Introducing EOS.IO Storage

Abstract: EOS.IO Storage is a proposed decentralized file system designed to give everyone the ability to permanently store and host files accessible by any web browser. Unlike some other proposed alternatives, there would be no upfront fee or ongoing charge for storage or bandwidth on EOS.IO Storage aside from a completely refundable deposit. Users must hold tokens while they need storage and bandwidth and may sell tokens when storage and bandwidth is no longer required. Built on the InterPlanetary File System (IPFS) and the EOS.IO software, EOS.IO Storage will be a service provided by block producers for those who hold tokens on a blockchain that adopts the EOS.IO software. The block producers would be incentivized to replicate and host these files, allowing anyone with an Internet browser to access them.

Background

IPFS

IPFS is an emerging standard for storing content addressable files. Content-addressable storage is a mechanism for storing information that can be retrieved based on its content rather than its location. Stated another way, all files stored using IPFS are given names derived from the hash of their content.

What this means is that the same file will have the same name on every computer, and the contents of that file can never change without also changing the name of the file. It also means that when you download a file from a server you can verify that it is the exact file you requested by recalculating the name based on the content provided by the server.

IPFS also provides a peer to peer (P2P) network layer that allows computers to discover and share files based on their deterministic names. However, this P2P network layer does not provide or guarantee storage, hosting, or bandwidth. As it is currently structured, the IPFS network expects users to provide their own servers and related infrastructure.

EOS.IO

EOS.IO is software designed to allow anyone to create and launch their own smart contract platform. A smart contract is self-executing computer code that automatically enforces its terms and validates user actions. Blockchains are secured by reaching consensus on the order of valid user actions and then applying their deterministic state machine to derive the current application state.

Because the security of a blockchain is highly dependent upon it being heavily replicated and 100% available, it is not suitable for storing large, potentially prunable, files. For example, a high performance blockchain processing 1 million transactions per second will grow at over 100 MB per second, assuming 100 bytes per transaction. To remain practical, these blockchains may periodically truncate their transaction history and take snapshots of the state. Furthermore, the blockchain ledger is replicated to every node which creates an unnecessary level of replication overhead. Storing bulk data in either the transaction log or the blockchain state is neither a practical nor a scalable solution to decentralized file storage.

To address this problem, some blockchain applications have opted to store the IPFS file names. This process ensures that the smart contracts are referring to specific and incorruptible files, but makes no guarantees about the availability of those files.

IPFS does not guarantee the availability of files; a file may disappear if nodes decline to make it available. An inaccessible file may ultimately break the utility and purpose of a smart contract as parties are no longer able to verify the meaning of the file. For example, consider a smart contract that references a will by its IPFS name. That contract may fail if the file containing the will is unavailable, which could happen if someone forgets to pay for ongoing file hosting, or if the deceased person's estate fails to arrange to pay for file hosting. Smart contracts cannot simply store IPFS filenames and be confident that the file will always exist and be accessible when needed.

Filecoin, Maidsafe, Siacoin, and Storj

Filecoin is a decentralized storage network created by the team behind IPFS for the purpose of incentivising the storage of files on IPFS. This protocol creates a blockchain that utilizes the latest advancements in cryptographic proofs to generate trustless proof-of-storage and proof-of-replication. The protocol then incentivises individuals to run auditors that spot-check storage providers.

Filecoin is the currency that storage providers are paid when someone wishes to store or fetch a file from the network. The underlying idea is that there are vast quantities of unused storage sitting on home computers and servers around the globe. Filecoin aims to enable the owners of this unused storage to monetize it, while eliminating any need for 3rd parties to trust the storage providers, or vice versa.

The model adopted by Filecoin is similar to other decentralized storage solutions such as Maidsafe, Storj, and Siacoin. They all attempt to collect micropayments for both the storage and retrieval of data, and they all create their own dedicated currency. Furthermore, all of these products target home computer storage providers renting out space located behind slow internet connections. Lastly, they all require users to continually purchase cryptocurrency to pay for storage and bandwidth. This means the files may not be available for the general public to access for free via their browser.

The cost of storage and bandwidth on these networks may be higher than that offered by cloud service providers such as Amazon S3. For example, at the time of this writing, Storj charges $0.05 per GB of download whereas Amazon charges $0.01 per GB downloaded. Storj charges $0.015 per GB per month whereas Amazon charges $0.0125 per GB per month for infrequently accessed storage (Glacier).

It is not clear that the designs of Filecoin, Maidsafe, Siacoin or Storj scale up to many users and many accesses. As the number of users and files grows, the number of recurring payments will also grow. This will place increasing stress on their single-threaded blockchains as the base transaction load grows just to maintain the status quo. Users wanting to store files will need to set up their own server to make automatic crypto payments or they will have to log in every month to do it manually. The overhead of zero-knowledge proofs and spot checks consume bandwidth and CPU resources whose cost may be greater than the actual cost of the storage and bandwidth being managed.

DropBox, Mega, GoogleDrive, and iCloud

These services provide users 2GB to 50GB of free storage and some bandwidth. These services are freemium products used to upsell their paid products. Unfortunately, these services do not have a common file naming system, like IPFS, nor do they integrate with an open P2P network, nor are they decentralized. Each is entirely controlled by its respective single legal entity, and it is not uncommon for one of these services to have some down time or to change their pricing model.

The Design of EOS.IO Storage

For the purpose of this paper we will assume someone has deployed an EOS.IO based blockchain with native tokens called TOK. A filesystem smart contract, @storage, is deployed to the TOK blockchain, this smart contract allows every user to define a directory structure where all files are links to an IPFS file.

A user creates a link to an IPFS file by signing a transaction that is broadcast to the TOK blockchain. The transaction includes the path relative to the user's "home directory", the corresponding IPFS file name, and the size of the file. The user also specifies whether they want the file be stored and hosted by the TOK block producers.

The user will then upload the file to one of the block producers via standardized REST Application Programming Interfaces (APIs) defined by the EOS.IO Storage software. Once the producer verifies the file has the size and the IPFS name indicated by the user, the producer will broadcast a transaction to the TOK blockchain indicating the file has been received. The other block producers will then replicate the file over an IPFS network.

Storage Quota

Collectively, the block producers vote on how much total storage capacity they would like to offer. The median value of the producer votes is the expected capacity that all producers will have to provide. Block producers are incentivised to increase capacity as they compete for votes from TOK holders. A grace period may be offered during which those below the mean can increase their available capacity.

In order for a user to utilize storage, they must first reserve it by locking TOKs in the @storage smart contract - essentially a fully refundable security deposit. A user may unlock their TOKs by releasing the block producers from the requirement to store and host the files, although these files may still be available via other IPFS hosts. Assuming the price of TOK is constant, the ongoing cost of storage and bandwidth is 0. The market value of TOK may rise or fall while one's files are stored. Either way, an individual will pay 0 net TOK for their storage and bandwidth usage.

The amount of storage available per TOK token is determined using the Bancor algorithm that maintains a Constant Reserve Ratio (CRR) of 10. A CRR means that the storage will never be completely consumed, as the price (locked TOK per megabyte) will rise as free capacity shrinks. A CRR of 10 is based on the fact that most TOK holders will not demand access to all of their storage, which therefore minimizes the cost of over-provisioning the network.

The equation to the right defines Balance as the total amount of storage consumed by all parties. Supply is the total amount of storage the block producers physically have, and CCR is the constant reserve ratio.

$$Price = \frac{Balance}{Supply \times CRR}$$

Collectively the block producers may adjust the CRR (up or down), or adjust the total storage supply (up or down), but may never decrease the storage supply below what has already been claimed (balance).

Objectionable Data

EOS.IO software is designed to combine smart contracts with legally binding arbitration. In addition to having code, these contracts can also impose subjective requirements on the parties. Block producers and storage users enter into a smart contract paired with a legal contract that agrees block producers may be responsible for controlling objectionable content. Pursuant to the arbitration dispute resolution mechanism provided by the network, anyone can seek a ruling that any stored file is objectionable and should be deleted if its storage and hosting is in violation of laws or other contracts.

The EOS.IO Storage protocol will allow a block producer to delete any file where required by law or arbitration. Not all block producers will be subject to the same laws and regulations; therefore, it will be up to the community of TOK holders to determine whether block producers are deleting files fairly and reasonably. Misbehaving producers can be voted out and/or brought before arbitration under the blockchain's constitution.

It is important to understand that the use of the IPFS network places fundamental limits on the ability of EOS.IO Storage to censor data. While the block producers may no longer store or serve a particular file, the file may still be available if someone else hosts it on the IPFS network. The identifier remains an accurate descriptor of the file, and any independent full node may also employ an independent IPFS node to access the file. An individual may choose to host it themselves or pay someone else to host the file on their behalf. In this case the individual or their service provider will assume the liability for hosting and serving the file.

Privacy

EOS.IO Storage is a platform for hosting public data. Users who need privacy may apply an encryption algorithm prior to uploading their files. While the content of the encrypted file will be private, the identity of the blockchain account that uploads the file will still be visible to everyone.

Decentralization and Replication

The core of EOS.IO Storage will be IPFS, which provides a decentralized network where anyone can host files that are discoverable by their address. The block producers represent 20 or more unique and independent individuals or organizations each of which could replicate and host the data in different jurisdictions around the globe. These producers could already be located in data centers capable of supporting high-throughput EOS.IO transaction volume. As long as at least 1 of the 20 block producers is online and makes the document available, then the file will be available to everyone.

This approach will provide a level of replication and bandwidth availability that is significantly greater than other decentralized solutions which use a lower level of replication. The reliability of the service will also be significantly greater because block producers need to maintain uptime to keep their votes and get paid for producing blocks.

According to the proposed storage smart contract and corresponding legal obligations, block producers who are not in the top 25 by total votes will not be obligated to offer the EOS.IO Storage service; however, they should indicate their ability to enable the service quickly once they are voted into the top 25.