IpfsCloud

Abstract

In the recent years, concerns about the data governance and centralized cloud services have leads to the innovation of new distributed technologies. Blockchain is the emerging technology which attempts to solve the issue of data tampering and IPFS is a protocol designed to store hypermedia in a peer-to-peer distributed file storage with content addressability. These innovations are trying to tackle technical challenges of existing centrailzed services and giving data governance power to the end users. The framework define in this paper attempts to combine both these technologies and create an open, decentalized, anonymous storage and computation platform. The aim of this framework is to create an open marketplace for data and computational power. The framework also proposes a set of applications and their APIs for making it easier for other developers to utilise the true potential of these amazing technology.

Keywords

Blockchain, IPFS, Cloud, Storage, Distributed Ledger Technology, P2P System

Introduction

Todays, Internet is mostly govern by centralized cloud based web services and in the recent years thera are increasing concerns on the way they control and missuse user priavte information. Decentralized systems have tried to tackle these issues through providing a secure, immutable track of records in the ledger. However, they are still hindered by several drawbacks, such as the ease to use, control and the lack of interoperability of the data between different distributed ledger applications. Full decentralization would be certainly useful, especially for user centric applications where user can govern and control the data.

To make adoption of decentralized systems easier, there is a need for models and frameworks that explore how blockchain and IPFS technologies can be combined, in order to unveil the possibilities of recent innovations. This work proposes a framework for the design and development of open distributed application and offers a wide range of developer tools and infrastructure support that make it easier for the creation of robust dapps creation which can be used in day-to-day life.

The rest of the paper is structured as follows. In the next section an introduction to the Blockchain and IPFS technologies is given. Next an introduction to the proposed tools, framework and applications created using the proposed tools and framework have been given, follows by conclusion.

Background of IPFS

IPFS is a peer-to-peer distributed file system that seeks to connect all computing devices with the same system of files. In some ways, IPFS is similar to the World Wide Web, but IPFS could be seen as a single BitTorrent swarm, exchanging objects within one Git repository. In other words, IPFS provides a high-throughput, content-addressed block storage model, with content-addressed hyperlinks.

What is IPFS?

IPFS is a distributed file system that seeks to connect all computing devices with the same system of files. In some ways, this is similar to the original aims of the Web, but IPFS is actually more similar to a single bittorrent swarm exchanging git objects. IPFS could become a new major subsystem of the internet. If built right, it could complement or replace HTTP. It could complement or replace even more. It sounds crazy. It is crazy.

HTTP is inefficient and expensive HTTP downloads a file from a single computer at a time, instead of getting pieces from multiple computers simultaneously. With video delivery, a P2P approach could save 60% in bandwidth costs.
The web’s centralization limits opportunity. The Internet has been one of the great equalizers in human history and a real accelerator of innovation. But the increasing consolidation of control is a threat to that.
Our apps are addicted to the backbone Developing world. Offline. Natural disasters. Intermittent connections. All trivial compared to interplanetary networking. The networks we’re using are so 20th Century. We can do better.

How IPFS works

1.Each file and all of the blocks within it are given a unique fingerprint called a cryptographic hash.
2.IPFS removes duplications across the network and tracks version history for every file.
3.Each network node stores only content it is interested in, and some indexing information that helps figure out who is storing what.
4.When looking up files, you’re asking the network to find nodes storing the content behind a unique hash.

Introduction to IpfsCloud

Uber for Data and Computation Power Economy

Today data and computational power services are highly centralized both technically and economically. This model has a number of drawbacks. Due to this economic centralization, a few players control the market and charge prices as much as they want. Also, as there are less compititors leads to less compitition. Technically, centralized systems are based on trust that the provider does what it promises and also, this model is more prone to hacks and failures.

We propose a system with following design goals:

* **Secure**: Consensus Algorithm which not only ensures security, but also the work done in process of achieving consensus is not wasted(as in POW).

* **Avaliable**: The network should ensure that the data will be available under some assumptions(depends on number of faulty nodes, replication factor).

* **Auditable:** Using novel algoritms like Proof of SpaceTime, proof-of-storage, proof-of-replication to ensure auditability.

* **Incentivised**: Using Filecoin to introduce incentivisation for following the protocol rules and deincentivising defaulters.

* **Programmable**: While the above(IpfsCloud) supports primitive ways to interact with the network, we enable for more complex operations to be designed on top of primitive methods by supporting a deployment of smart contracts. Users can program new fine-grained storage/retrieval requests that we classify as File Contracts as well as generic Smart Contracts. We integrate a Contracts system and a Bridge system to bring IpfsCloud in other blockchain, and viceversa, to bring other blockchains' functionalities in IpfsCloud.

* **Interoperable:** The reulting system can be used with other blockchain systems which require storage.

IpfsCloud uses Filecoin, an opensource protocol to accomplish some of the above design goals.

Filecoin is a decentralized storage network/protocol that turns cloud storage into an algorithmic market. The market runs on a blockchain with a native protocol token (also called "Filecoin"), which miners earn by providing storage to clients. Essentially, filecoin is the airbnb of data storage—a peer to peer based exchange that accepts asks and bids to settle decentralized data storage transactions on blockchain.

Filecoin works as an incentive layer on top of IPFS, which can provide storage infrastructure for any data. It is especially useful for decentralizing data, building and running distributed applications, and implementing smart contracts.

There are 3 types of entities in the system:

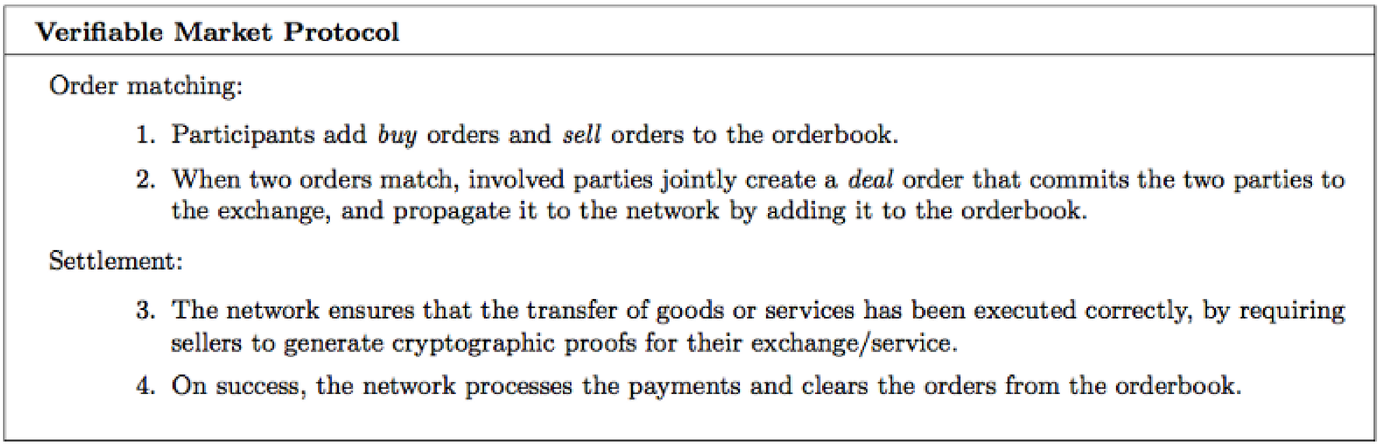
•**Client**—Pay to store data and to retrieve data in the DSN, via Put and Get requests

•**Miner**—Storage Miners provide data storage to the network. Storage Miners participate in Filecoin by offering their disk space and serving Put requests. To become Storage Miners, users must pledge their storage by depositing collateral proportional to it.

•**Retrieval Miner**—Retrieval Miners provide data retrieval to the Network. Retrieval Miners participate in Filecoin by serving data that users request via Get. Unlike Storage Miners, they are not required to pledge, commit to store data, or provide proofs of storage.

Clients pay a network of miners for data storage and retrieval; miners offer disk space and bandwidth in exchange of payments. Miners receive their payments only if the network can audit that their service was correctly provided.

Design Goal: Secure(Consensus Protocol)



Verifiable Market Protocol

Order matching:

1. Participants add *buy* orders and *sell* orders to the orderbook.
2. When two orders match, involved parties jointly create a *deal* order that commits the two parties to the exchange, and propagate it to the network by adding it to the orderbook.

Settlement:

3. The network ensures that the transfer of goods or services has been executed correctly, by requiring sellers to generate cryptographic proofs for their exchange/service.
4. On success, the network processes the payments and clears the orders from the orderbook.

•**The Filecoin protocol is a Decentralized Storage Network construction built on a blockchain and with a native token.** DSNs aggregate storage ordered by multiple independent storage providers and self-coordinate to provide data storage and data retrieval to clients.

Design Goal: Auditable

In the Filecoin protocol, storage providers must convince their clients that they stored the data they were paid to store; in practice, storage providers will generate Proofs-of-Storage (PoS) that the blockchain network (or the clients themselves) verifies.

Proofs-of-Storage(PoS) schemes such as **Provable Data Possession**(PDP) [2] and**Proof-of-Retrievability**(PoR) [3, 4] schemes allow a user (i.e. the verifier V) who out sources data D to a server (i.e. the proverP) to repeatedly check if the server is still storing D. The user can verify the integrity of the data outsourced to aserver in a very efficient way, more efficiently than downloading the data. The server generates probabilistic proofs of possession by sampling a random set of blocks and transmits a small constant amount of data in a challenge/response protocol with the user

Proof-of-Replication(PoRep) is a novel **Proof-of-Storage** which allows a server (i.e. the prover P) to convincea user (i.e. the verifier V) that some data D has been replicated to its own uniquely dedicated physical storage. Our scheme is an interactive protocol, where the prover P:

(a) commits to storendistinctreplicas (physically independent copies) of some dataD, and then

(b) convinces the verifier V, that P is indeed storing each of the replicas via a challenge/response protocol.

To the best of our knowledge,PoRep improveson PoR and PDP schemes, preventing Sybil Attacks,Outsourcing Attacks, and Generation Attacks.

Design Goal: Programmable

Smart Contracts enable users of Filecoin to write stateful programs that can spend tokens, request storage/retrieval of data in the markets and validate storage proofs. Users can interact with the smart contracts by sending transactions to the ledger that trigger function calls in the contract. We extend the Smart Contract system to support Filecoin specific operations (e.g. market operations, proof verification).

Filecoin supports contracts specific to data storage, as well as more generic smart contracts:

- **File Contracts**: We allow users to program the conditions for which they are offering or providing storage services. There are several examples worth mentioning: (1) contracting miners: clients can specify in advance the miners offering the service without participating in the market, (2) payment strategies: clients can design different reward strategies for the miners, for example a contract can pay the miner incresignly higher through time, another contract can set the price of storage informed by a trusted oracle, (3) ticketing services: a contract could allow a miner to deposit tokens and to pay for storage/retrieval on behalf of their users, (4) more complex operations: clients can create contracts that allow for data update.

- **Smart Contracts**: Users can associate programs to their transactions like in other systems (as in Ethereum [18]) which do not directly depend on the use of storage. We foresee applications such as: decentralized naming systems, asset tracking and crowdsale platforms

Design Goal: Available

- Clients can select replication parameters to protect against different threat models.

- Clients can eventually retrieve data from miners.

Impacts on other fields

Fields like machine learning, gaming, graphic designing(3d graphics); In general any field which needs large storage will be disrupted. Like in machine learning for a normal user, training an AI is a computationally hard task, as the methods/models are power hungry. Also, not every one has access to the powerful servers hosted by centralized entities. A platform like this will provide a cheap and powerful solutions to individuals or groups.

COSMOS

Peer-to-peer communication Engine.

Current collaboration systems, from productivity suits including documents, presentations and spreadsheets collaboration, to online gaming, face frequent problems in syncing and communication data accross the devices, and are prone to hacks and faliure to their centralized model. Also, to maintain such systems, heavy infrastructure is required.

COSMOS is a **peer-to-peer communication engine**. It uses IPFS pub-sub for sub-millisecond peer-to-peer communication.

What is pub-sub?

Publish-Subscribe, called ‘pubsub’ for short, is a pattern often used to handle events in large-scale networks on IPFS. ‘Publishers’ send messages classified by topic or content and ‘subscribers’ receive only the messages they are interested in, all without direct connections between publishers and subscribers. This approach offers much greater network scalability and flexibility.

Applications

Some **applications** include

* collaborative document editing,

* “dynamic” website content,

* chat applications,

* multiplayer games,

* collaborative AR, VR experiences,

* continuously evolving datasets, and webservice workers passing around messages,

* lighting fast domain record(IPNS) updates.

It gives us ways to make IPFS fast for large-scale networks such as datacenters, local area networks, and large p2p applications.

DappBase(FireBase on Ipfs)

*** Talk about what is the problem with firebase(drawbacks of Firebase and centralized systems. Get some evidence of hacks and drawbacks. They are slow and firebase is prone to hacks)**

problems with firebase:

Migrating data is

DappBase is the child of IpfsCloud and COSMOS.

Storage comparison

Authentication method comparison

Domain records Comparision

Costs Comparison

Decentralization comparision

Diagram of the DappBase Stack and it’s comparison with Firebase

*** Storage(static content):** Decentralized Storage market place using IpfsCloud v.s Centralized storage.
* **Database(dynamic content):** Decentralized Storage market place using IpfsCloud and lighting fast updates accross the devices using COSMOS v.s. traditional database architecture.
*** Domain records:** IPNS for managing domain records and COSMOS for lighting fast domain record updates v.s. traditional expensive and in-efficient domain records system.

Study **BlockStack(https://blockstack.org/)**

Apps on DappBase

IpfsDocs:

IpfsDocs is an open, decentralized document collaboration platform. It is a realtime P2P collaborative editing tool, powered by IPFS and CRDTs.

Storage Comparison: Centralized google servers v.s. decentralized servers

Authentication Comparision: Google Authentication v.s. Decentralized Auth

IpfsHost:

IpfsHost is an open, decentralized website/webapp building and hosting platform.

Decentralized Search Engine

Write what will it look like, it’s benefits, some economics. More specifics will be revieled in the separate paper.

Decentralized AD Network

Write what will it look like, it’s benefits, some economics. More specifics will be revieled in the separate paper.

Conclusion

Reference

Deisgn goals, must guarenteed