control maintained by social media companies and the significant increase in the number of social media users, this problem has received a lot of attention. According to reference [10], a Global Web Index survey done in July 2020 revealed a 10.5% rise in social media usage compared to July 2019. Furthermore, as noted in reference [11], the global number of people using social media has grown to 4.48 billion, more than double the 2.07 billion figures recorded in 2015.

1.1 Background Literature

Numerous researchers are currently engaged in discussions regarding the implementation of decentralized social media platforms. Because of the overwhelming control maintained by social media companies and the significant increase in the number of social media users, this problem has received a lot of attention. According to reference [10], a GlobalWebIndex survey done in July 2020 revealed a 10.5% rise in social media usage compared to July 2019. Furthermore, as noted in reference [11] the global number of people using social media has grown to 4.48 billion, more than double the 2.07 billion figures recorded in 2015.

The most popular systems were created on centralized servers which store data and information for users, according to research [12]. Such systems are vulnerable to a number of threats because of their centralized architecture. Data can easily be managed, sold, or stolen if the owner is not actively in charge of it. The increase in social media issues has been combined with blockchain technology. The Blockchain idea was first introduced in 2008 when Bitcoin, the first cryptocurrency without a centralized authority, adopted it as a key component. Previous research has shown that this technology has the ability to challenge existing financial procedures and change the nature of investment.

Web 3.0 is predicted to be the web's future in [13], as noticed in the conclusion with technological developments. However, some people are doubtful that a decentralized Web 3.0 framework can replace the centralized Web 2.0 design. The goal of this research was to develop a decentralized social network architecture based on the Ethereum platform that can enable greater privacy, data ownership, and community-driven services. However, because Ethereum and other blockchains have well-known limits, the platform can be improved in the future. Massive organizations like Facebook, Google, and Microsoft are also working on and testing potential decentralized computing solutions.

In [8], the following are some of the disadvantages of blockchains: One is the low transaction-processing pace of blockchain systems. While other payment systems like

VISA can process hundreds of transactions per second, the Ethereum network can only process about 15 transactions per second. As the number of connections grows, this problem becomes more challenging to solve. Another issue with blockchain networks is scalability.

On-chain storage can result in significant cost increases. [14] The yellow paper for Ethereum states that keeping a 1 GB file on the Ethereum blockchain costs around $316,416 USD. Chain storage costs several orders a thousand times more than centralized storage (for example, Amazon S3 HOT storage costs $0.023GB per month). Second, storing huge amounts of data on-chain not only hinders the synchronization of new blockchain nodes but also consumes a significant amount of transaction throughput.

In addition to reviewing research papers, current decentralized social media platforms were examined to discover features that had previously been implemented as well as improvements that needed to be created.

1.1.1 Comparison of Existing Decentralized Social Media Platforms

1.1.1.1 Mastodon

Figure 1.1.1.1.1: Mastodon Logo

Mastodon is an open-source platform that has resulted in a succession of alternative social media sites. The social networking sites are hosted on decentralized servers, signifying that they are not owned by a single person or organization. The free and open-source software project was launched in 2016, with a stable release scheduled for November 2022. [15]

Mastodon uses a decentralized architecture concept that makes use of a hosted server to provide decentralized data storage. Any user can self-host a Mastodon server. As noticed in [16], Mastodon needs scalability, which will be required when the platform is used by a large number of users. Furthermore, Mastodon specifically offers mobile app clients to their clients. Mastodon does not support caching or control of the caching mechanism. [17]

1.1.1.2 Deso

Figure 1.1.1.2.1: Deso Logo

DeSo is a relatively new blockchain that was designed from the ground up to decentralize social media. The idea is to give more power to creators, consumers, and developers rather than the large tech giants that control the majority of what we see on the internet today. [18] Nader Al-Naji launched DeSo in 2019, and the DeSo Foundation promotes blockchain technology. The Octane Fund, a $50 million developer fund, was announced for 2021. One of the most notable benefits of DeSo is that it is not a general-purpose blockchain, which means that it offers a limited range of social- oriented features, making it highly scalable and cost-effective for on-chain storage. [19] DeSo also follows a decentralized social media platform which uses a blockchain built itself. Because of this DeSo can only be self-hosted partially as it has to use the DeSo blockchain.

1.2 Research Gap

There has been a lot of research done recently on decentralized social media systems. The majority of these studies utilize blockchain to achieve decentralized architecture within the social media network. Centralized social media platforms have long governed the current digital environment, but they are not without faults. An increase in research and development with the goal of decentralizing social media networks has been triggered by concerns about data privacy, censorship, and control over generated by users' material. Blockchain technology, which has become an effective mechanism for changing the design of social media platforms, is an essential player in this movement.

Blockchain can be used to create decentralization, however maintaining massive amounts of data will be challenging. There is currently research being done to address this problem. It is becoming increasingly clear that blockchain technology is essential to changing the environment as demand for decentralized social media platforms increases. An important finding is that a significant number of these decentralized social media platforms use blockchain technology to take advantage of the many benefits of decentralization.

	Mastodon [17]	**DeSo** [19]	**DeMedia**
User Data Storage	In hosted server	In DeSo blockchain	In user's device
Support decentralized data storage	No	No	Yes
Self-hosting	Yes	Partially	Yes
Scalability	No	Partially	Yes
Caching	No	No	Yes
Installable client	Only mobile client	No	Yes

Table 1.2.1.2.1: Comparison of existing decentralized social media platforms

1.3 Research Problem

To address the challenges of existing social media platforms, a social media platform without centralized authority that is governed by the users of the platform must be established. Despite the fact that several decentralized social media systems have been established to address these difficulties, some issues remain unresolved. The main disadvantages include that maintaining a large quantity of data on a blockchain will be an expensive and ineffective procedure.

The identified challenges that arise in the context of a user data decentralization protocol for social media provide multiple challenges that must be addressed in order for such a protocol to be successfully implemented.

- Complete ownership of user data is a key challenge in decentralized social media. Users need to have full control and ownership over their data to ensure privacy and security. Therefore, measures must be taken to provide users with complete ownership of their data, including the ability to manage, access, and delete their data as they see fit. This requires a clear understanding of how user data is collected, processed, and stored and implementing appropriate measures to ensure that users retain control of their data.
- The limitations of blockchain as a storage solution for large quantities of data faces a significant challenge. While blockchain technology is a promising solution for decentralized data storage, its limitations in terms of scalability and efficiency when it comes to storing large amounts of data need to be addressed. Therefore, potential solutions need to be explored to overcome these limitations and ensure that the user data is stored securely and efficiently.
- Enhancements to current decentralized social media networks need to be made to provide users with a seamless and efficient experience. This includes addressing the limitations of existing decentralized social media platforms, such as slow transaction times, low throughput, and limited scalability. Enhancements such as improving the speed of data transfer, reducing transaction costs, and improving the overall user experience will help ensure that users have a positive experience with decentralized social media networks.

Addressing these challenges will require a comprehensive understanding of the technical, regulatory, and social aspects of user data decentralization protocols for social media. Therefore, the development of a user data decentralization protocol that addresses these challenges is crucial to enable the successful implementation of a decentralized social media network.

1.4 Research Objectives

1.4.1 Main Objective

The main objective of this research is to create a new decentralized social media protocol that offers users a more secure, private, and decentralized alternative to traditional social media platforms. The proposed protocol is designed to store user data on a user's device rather than on a centralized server, ensuring that the user retains complete control over their data.

To achieve this objective, this research will involve the development of a user data decentralization protocol that outlines how the user data will be stored, accessed, and shared among peers on the network. This protocol will ensure that user data is encrypted and secured while maintaining its accessibility and availability to the user.

In addition to developing the user data decentralization protocol, this research will also involve the design and implementation of a client application capable of smoothly communicating with other network peers via a decentralized protocol.

The client application will be designed to provide a user-friendly interface that allows users to easily create and share content with others on the network. It also includes features such as encryption, authentication, and verification mechanisms to ensure that all user data and communications are secure and private.

Overall, the main objective of this research is to provide users with a decentralized social media platform that offers them complete control over their data, while also providing a smooth and user-friendly experience. By developing a user data decentralization protocol and a client application that utilizes this protocol, the proposed platform offers a more secure, private, and decentralized alternative to traditional social media platforms.

1.4.2 Sub Objectives

- Examine the opportunities and challenges that user data decentralization technologies present:

This sub-objective involves evaluating the advantages and disadvantages of the user data decentralization protocols. The aim is to identify the strengths and weaknesses of the existing protocols to determine their feasibility and effectiveness in addressing the problem at hand. For instance, the existing protocols' ability to provide security and privacy for user data and compare it with the required standards will be analyzed. Scalability, interoperability, and ease of implementation of the existing protocols will also be assessed. The outcome of this objective will be used to inform the decision-making process regarding the selection of a suitable protocol for implementation.

- Determine the potential improvements and additional features that could be integrated into a protocol for storing user data on the user's device:

This sub-objective focuses on identifying the possible modifications or improvements that can be integrated into an existing user data decentralization protocol to enhance its functionality. The current limitations and requirements for storing user data within the user's device and identifying possible enhancements to address these limitations will be analyzed under this objective. The necessity for a more effective security system for protecting user data will be analyzed, and the implementation of backup mechanisms may be suggested to ensure data availability. This sub-objective will produce a list of possible modifications to the current protocol that will be considered during the phase of implementation.

- Implement a protocol that implements the required changes while following to best practices and includes the following functionalities:

This sub-objective involves the actual implementation of the protocol with the identified modifications. The implementation of a protocol follows best practices in order to ensure its compatibility, scalability, and security. The functionalities identified, such as enabling the user to control user data, utilizing user devices to store user data, improving limitations in current decentralized social media platforms, and making user data highly available, will be incorporated into the implementation. The researcher will also test the protocol to ensure it is working as intended.

- Enable the user to control user data:

This functionality involves giving the user control over their data, including the ability to add, modify, or delete data. The user will also have the option to restrict access to their data or grant access to specific users or applications.

- Utilizing user devices to store user data:

This functionality involves storing user data on the user's device instead of centralized servers. This approach enhances data privacy and security, and it also reduces the risk of data breaches.

Improving limitations in current decentralized social media platforms.

This functionality involves identifying and addressing the limitations in current decentralized social media platforms, such as slow transaction times, lack of scalability, and high energy consumption.

- Making user data highly available:

This functionality involves ensuring that user data is highly available by implementing a backup mechanism to prevent data loss or corruption. The protocol will also ensure that data is easily accessible to the user, even in the absence of an internet connection. Overall, the sub-objectives mentioned above will contribute to the successful implementation of a protocol for storing user data within the user's device, while ensuring security, privacy, and availability.

2.0 METHODOLOGY

2.1 Requirement gathering

The first stage of this research component is to gather requirements for the new data decentralization protocol as well as the new client that will be created to use the protocol and connect to a peer-to-peer network. The requirements are derived from current peer-to-peer apps and peer-to-peer protocols. And also, previous research on existing data decentralization techniques and client programs for connecting to a peer-to-peer network will be reviewed.

The analysis is performed to identify the challenges in present research and the improvements that need to be made. The research criteria are usually gathered by a review of previously published research articles in the topic area. In addition, I searched several blogs, papers, and industry professionals in this subject to learn more about related works. The primary goal of requirement gathering is to define and comprehend the study issue, as well as previously used methodologies, approaches, procedures, and algorithms.

2.1.1 Past Paper Analysis

Analyzing research articles will help with data collection. Analyzing previously published research articles allows for an analysis of current solutions as well as the advantages and disadvantages of past methodologies. Analyzing previous research is useful for selecting the research procedure. Peer-to-peer network, data integrity, caching mechanism and other specialized topics were all examined in the research publications.

2.1.2 Refer Official Documentation

The official documentation is an excellent resource to keep up to date on the technologies that we are interested in using in our system development. Unlike prior research publications, which may contain outdated information, official documents provide immediate and accurate insight into the most recent technology changes, features, and best practices. It is produced and maintained internally by professionals who understand the technology, and it provides thorough protection, troubleshooting tips, and essential safety updates, making it an essential guide for constructing an effective and current system.

2.1.3 Identify Existing Methodologies

This is especially used when analyzing present systems because it allows for identifying research requirements. It also helps in the analysis of the limitations and weaknesses of existing systems and their application to the newly developed system. Existing systems can be found online, and our research will help to design a high-quality solution that is free of current weaknesses and limits.

2.2 Feasibility Study

2.2.1 Technical Feasibility

The initial phase of the project involves the development of a client application with the primary goal of facilitating smooth connections with network nodes and peers. This client application will be the primary interface through which users will interact with the network. Furthermore, it will put emphasis on security by implementing strong mechanisms for securely keeping user data on the user's local device storage. This data storage option not only protects sensitive user information but also improves user control and privacy.

The suggested technological stack gives a strong basis to accomplish these objectives. We can use this system to develop and implement a client-side application that not only properly interfaces with network components but also smoothly incorporates security measures into its architecture. This technological feasibility validates our strategy by allowing us to create a client application that properly corresponds to the project's objectives while maintaining a strong focus on data protection and user experience.

2.2.2 Schedule Feasibility

The proposed solution must be implemented in the allocated seven months, with around two months allocated to gathering and analyzing requirements and the remaining three months to system development. The last two months have been allocated to testing the system since testing is essential to a project's success and during this period it may be further adjusted to provide an improved outcome.

Since the research was carried out using a software-based method, a requirement analysis allowed for the development of a schedule as well as technical and economic feasibility. This facilitated the interaction of research objectives, ensuring that they weren't exceeding technical capabilities or expected timelines.

The main goal of the literature review was to evaluate the research methods and tools used, along with what was discovered from the research that has been done in this field of study. This analysis helped in the research work's decision-making process by assisting with determining the direction the research should take because it gave reliable information. The existing features and possible limitations of the products were determined with the assistance of a review of web resources associated with currently offered monitoring solutions.

2.2.3 Economic Feasibility

The first phase of the project has no budgeted tasks since it focuses on the essential requirement-gathering process, which is accomplished through thorough analysis and a review of previous research and field investigations. The project uses open-source technology and tools effectively, avoiding direct costs throughout this period. However, as we get into the development phase, costs become necessary. This phase requires the use of cloud service providers' services, which are required for infrastructure and application development and provide scalability and flexibility. These costs will be the primary focus of the thesis's next Economic Feasibility analysis, ensuring a thorough assessment of the project's financial consequences.

2.3 Requirement Analysis

The requirement analysis phase was critical to this research project since it allowed for the identification of several components that should be taken into consideration during the research's implementation process.

The data gathered from various sources during the requirement gathering phase was examined during this phase. As a result, essential aspects concerning potential challenges, as well as knowledge of how to proceed and an understanding of the usage of possible devices and technologies, were quickly recognized.

This step also allowed for an assessment of the scope and feasibility of the research. Furthermore, as the research evolved, requirement analysis assisted in identifying current research gaps and provided guidance for identifying the fundamental research problem.

2.4 Software Development Life Cycle (SDLC)

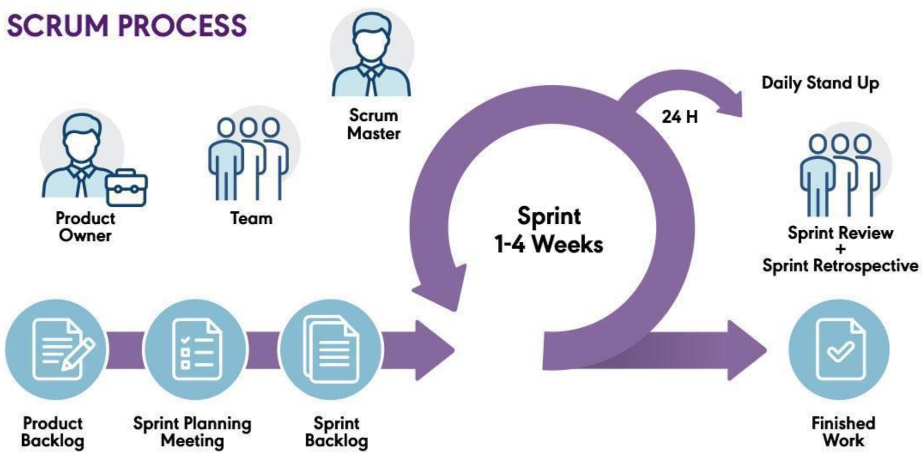


Figure 2.4.1: Software Development Life Cycle

The SCRUM framework, an Agile software development framework, will be used as the primary software project management framework throughout the research. The reason to choose agile methodology over other software project management methodologies such as Lean, Waterfall model, and Six Sigma is because it is best adapted for rapid and effective software development. According to the article [20]SCRUM is a popular agile framework because it defines the systems development process as a loose collection of activities combining the finest tools and techniques a development team can devise to create a system. According to additional information in the same article [20], SCRUM implies that the systems development process is unpredictable, complex and can only be described as an overall progression.

A systematic allocation and organization of the work have been used to achieve the research's outlined objectives and achieve the desired outcomes. A detailed schedule, complete with a Gantt chart, has been made to give each part of the research sufficient time to be finished on time. In addition, the selection of appropriate technologies to effectively implement the proposed solution and demonstrate the intended results of this research has been carefully considered. As evidenced by the detailed preparation and strategic decisions made throughout this research, each step has been taken to ensure a well-structured and systematic approach to achieving the research objectives.

2.5 System Architecture

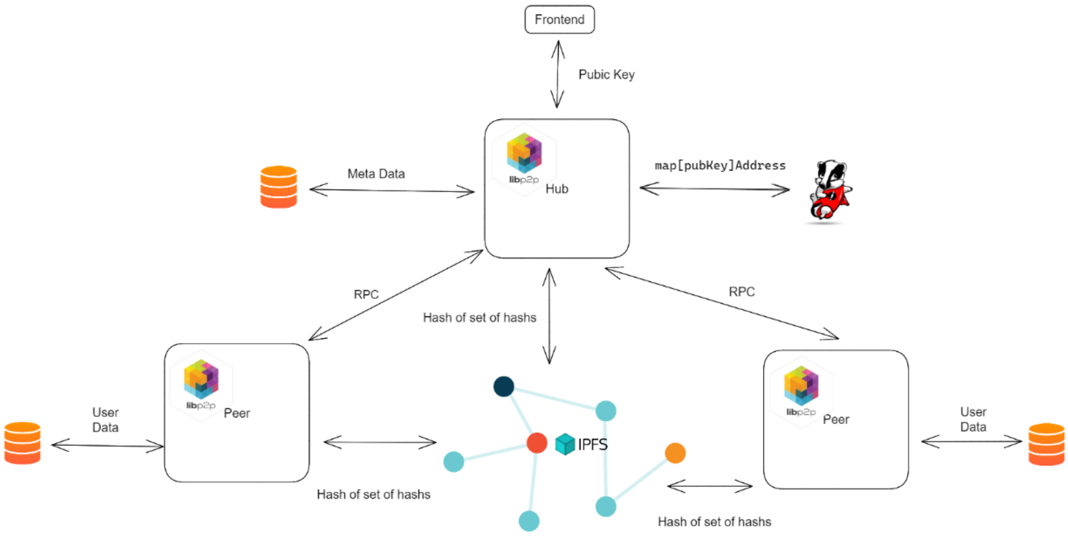


Figure 2.5.0.1:Overall System Architecture

The provided explanation outlines the components of the DeMedia architecture, which is a proposed infrastructure for a decentralized social media platform. The diagram accompanying the explanation provides a high-level overview of the architecture, which includes four major components: data decentralization protocol, peer-to-peer communication, decentralized data storage, and data integrity in a decentralized network. The first component of the DeMedia architecture is the data decentralization protocol, which will enable the platform to store user data within the users' devices rather than on centralized servers. This decentralization of data will enhance user control over their personal data, as well as increase data privacy and security. The second component of the DeMedia architecture is peer-to-peer communication, which will allow users to communicate with each other through a hub which will act only as a communicator. This will increase the speed and efficiency of communication while reducing the dependence on centralized servers. The third component of the DeMedia architecture is decentralized data storage, which will be accomplished using InterPlanetary File System (IPFS). IPFS is a protocol that enables the creation of a decentralized file-sharing network, which is more secure and fault-tolerant than centralized file-sharing networks. The fourth and final component of the DeMedia architecture is data integrity in a decentralized network. This involves ensuring that the data stored on the decentralized network is secure, reliable, and tamper-proof. This is accomplished through the use of cryptography and other security measures to ensure that the data cannot be tampered with or compromised. Overall, the DeMedia architecture aims to create a decentralized social media platform that offers increased privacy, security, and user control over personal data. The infrastructure will consist of hubs, peers, and a decentralized data storage network, which will be illustrated in the high-level architectural diagram.

2.5.1 System Architecture (Individual)

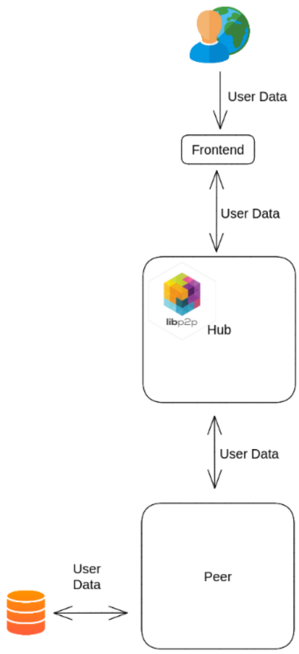


Figure 2.5.1.1: Individual System Architecture

The component mentioned in the explanation refers to the development of a protocol that enables the decentralization of user data. In other words, the focus is on creating a system that allows users to store their data on their own devices instead of centralized servers. This protocol will be accompanied by a client program that will demonstrate its capabilities. One of the main functions of the client program will be to establish communication with hubs, which are nodes on the network that facilitate communication between peers. By connecting to hubs, the client program can then connect to other peers on the network, allowing for the exchange and sharing of data. Another important function of the client program will be to securely store user data on the user's device. This means that the data will be stored within user’s device to ensure that it cannot be accessed by unauthorized parties. By storing data on the user's device, the protocol will minimize the risk of data breaches and increase user control over their own data. Lastly, the client program will display connection information with hubs and peers, allowing users to monitor their connections and manage their data effectively. This will enable users to have greater control over their data and the information they share on the network.

2.6 Commercialization aspects of the product

DeMedia will provide two paid models in addition to its free model: a membership-based subscription model and an advertising-based income decentralized social media platform.

In addition to the free model, DeMedia will also offer two paid models.

- Subscription-based membership model
- Advertising-based revenue model

The host can govern these paid models, allowing them to generate revenue from their platform. Therefore, one could describe DeMedia as a research project focused on commercializing the development of decentralized social media platforms. The project is developing a range of monetization models that hosts can use to generate revenue from their platforms, thereby enabling the commercialization of the technology.

To further elaborate on the commercialization aspect of DeMedia, it's essential to understand that the project is focused on creating technology that can enable the development of decentralized social media platforms. By providing a free base model, DeMedia is making it easier for individuals or organizations to create their own social media platforms that are not controlled by a centralized authority. However, to sustain and grow these platforms, there needs to be a way to generate revenue.

The subscription-based membership model allows hosts to charge users for access to premium features or content on their platform. This revenue can be used to cover the costs of hosting and maintaining the platform.

On the other hand, the advertising-based revenue model enables hosts to generate revenue by displaying ads on their platforms. Hosts can charge advertisers to display their ads on the platform, and this revenue can be used to cover the costs of hosting and maintaining the platform and generate profits.

As a summary of this proposed system, DeMedia is enabling the commercialization of decentralized social media platforms by providing a technology that allows anyone to create their own platform and monetization models that enable hosts to generate revenue and sustain their platforms. This could lead to a more diverse and decentralized social media ecosystem, with a more excellent range of platforms catering to specific sectors and communities.

2.7 Implementation

2.7.1 Background Establishment for Implementation

The team went through a careful planning and decision-making process on the way to putting the decentralized social media protocol into implementation. This phase was thought to be essential for laying a strong basis for the next development and testing phases. In-depth discussions of the decisions made regarding the project's programming language, database system, communication tools, cloud provider, and containerization technologies will be discussed in this section. Each of these choices was crucial in determining the project's ultimate outcome.

2.7.1.1 Programming Language Selection: Go Lang

After a thorough consideration, the decision was made to utilize the Go programming language, commonly known as Go Lang. There were multiple solid factors that led to this decision.

Go Lang is popular for its efficiency, simplicity, and scalability, making it an ideal choice for building decentralized systems. Its concurrency support allows multiple operations to be executed concurrently, enabling the protocol to handle a large number of users and interactions efficiently. Moreover, Go Lang offers a strong standard library and a rich ecosystem of packages, significantly expediting the development process.

The inherent support for robust error handling was considered advantageous, as it enhances the protocol's resilience and fault tolerance. In essence, Go Lang not only met the technical requirements but also aligned with the goal of creating a stable and reliable decentralized social media platform.

2.7.1.2 Database System: PostgreSQL

The selection of an appropriate database system is critical in ensuring the efficient storage and retrieval of data in any project, including the decentralized social media protocol. PostgreSQL was chosen as the database management system, a decision influenced by several key considerations.

First and foremost, PostgreSQL is recognized as a robust open-source relational database system. Its support for complex data types and advanced indexing mechanisms makes it well-suited to handling the diverse data structures prevalent in social media applications. Additionally, strong data integrity and security features offered by PostgreSQL ensure the privacy and reliability of user data.

Furthermore, PostgreSQL's extensibility through user-defined functions and a vibrant community of contributors made it adaptable to evolving project needs. The database's ability to perform efficiently even under high loads was a crucial factor in the decision, given the typically rapid and unpredictable user activity on social media platforms.

2.7.1.3 Collaboration and Version Control: GitHub

Effective collaboration and version control are indispensable for a team working on a complex project like a decentralized social media protocol. To achieve this, GitHub was adopted as the primary platform for code collaboration and integration. The reasons behind this choice were straightforward yet compelling.

GitHub provides an intuitive and user-friendly interface, simplifying the processes of code sharing, reviewing, and merging. This streamlined the development workflow, ensuring that team members could collaborate seamlessly. Its version control capabilities allowed changes to be tracked, conflicts to be managed, and a comprehensive history of the codebase to be maintained, enhancing transparency and accountability.

Moreover, GitHub's robust issue tracking system enabled effective task management and prioritization, particularly valuable in a project of this scale, where numerous features and components needed to be developed and integrated.

2.7.1.4 Cloud Deployment: Amazon Web Services (AWS)

In the modern era of software development, cloud computing has emerged as a game-changer, offering unparalleled scalability, reliability, and flexibility. The benefits of deploying development and production environments in the cloud were recognized, and after careful evaluation, Amazon Web Services (AWS) was selected as the cloud provider.

AWS's vast array of services and global infrastructure ensured that the decentralized social media protocol could scale seamlessly to accommodate increasing user loads. The pay-as-you-go pricing model allowed costs to be optimized while benefiting from world-class infrastructure and support.

Furthermore, AWS offered a range of tools and services for tasks such as server provisioning, load balancing, and auto-scaling. These features simplified the management of cloud infrastructure, freeing up valuable time and resources that could be redirected toward enhancing the protocol's functionality and performance.

2.7.1.4.1 Deployment Using Amazon EC2 Instances

For deploying Docker containers and the PostgreSQL database, an Amazon Elastic Compute Cloud (Amazon EC2) instance was utilized. Amazon EC2 instances offered several advantages, including scalability, flexibility, and ease of configuration. Below are the specifications of the Amazon EC2 instance used for this deployment:

Type	t2.micro
Number of vCPUs	1
RAM (GiB)	1.0
Storage (GiB)	30
Operating System	Ubuntu

Table 2.7.1.4.1.1: Specifications of EC2 instance

2.7.1.5 Containerization: Docker

To achieve consistency in deployment and streamline the process, Docker, a containerization technology, was leveraged. Docker provided several advantages that significantly facilitated the implementation.

Docker's fundamental benefit lies in its ability to encapsulate applications and their dependencies within lightweight containers. Each component of the decentralized social media protocol was integrated and packaged as a Docker image. This approach not only simplified deployment but also ensured that each component could run consistently across various environments.

One of Docker's key strengths is its isolation capabilities, allowing conflicts between different components of the system to be avoided. This isolation reduced the risk of compatibility issues and contributed to the overall stability of the protocol.

Additionally, Docker's portability made it possible to develop and test components independently before seamlessly integrating them into the larger system. This modular approach enhanced development agility and minimized disruptions during the implementation phase.

2.7.1.5.1 Deployment as Docker Containers

The decision to employ Docker containers for deployment resulted in the successful deployment of the implemented demo social platform. This approach offered several tangible benefits:

- **Consistency**: Docker containers ensured that the demo platform ran consistently across different environments, eliminating the notorious "it works on my machine" problem.
- **Scalability**: Docker's scalability features allowed adaptation to varying levels of user demand effortlessly. As user activity increased, the application could be scaled horizontally by adding more containers.
- **Resource Efficiency**: Docker containers, being lightweight and sharing the host OS kernel, resulted in minimal overhead. This translated to efficient resource utilization, reducing infrastructure costs.
- **Quick Deployment**: Docker's quick start-up time enabled the rapid deployment of new updates and features, minimizing downtime and user disruption.

In conclusion, the strategic choices made in setting up the background for implementation were instrumental in ensuring the success of implementation of the decentralized social media protocol. The adoption of Go Lang, PostgreSQL, GitHub, AWS, and Docker laid a strong foundation for the subsequent phases of development and testing. These decisions, rooted in practical considerations and a deep understanding of the requirements unique to the project, enabled the creation of a robust, scalable, and efficient platform that met the goals and expectations.

2.7.2 Setting Up CI/CD Pipeline

Due to the frequent releases during the development process, the team realized the necessity for a CI/CD (Continuous Integration/Continuous Deployment) pipeline. As a result, it was decided to create a reliable CI/CD pipeline that incorporated both CI and CD components. This section describes the pipeline's implementation, which was crucial in accelerating the development and deployment procedures.

2.7.2.1 Continuous Integration (CI) Pipeline

For the CI portion of the pipeline, GitHub Actions was chosen as the preferred tool. This decision was well-founded for several reasons.

GitHub Actions offers a seamless and integrated approach to automating our software development workflows. It allowed us to define, customize, and automate various tasks, such as code compilation, testing, and code quality checks, directly within our GitHub repository.

Moreover, GitHub Actions integrates seamlessly with our codebase hosted on GitHub, making it a natural choice for our CI needs. The ease of configuration and extensive library of pre-built actions simplified the setup of our CI workflow.

2.7.2.2 Continuous Deployment (CD) Pipeline

In parallel with the CI pipeline, a CD pipeline was established using GitHub Actions and Watchtower. Each was chosen for distinct but complementary reasons.

GitHub Actions was extended to serve as our CD tool, ensuring the seamless deployment of our application. GitHub Actions' continuous deployment capabilities allowed us to automate the deployment process, ensuring that every successful CI build was automatically deployed to our production environment. This reduced manual intervention and minimized deployment errors.

Watchtower, on the other hand, played a crucial role in automating container updates. It continuously monitored our Docker containers for new image versions and automatically updated running containers to the latest versions when changes were detected. This ensured that our application was always running with the most up-to-date code, enhancing security and reliability.

Direct integration to the GitHub repository and the user-friendly configuration of

GitHub Actions made them the best option for both CI and CD. It eliminated the need for third-party tools, simplifying our development workflow and ensuring that code changes were rapidly and reliably integrated and deployed.

Watchtower's selection was driven by its ability to automate container updates, reducing the need for manual intervention, and ensuring our application's consistent and secure operation.

The implementation of this CI/CD pipeline provided several benefits including:

- **Automation:** The CI/CD pipeline automated critical aspects of the development and deployment processes, reducing the manual effort required.
- **Speed and Efficiency:** Rapid and automated testing, integration, and deployment improved the overall speed and efficiency of the development cycle.
- **Consistency:** With CI/CD, every code change was subjected to the same automated testing and deployment process, ensuring consistency and reliability.
- **Reduced Errors:** Automation minimized the potential for human errors during deployment, enhancing the overall quality of our application.
- **Frequent Releases:** The CI/CD pipeline facilitated frequent and reliable releases, crucial for the development requiring rapid iterations and updates.

The establishment of the CI/CD pipeline using GitHub Actions and Watchtower significantly enhanced the development and deployment processes, aligning with the unique demands of the project. This automation not only improved efficiency but also contributed to the consistency and quality of our work, enabling to meet project milestones and deliver robust outcomes.

2.7.3 Development

This section explains the development effort for creating a Data Decentralization Client in a decentralized social media system. The client enables users securely store and retrieve their data from their devices, reducing the need for centralized servers. Below are the key phases of implementation:

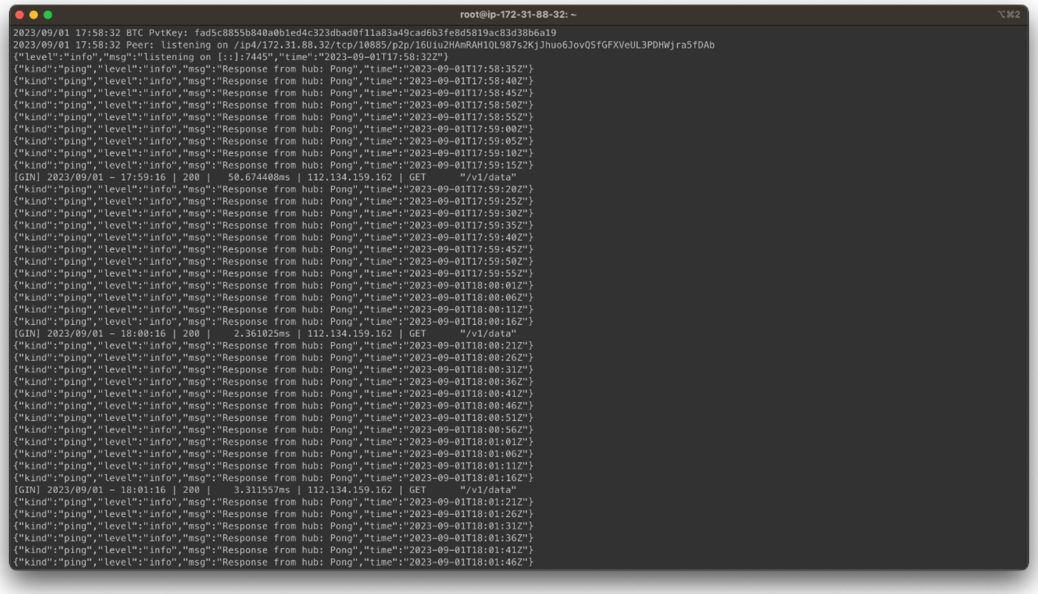
Initialization and Configuration: Client need to initialize with the required environment variables. Users have the flexibility of deciding a database system. The client is made to work with both PostgreSQL and Elasticsearch, giving users the choice that fits them best.

Peer Initialization and Communication: As the client enters decentralized communication, it becomes a network peer, joining to hub to do peer-to-peer interactions. To ensure secured communication and a unique network ID, a special private key is created, used for security and network ID. Picking a network port for data exchange is a strategic decision to be made. Creating communication channels with other peers is crucial for decentralization.

Data Decentralization: The main goal is to hand over the control of user data to user by using the user's device for secure data storage and retrieval. The primary function of the client is to assist in the on-device storage and retrieval of user data, eliminating the need for centralized servers and reducing the risk of outages or data breaches. Client will assist in integrating with database chosen by the user. And also, it will store and retrieve data when required.

Integration and Compatibility: The client's strength is its smooth connection to PostgreSQL or Elasticsearch, depending on the chosen database. This makes data storage and retrieval efficient, using the chosen database's strengths. The client interacts with other parts of the decentralized social media system, like sharing messages, events, or content with other peers. It also watches over its peers to keep the network stable. Strong error-handling methods are ready for unexpected issues, with detailed logs for finding and fixing problems.

The creation of the Data Decentralization Client within the decentralized social media protocol is an essential step toward the development of a more reliable and secure system for data storage and retrieval. This client reduces dependence on centralized servers, boosting data security and availability by giving users control over data management and utilizing their devices. Its flexibility, which enables customers to select the database system that best meets their needs, is ensured by its compatibility with both PostgreSQL and Elasticsearch.



```
root@ip-172-31-88-32: ~
2023/09/01 17:58:32 BTC PvtKey: fad5c8855b840a0b1ed4c323dbad0f11a83a49cad6b3fe8d5819ac83d38b6a19
2023/09/01 17:58:32 Peer: listening on /ip4/172.31.88.32/tcp/10885/p2p/16Uiu2HAmRAH1QL987s2KjJhuo6JovQSfGFXVeUL3PDHWjra5fDAb
{"level":"info","msg":"listening on [::]:7445","time":"2023-09-01T17:58:32Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:58:35Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:58:40Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:58:45Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:58:50Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:58:55Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:00Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:05Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:10Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:15Z"}
[GIN] 2023/09/01 - 17:59:16 | 200 |   50.674408ms | 112.134.159.162 | GET      "/v1/data"
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:20Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:25Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:30Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:35Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:40Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:45Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:50Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T17:59:55Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:01Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:06Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:11Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:16Z"}
[GIN] 2023/09/01 - 18:00:16 | 200 |    2.361025ms | 112.134.159.162 | GET      "/v1/data"
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:21Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:26Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:31Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:36Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:41Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:46Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:51Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:00:56Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:01Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:06Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:11Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:16Z"}
[GIN] 2023/09/01 - 18:01:16 | 200 |    3.311557ms | 112.134.159.162 | GET      "/v1/data"
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:21Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:26Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:31Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:36Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:41Z"}
{"kind":"ping","level":"info","msg":"Response from hub: Pong","time":"2023-09-01T18:01:46Z"}
```

Figure 2.7.3.1: Logs from peer container

2.8 Testing

In terms of testing, the approach used was intended to guarantee the accuracy and dependability of the research project. Each development effort was verified through a methodical testing process before it was integrated into the project's codebase hosted on GitHub and after the deployment.

2.8.1 Unit Testing

Unit tests played an important role in the testing process. These tests were conducted before code was pushed to GitHub, serving as a initial validation step. Unit tests are small, focused tests that verify the correctness of individual components or functions within the code.

Through the conduct of unit tests, potential issues were detected and rectified at an early stage, preventing them from propagating into the codebase. This approach ensured that each development effort underwent validation for correctness and functionality, contributing to the overall stability of the project.

2.8.2 Continuous Integration (CI) and Deployment (CD)

The testing process was tightly integrated with the CI/CD pipeline, streamlining the testing and deployment of developed features. Here's how it worked:

1. **Continuous Integration (CI):** Upon completion of development efforts, the CI pipeline automatically integrated these features. It compiled the code, ran unit tests, and ensured that the new code did not introduce regressions or errors.
2. **Docker Image Building:** Following successful CI, the pipeline proceeded to build Docker images. Docker images are like packaged containers that encapsulate the application and its dependencies.
3. **Continuous Deployment (CD):** A crucial aspect of the testing process was the CD pipeline. This pipeline utilized Watchtower, an automated container updating tool. When new Docker images were created, Watchtower was triggered as part of the CD process.
4. **Watchtower Deployment:** Watchtower's role was to pull the latest Docker images with updated code and deploy them on Amazon Elastic Compute Cloud (Amazon EC2) instances. This automated process ensured that the application

consistently ran the latest code changes.

2.8.3 Smoke Testing

After each deployment, a critical step was the execution of smoke tests. Smoke tests are a set of initial tests designed to verify that the newly deployed release is stable and functional. They are meant to ensure that the release is "smoke-free," indicating that it is ready for further testing and use.

The testing approach offered several notable benefits to the project:

- **Early Issue Detection**: Unit tests allowed for the early detection and resolution of issues in the development process, reducing the likelihood of critical errors in the final product.
- **Quality Assurance**: Through thorough unit tests and continuous integration, a high level of code quality and reliability was maintained.
- **Efficiency**: Automation within the CI/CD pipeline reduced manual testing efforts, enabling faster development cycles and quicker deployment of new features.
- **Consistency:** The automated deployment process with Watchtower ensured that the application consistently operated with the latest code changes, enhancing overall reliability.
- **Rapid Updates**: The testing and CD approach facilitated swift updates and deployments, essential for an academic research project requiring frequent iterations and enhancements.

In conclusion, the testing approach, included unit testing, seamless integration with the CI/CD pipeline, and smoke testing, was important in maintaining the quality, reliability, and efficiency of the project. It enabled the validation of each development effort, early error identification, and the delivery of a robust and continuously evolving research platform.

2.8.4 Validation using practical implementations.

In order to assess the practical implementations of the project, the development of a demo social media platform and a separate benchmarking application was initiated. These efforts allowed valuable insights into the performance and functionality of the built protocol.

2.8.4.1 Benchmarking with Infrastructure as Code (IaaC) Approach

To deploy the benchmarking application, along with the required infrastructure and dependent applications, an Infrastructure as Code (IaaC) approach was followed. This approach brought significant advantages:

- **Reproducibility**: IaaC allowed the definition and recreation of the entire infrastructure consistently, minimizing discrepancies between deployments.
- **Version Control**: Infrastructure configurations were version-controlled, enabling the tracking of changes, effective collaboration, and maintenance of a comprehensive history.
- **Scalability**: With IaaC, easy scaling of the infrastructure up or down in response to varying workloads was possible, ensuring optimal performance.

2.8.4.1.1 Terraform as the IaaC Tool

For implementing the IaaC approach, terraform was selected as the tool of choice. Terraform provided numerous benefits, including:

- **Infrastructure ambiguity**: Terraform supports multiple cloud providers and infrastructure types, giving flexibility in choosing the best-suited resources.
- **Declarative Syntax**: Terraform’ s declarative syntax made it easy to define and manage infrastructure configurations, enhancing readability and maintainability.
- **Modularity**: Terraform allowed the creation of reusable modules, simplifying the deployment of complex infrastructure components.

Along with these 2 practical implementations, testing is carried out to validate the functionality, usability, and reliability of the protocol.

3.0 RESULTS AND DISCUSSIONS

3.1 Results

The implementation of both the demo social media platform and the benchmark application played an important role in validating various aspects of the project. Here, present the results obtained from these implementations, highlighting the project's key achievements and insights.

3.1.1 Demo Social Media Platform Results

The development of the demo social media platform on top of the implemented decentralized social media protocol produced below noticeable results:

- **User Engagement**: The platform successfully facilitated user engagement, demonstrating the protocol's effectiveness in creating a user-friendly social media environment.
- **Feature Integration**: Various features, such as user profiles, posts, and interactions, were seamlessly integrated and validated, showcasing the versatility of the protocol.
- **User Experience**: Users were provided a positive and intuitive experience while navigating and interacting with the demo social media platform, highlighting its user-centric design.
- **Stability**: The platform established stability and robustness, with minimal downtime or disruptions during usage, ensuring a reliable user experience.

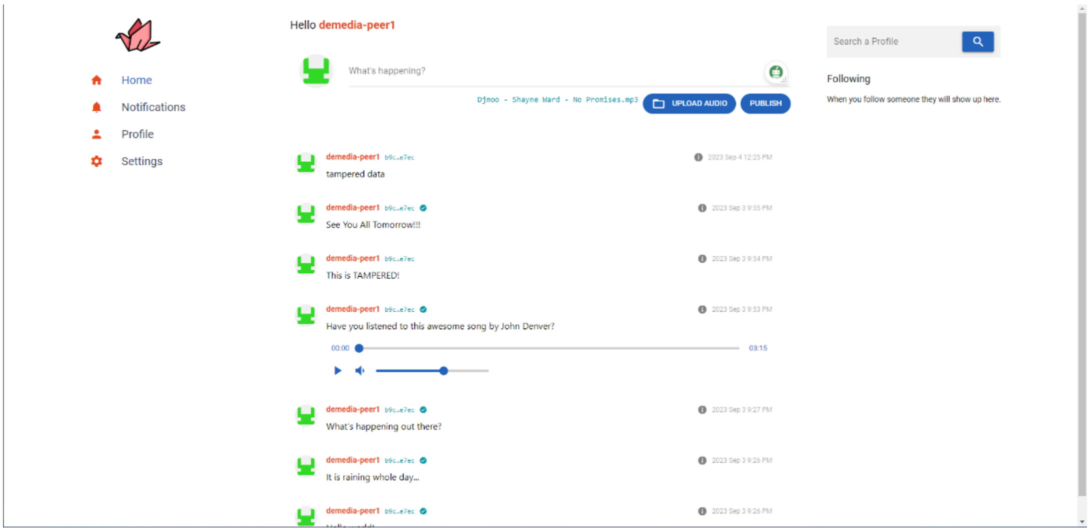


Figure 3.1.1.1: Screen capture of demo social media platform

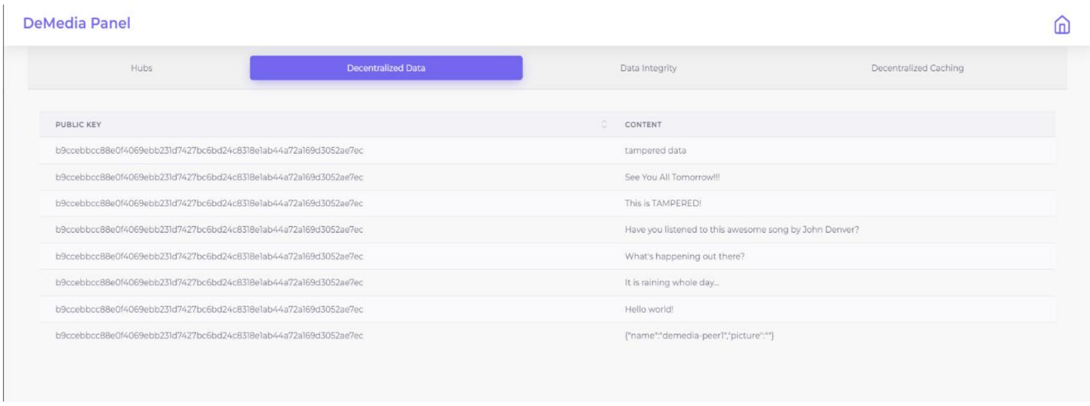


Figure 3.1.1.2: The panel of Decentralized Data

3.1.2 Benchmarking Results

The benchmarking application provided insightful results on the performance and capabilities of the built protocol. These results, which are detailed below, are integral to the project's evaluation and future improvements:

METHOD	REST	IPFS/DEMEDIA
REST API Call vs DeMedia API Call	182.883684ms	432.263605ms
Direct DB Fetch vs DB Fetch Through DeMedia	3.353261ms	432.263605ms
Direct DB Fetch vs REST API Call	3.353261ms	182.883684ms
Time Elapsed To Fetch Image 01	36.62113ms	23.011138ms
Time Elapsed To Fetch Image 02	4.819432ms	9.295224ms

Figure 3.1.2.1: Output from benchmark application

3.1.3 Results From Overall Practical Implementations

The development of the demo social media platform and the benchmark application allowed for the validation of several critical aspects of the project, including:

- **Functionality**: Through rigorous testing and usage of the demo social media platform, validated the core functionality of the implemented decentralized social media protocol.
- **Scalability**: By deploying the demo platform and benchmark application at scale, assessed the protocol's ability to handle a substantial user load and interactions effectively.
- **Security**: Security features and measures were thoroughly examined and validated through the development of the demo platform, ensuring the protection of user data and interactions.
- **Performance**: Performance metrics were gathered and analyzed to evaluate the speed and efficiency of the decentralized social media protocol in real-world scenarios.

3.2 Research Findings

User data decentralization technology enables better data governance, privacy, and potential data commercialization. However, they are challenged with difficulties, including data availability, scalability, security, portability, and acceptance by users. Potential improvements include data sharing, incentive systems, learning federation, blockchain integration, offline access assistance, compatibility, data versioning, and user-friendly interfaces to address these issues and maximize positive aspects. To ensure high data availability, an effective protocol in this context should include capabilities for a distributed storage system with user-controlled data access, secure authorization, fine-grained sharing rights, redundancy, replication, data caching, and peer-to-peer networking.

3.3 Discussion

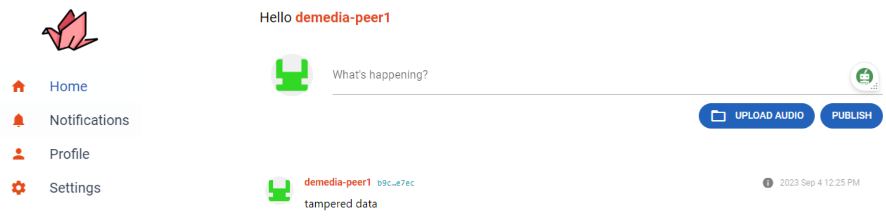


Figure 3.3.1: Screen capture of the demo social media platform

The development of a demo social media platform provided a practical framework to evaluate the effectiveness and usability of the DeMedia protocol. Through the deployment and testing of this platform, the protocol's features and mechanisms were put to the test in a real-world setting. Users engaged with the platform interacted with peers, and shared content, thereby simulating the actual usage scenario. This usability assessment served as a validation of the protocol's design principles and its applicability to real-life social media interactions.

4.0 SUMMARY OF EACH STUDENT'S CONTRIBUTION

Student Number	Name	Tasks
IT2159726	Bandara A. M. C. A.	• Conduct a comprehensive review of user data decentralization protocols and associated research. • Assess the opportunities and challenges present in existing user data decentralization protocols. • Determine the potential enhancements and new capabilities that could be integrated into a mechanism for data caching in a peer-to-peer network. • Determine the potential enhancements and new capabilities that could be integrated into a protocol for storing user data within the user's device. • Implement a protocol with the identified modifications while adhering to best practices.

Table 4.1: Contribution summary

5.0 CONCLUSION

In conclusion, the goal of this research project is to design a decentralized social network architecture that successfully manages all of the essential components of a typical social media platform, all the while placing a top priority on protecting the security and privacy of its users' data. The development of a solid protocol for decentralized data management and the creation of a client application that is user-centric are the two key aspects of this research. Particularly, the information would be kept locally on users' devices, improving data privacy by giving them more control over their personal data.

This research's main objective is to create a decentralized social network architecture that prioritizes the community's demands and concerns while operating securely. Therefore, it attempts to find a balance between the innovative possibilities of decentralization and the fundamental features customers have come to expect from standard social media platforms.

In more general terms, this research represents an important turning point in the development of a digital world where people are fully in charge of their personal data. With the knowledge that their data will be handled with the greatest confidentiality and that their privacy and security are maintained as fundamental values, it seeks to establish a space where individuals may communicate with one another in confidence online. In the end, this vision is in line with how the digital landscape is changing, where user engagement and data protection have become more and more important factors in our connected society.

REFERENCES

[1] B. i. Marketing, "The evolution of social media: How did it begin and where," Maryville Online, 03 March 2021. [Online]. Available: https://online.maryville.edu/blog/evolution-social-media/. [Accessed 08 September 2023].

[2] "What are decentralized social networks?," 06 February 2023. [Online]. Available: https://cointelegraph.com/explained/what-are-decentralized-social-networks#:~:text=Hive%20is%20a%20decentralized%20social,free%20speech. [Accessed 08 September 2023].

[3] Nonstack, "What is web 3.0: A layman's guide to the decentralized web," Developer DAO, 23 June 2023. [Online]. Available: https://blog.developerdao.com/what-is-web-30-a-laymans-guide-to-the-decentralized-web. [Accessed 10 September 2023].

[4] Sketchar, "What advantages does web 3.0 give to creators?," 03 March 2023. [Online]. Available: https://blog.sketchar.io/what-advantages-does-web-3-0-give-to-creators/. [Accessed 08 September 2023].

[5] "Web3 social media platforms in 2023: Why they are the future," GamesPad, 30 March 2023. [Online]. Available: https://gamespad.io/web3-social-media-platforms-in-2023-why-they-are-the-future/. [Accessed 08 Septermber 2023].

[6] "Decentralized social media," Mastodon, [Online]. Available: https://joinmastodon.org/. [Accessed 08 September 2023].

[7] S. Freight, "The decentralized social blockchain," Deso, [Online]. Available: https://www.deso.com/. [Accessed 08 Septermber 2023].

[8] "Top decentralized social networks startups," [Online]. Available: https://tracxn.com/d/trending-themes/Startups-in-Decentralized-Social-Networks. [Accessed 08 September 2023].

[9] K. Z. R. Nourmohammadi, "An On-Chain Governance Model Based on Particle Swarm Optimization for Reducing Blockchain Forks," vol. 10, pp. 118965-118980.

[10] V. Snyder, "How covid-19 has changed social media habits," [Online].

Available: https://www.business.com/articles/social-media-patterns-during-the-pandemic/. [Accessed 08 September 2023].

[11] "Social Media Statistics Details - undiscovered Maine - University of Maine," 08 October 2021. [Online]. Available: https://umaine.edu/undiscoveredmaine/small-business/resources/marketing-for-small-business/social-media-tools/social-media-statistics-details/. [Accessed 08 September 2023].

[12] S. Q. E. Kadena, "Blockchain in social media: Eliminating centralized," in *2022 IEEE 10th Jubilee International Conference on Computational Cybernetics and Cyber-Medical Systems (ICCC)*, 2022.

[13] Z. H. S. L. W. C. Z. Z. Y. Y. T. Cai, "SocialChain: Decoupling Social Data and Applications to Return Your Data Ownership," vol. 16, pp. 600-614, 2023.

[14] T. Sarathchandra and D. J. , "A decentralized social network architecture," pp. 251-257, 2021.

[15] "What is the Mastodon social media platform?," Internet Matters, 15 March 2023. [Online]. Available: https://www.internetmatters.org/hub/news-blogs/what-is-mastodon-social-media/. [Accessed 10 September 2023].

[16] A. Ronacher, "Scaling mastodon is impossible," Armin Ronacher's Thoughts and Writings, [Online]. Available: https://lucumr.pocoo.org/2022/11/14/scaling-mastodon/. [Accessed 10 September 2023].

[17] Adam, "The architecture of Mastodon," SoftwareMill Tech Blog, 08 September 2023. [Online]. Available: https://softwaremill.com/the-architecture-of-mastodon/. [Accessed 10 September 2023].

[18] "The Decentralized Social Blockchain (DeSo)," The Reading Ape, 26 April 2022. [Online]. Available: https://thereadingape.substack.com/p/the-decentralized-social-blockchain. [Accessed 10 September 2023].

[19] HattyHats, "DeSo - a deeper look into the best social blockchain project," Publish0x, [Online]. Available: https://www.publish0x.com/beginning-crypto/deso-a-deeper-look-into-the-best-social-blockchain-project-xxzelmn. [Accessed 10 September 2023].

[20] R. Keefe, "5 popular project management methodologies and when to use them," Toggl Blog, 13 September 2022. [Online]. Available: https://toggl.com/blog/5-popular-project-management-methodologies-and-when-to-use-them. [Accessed 10 September 2023].

GLOSSARY

- User Data Decentralization Protocol (UDDP): The general protocol allows for the decentralized storage and management of user data, giving consumers more control over their information.
- Decentralized Storage: The practice of distributing user data over a network of nodes or servers in order to reduce dependency on centralized data repositories while improving data security and privacy.
- Node: A device or computer operating in a decentralized network that manages and saves user data snippets, enhancing the adaptability and availability of the network as an entire system.
- Data Ownership: The concept is that users maintain ownership and control over their data, giving them the ability to grant or deny access within the distributed system as they see accurate.
- Blockchain: Transparency and immutability are guaranteed by a distributed database that keeps records of all data changes and transactions within the decentralized network.
- Data Privacy: The protection of personal data from unauthorized access, ensuring that only authorized users can access and make use of the information.
- Access Control: The system gives consumers complete control over data sharing by allowing them to specify who can view and interact with their data.
- Data Integrity: The assurance that user data is protected and kept private within the decentralized network, is frequently accomplished through cryptographic methods.
- Peer-to-peer (P2P) caching: a decentralized caching mechanism in which nodes in a network cache data for each other, reducing the load on the network and improving performance.

APPENDICES

USER DATA DECENTRALIZATION PROTOCOL

ORIGINALITY REPORT

5%	2%	2%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Sasha Shilina. "chapter 10 The Promise of Blockchain-Based Decentralized Social Networks", IGI Global, 2023 Publication	1%
2	Submitted to Mt San Antonio College Student Paper	1%
3	"Full Proceedings Printed", 2021 International Research Conference on Smart Computing and Systems Engineering (SCSE), 2021 Publication	1%
4	gamespad.io Internet Source	<1%
5	Submitted to Brigham Young University Student Paper	<1%
6	Submitted to Manchester Metropolitan University Student Paper	<1%
7	5dok.net Internet Source	<1%

dspace.cbe.ac.tz:8080

Appendices -A: Plagiarism Report