Devault (graduation project)

A Blockchain-based end-to-end encrypted cloud storage

Participants:
Abd El-Twab M. Fakhry
Hossam A. Eissa

Supervisor:
Dr. Abdurrahman Nasr

Al-Azhar University
Faculty of Engineering
Computers & Systems Engineering Department

July 23, 2022

Table of contents

Introduction

What is Web 1.0?

Basically, this first version of the web was designed to help people better find information. This web version dealt was dedicated to users searching for data. This web version is sometimes called *"the read-only Web"* because it lacks the necessary forms, visuals, controls, and interactivity we enjoy on today's Internet.

- No user-to-server communication.
- Static websites.
- Hyper-linking and bookmarking pages.
- Read-only Web.

Active 1989–2005

What is Web 2.0?

If Web 1.0 was made up of a small number of people generating content for a larger audience, then Web 2.0 is many people creating even more content for a growing audience. Web 1.0 focused on reading; Web 2.0 focused on participating and contributing.

- Functions such as online documents, video streaming, etc.
- Cloud computing operations
- Centralized data.
- Read and Write Web.

Active 1999-2012

What is Web 3.0?

Web 3.0, which is also referred to as Web3, is built on a foundation consisting of the core ideas of decentralization, openness, and more excellent user utility. Web 1.0 is the "read-only Web," Web 2.0 is the "participative social Web," and Web 3.0 is the "read, write, execute Web."

- The Internet of Things (IoT).
- Semantic searches.
- Decentralized processes.
- Read, Write, and Control Web.

Active 2006-ongoing

What is wrong?

- Censorship
- Giving over control of data
- Data Mismanagement

What is wrong?

- ✖ `Censorship`

  As the internet currently works on a centralized model, it is susceptible to censorship. However, this is an issue that can be easily mitigated with decentralization..
- ✖ `Giving over control of data`
- ✖ `Data Mismanagement`

What is wrong?

- ✖ Censorship
- ✖ Giving over control of data

  The biggest problem of third-party cloud storage services is that the company hands over their data to a third-party for storing services. Since the data is outside the company's control, the data privacy settings are beyond their control as well.
- ✖ Data Mismanagement

What is wrong?

- ✖ Censorship
- ✖ Giving over control of data
- ✖ Data Mismanagement

  media analytics company "Deep Roots Analytics," used the Amazon cloud server to store information about as much as 61% of the US population without password protection for almost two weeks. This information included names, email and home addresses, telephone numbers, voter ID, etc.

Proposed Solution

The **solution** we propose for such a problem is to use:

- A distributed database system that will store data in a peer-to-peer network where is no central authority with the right to modify or censor clients' data.
- Encryption, so that everything should be encrypted before being uploaded.
- A Blockchain and smart contract for identity without a central authority. Verification of data that cannot be faked or changed. Combine this with encryption, data ownership, and replication, and that's what true decentralization means for applications.

Proposed Solution

The **solution** we propose for such a problem is to use:

- A distributed database system that will store data in a peer-to-peer network where is no central authority with the right to modify or censor clients' data.
- Encryption, so that everything should be encrypted before being uploaded.
- A Blockchain and smart contract for identity without a central authority. Verification of data that cannot be faked or changed. Combine this with encryption, data ownership, and replication, and that's what true decentralization means for applications.

Proposed Solution

The **solution** we propose for such a problem is to use:

- ✅ A distributed database system that will store data in a peer-to-peer network where is no central authority with the right to modify or censor clients' data.
- ✅ Encryption, so that everything should be encrypted before being uploaded.
- ✅ A Blockchain and smart contract for identity without a central authority. Verification of data that cannot be faked or changed. Combine this with encryption, data ownership, and replication, and that's what true decentralization means for applications.

Proposed Solution

The **solution** we propose for such a problem is to use:

- ✅ A distributed database system that will store data in a peer-to-peer network where is no central authority with the right to modify or censor clients' data.
- ✅ Encryption, so that everything should be encrypted before being uploaded.
- ✅ A Blockchain and smart contract for identity without a central authority. Verification of data that cannot be faked or changed. Combine this with encryption, data ownership, and replication, and that's what true decentralization means for applications.

Background Materials

Need for decentralization

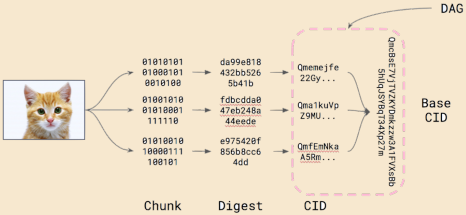
Decentralization is an ideology that advocates for a liberal style of administration in which no single authority has absolute power over all aspects of life.

Some of the benefits of decentralized cloud storage are listed below:

- Encrypted
- Secured
- Less Computer Power with Bandwidth
- No dedicated Servers for Storage
- Fast

The Interplanetary File System (IPFS) is a bundle of subprotocols and a project-driven by Protocol Labs, IPFS aims to improve the web's efficiency and to make the web more decentralized and resilient.

(a) From raw image to cryptographic digest to content id (multihash).



(b) Large files are chunked, hashed, and organized into an IPLD (Merkle DAG object).

Figure 1: IPFS base CID construction

What is Blockchain

A blockchain is a growing list of records, called blocks, that are securely linked together using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data (generally represented as a Merkle tree, where data nodes are represented by leafs).

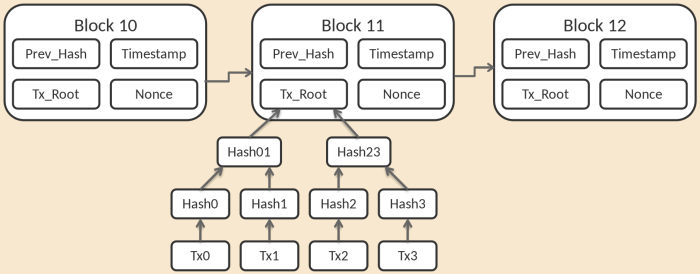


Figure 2: Bitcoin blockchain structure.

What is a Smart contract

A “smart contract” is simply a program that runs on the Ethereum blockchain. It’s a collection of code (its functions) and data (its state) that resides at a specific address on the Ethereum blockchain.

Smart contracts are a type of Ethereum account. This means they have a balance and they can send transactions over the network. However they’re not controlled by a user, instead they are deployed to the network and run as programmed.

Wallet Auth

To transact on Ethereum, you need an account. There is no MySQL “users” table. There is no email/password login.

To create an Ethereum account, you need to set up a crypto wallet like Metamask. The wallet will be responsible for generating and securing your crypto keys for signing transactions.

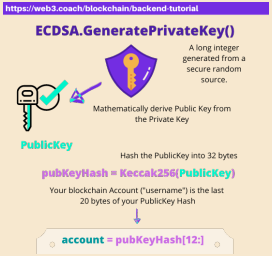


Figure 3: Generating keys and addresses.

Methodology

Development Methodology

In this project we use both waterfall and scrum methodologies. We used the waterfall methodology for developing the smart contract because we can not update it once we deploy it. And for the rest, we used the scrum methodology.

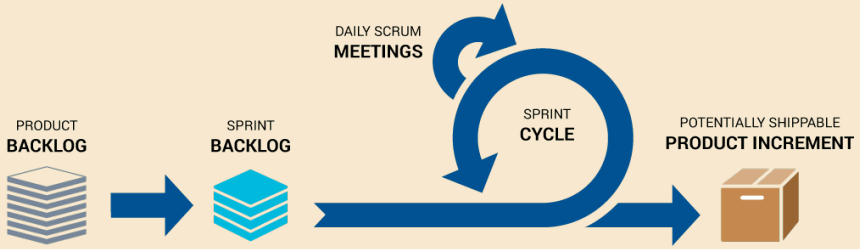


Figure 4: Agile scrum development process.

Use Case Modeling

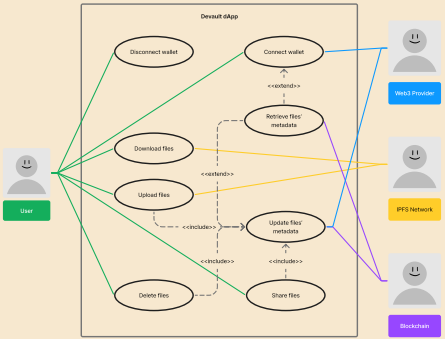


Figure 5: Devault dApp use case diagram.

Sequence Diagram

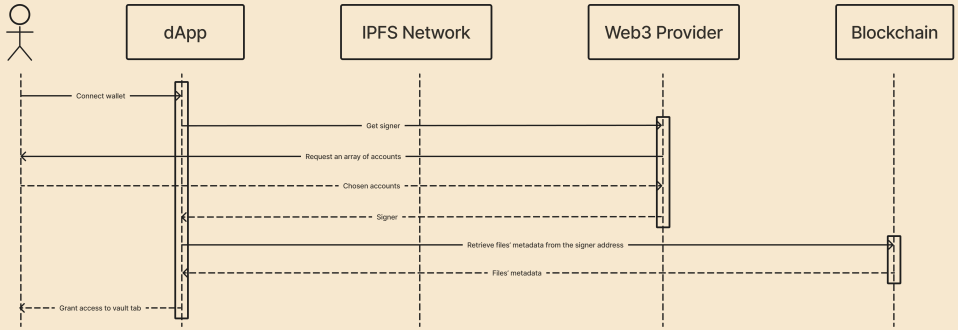


Figure 6: Connect wallet act in sequential order.

Sequence Diagram

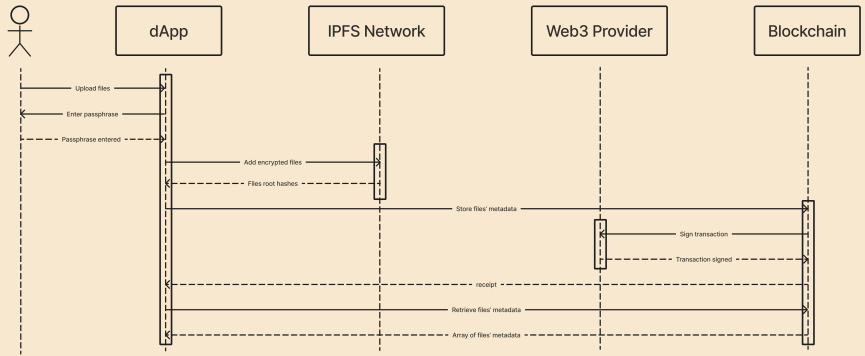


Figure 7: Upload files act in sequential order.

Sequence Diagram

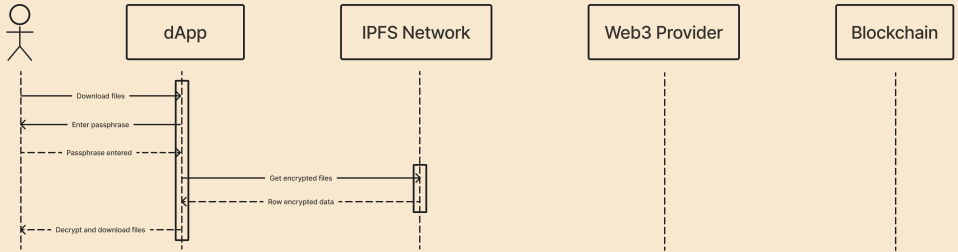


Figure 8: Download files act in sequential order.

System Design

System Architecture

Here is the conceptual model that defines the structure, behavior, and more views of a system.

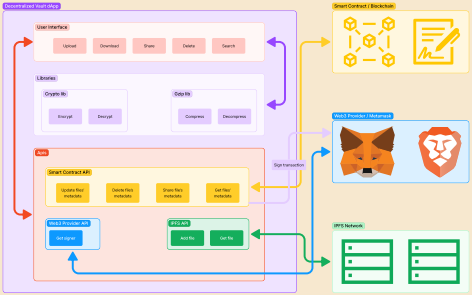


Figure 9: Decentralized vault dApp architecture.

Tools and Technologies

- JavaScript
- Next.js
- Solidity
- Hardhat
- Ethers.js
- Metamask
- Infura
- IPFS
- Etherscan
- GitHub Actions
- Docker
- Vercel
- Bash/Shell
- Emacs
- Neovim
- Figma
- LaTeX

Result and Discussion

Result

We tested the Devault with uploading, downloading, sharing, deleting, connecting, and disconnecting functionality on different browsers and devices. Below we show the results of two of these tests.

1. ***Brave browser, the desktop version and metamask***

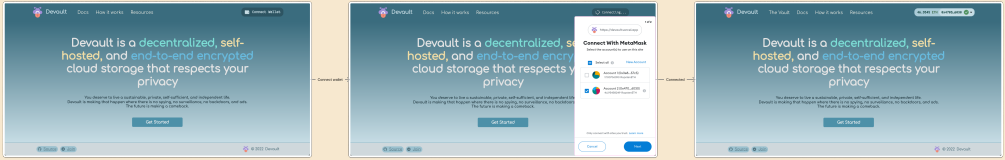


Figure 10: The user experience of connecting wallet.

2. *Brave browser, the mobile version and brave wallet*

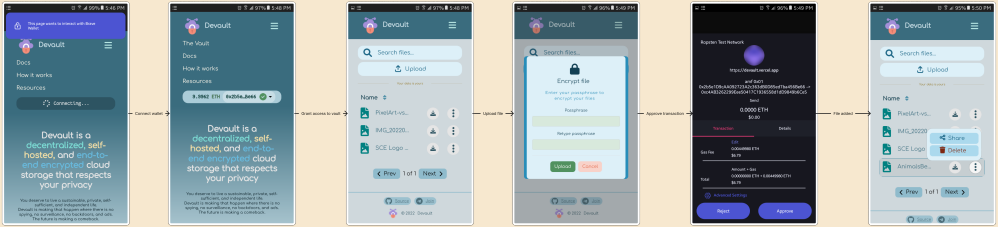


Figure 11: The user experience of uploading files.

Limitation

The limitations of our system are listed below:

- The user should have cryptocurrency to make transactions and to reward the entity for renting their hard disk, which is a constraint.
- The user should install specific software such as metamask or brave wallet to use our system.
- Cryptocurrencies are illegal in some countries, whereas our system needs those cryptocurrencies to work.
- The uploading of a file may take a while because of the encryption and block mining, which is not practical.

Future Work

The limitations of our system are listed below:

1. Support Arabic language.
2. Use Shamir's Secret Sharing algorithms for sharing files.
3. Add the search functionality.
4. Compress files before uploading.
5. Enhance the ui/ux.
6. Manipulate selected files and folders.
7. Deploy to the mainnet.
8. Use Filecoin instead of IPFS.
9. Use the advantages of private and public keys in encryption/decryption.
10. Give feedback if the key used for decryption is not the same key used for encryption.

Demo

Narayan Prusty.
Building Blockchain Projects.
Packt Publishing Ltd, B3 2PB, UK, 1 edition, 2017.

Tiana Laurence.
Blockchain For Dummies.
John Wiley & Sons, Inc, 111 River Street, Hoboken, NJ, 1 edition, 2017.

The Community.
Ethereum development documentation.
Available at `https://ethereum.org/en/developers/docs/` (April 22, 2022).

References ii

Protocol Labs Team.
Welcome to the ipfs docs.
Available at `https://docs.ipfs.io/` (April 22, 2022).

Pratima Sharma, Rajni Jindal, and Malaya Dutta Borah.
Blockchain-based decentralized architecture for cloud storage system.
Journal of Information Security and Applications, 62(8):102970, 2021.

Dylan Yaga, Peter Mell, Nik Roby, and Karen Scarfone.
Blockchain technology overview.
Technical report, National Institute of Standards and Technology, Gaithersburg, MD, 2018.

References iii

Satoshi Nakamoto.
Bitcoin: A peer-to-peer electronic cash system.
Technical report, www.bitcoin.org, 2009.

Soumik Sarker, Arnob Kumar Saha, and Md Sadek Ferdous.
A survey on blockchain & cloud integration.
In *International Conference on Computer and Information Technology*, Sylhet, Bangladesh, 2020.