Arweave: Permanent Censorship-Resistant Storage For The Web3 World

Rasheed Saleuddin

Jul 21, 2021

In the early hours of June 17th, 2021 the Hong Kong police raided the offices of Apple Daily and arrested five employees. Whatever you may think of Apple Daily's role in the 2019 protests, crypto networks like Bitcoin and Ethereum were devised initially to prevent censorship and the crypto community is generally hyper-pro free speech. One of the primary original goals in Satoshi's 2008 paper is for Bitcoin to be censorship-resistant.

The emergent decentralized web – termed Web3 – demonstrated its usefulness in countering censorship, in this case of the press, allowing Apple Daily to upload and preserve all of its content and files to a permissionless, censorship-resistant, and permanent storage network, Arweave.

The ability to prevent the censorship of information is possible thanks to a new type of network. Arweave founder, Sam Williams, initially devised the Arweave's "blockweave" network with the goal of preventing the control of autocratic governments in order to avoid the dystopia of George Orwell's 1984:

Source: Wikipedia under license via Creative Commons.

The term for a fraction of an Arweave token (AR), the Winston, is even named after 1984's protagonist Winston Smith.

Arweave is one of many growing solutions for perhaps the biggest limitation of the popular blockchains – limited, expensive, non-interoperable, and non-scalable storage. The most utilized blockchains are not designed to store data. Bitcoin has almost no room in its blocks

for storing anything more than keys and balances, while Ethereum frequently hits its limit processing simple smart contracts.

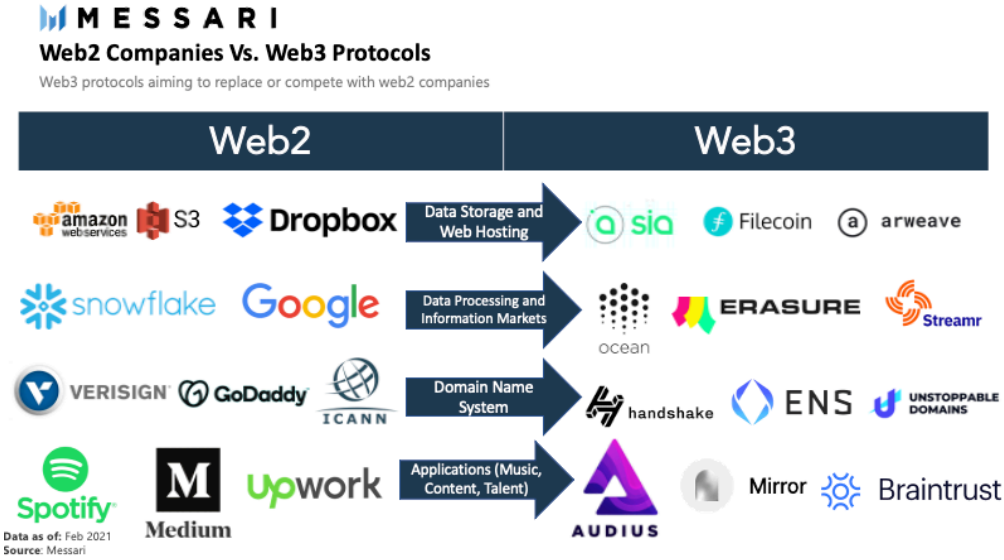
Scalable decentralized storage is a critical component of any crypto ecosystem, and will almost certainly become infrastructure for Web3, – the decentralized internet.

Web3 Storage

Decentralized storage platforms already serve real needs for Layer-1's and NFT markets (see below). More importantly, storage is required to form the backbone of Web3 and replace the heavily centralized and monopolized internet of the Web 2.0 era.

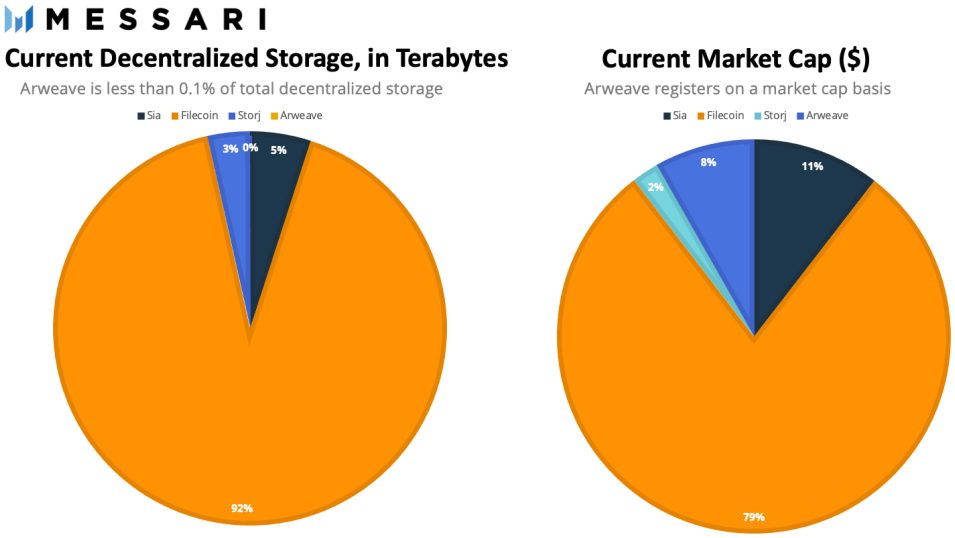
Growing demand by Netflix and other storage-intensive services has created a cloud computing and storage industry with annual revenues on track to hit $150 billion this year, with Amazon Web Services (AWS) dominant at 70% of the total market, Most crypto-based storage solutions, therefore, are attempting to decentralize the services AWS S3, Microsoft Azure, and Google Cloud provide.

Web3 versions of centralized platforms for social media, email, and music need the equivalent of AWS for data storage and tools to store and retrieve said data (e.g. Dropbox).



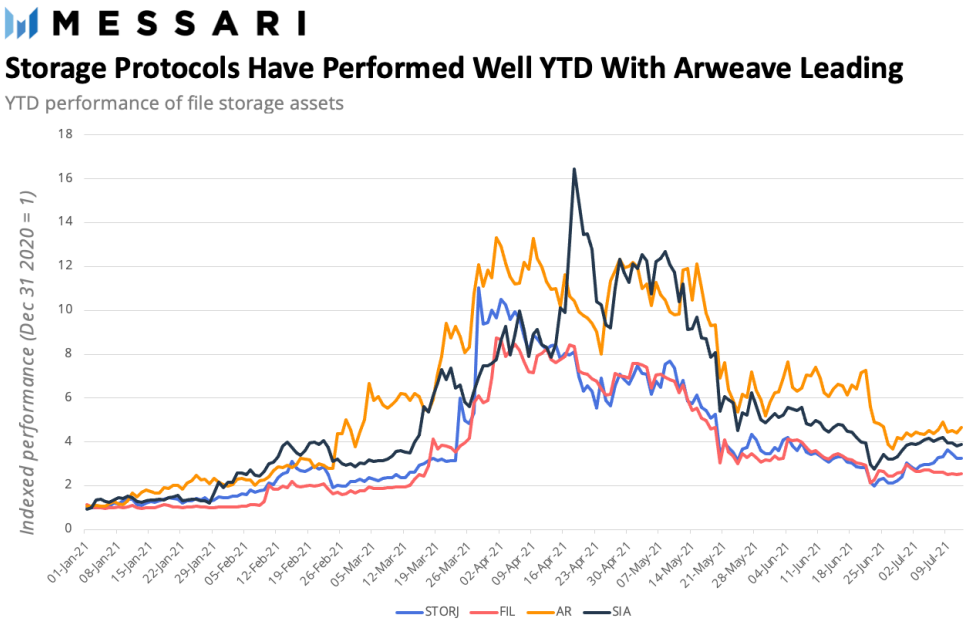
Source: Messari.

While each data storage network has its own strengths, ecosystems have so far coalesced around some large-scale solutions that compete with AWS S3. Sia, Filecoin, Storj, and Arweave are the four platforms to emerge to date. All pay "miners" to store data using their native token. The winner so far appears to be Filecoin, in terms of data storage volumes and market capitalization.



Data as of: 12 July 2021
Source: Filecoin.io. Storjshare.io. Siastats.info. Viewblock.io. Coingecko.

In spite of the size difference, however, Arweave's token has outperformed the others this year, up 364% year to date, even if market capitalization differences are still immense.



Data as of: 12 July 2021
Source: Coingecko.

It's easy for Arweave to be ignored given the age and dominance of Filecoin and the other existing players. But Arweave aims to create more than a decentralized AWS S3. Arweave's current competitive advantage is that it is the only storage solution to offer **data permanence**.

Data permanence – the ability for a piece of data to remain retrievable over time – is an issue in Web2 and may continue to proliferate into Web3. To date, one-third of all social media links are broken within two years. As with their centralized Web2 platform equivalents,

Filecoin, Sia, and Storj are not designed to guarantee access to data if the users stop paying.

A centralized company cannot be trusted to be around permanently – the average lifespan of companies listed in the S&P 500 is currently 18 years. While AWS and Google are dominant today, the lifespan of a company ultimately runs its course. However, a decentralized network, like Bitcoin, or Arweave's blockweave, has the capability to exist in perpetuity.

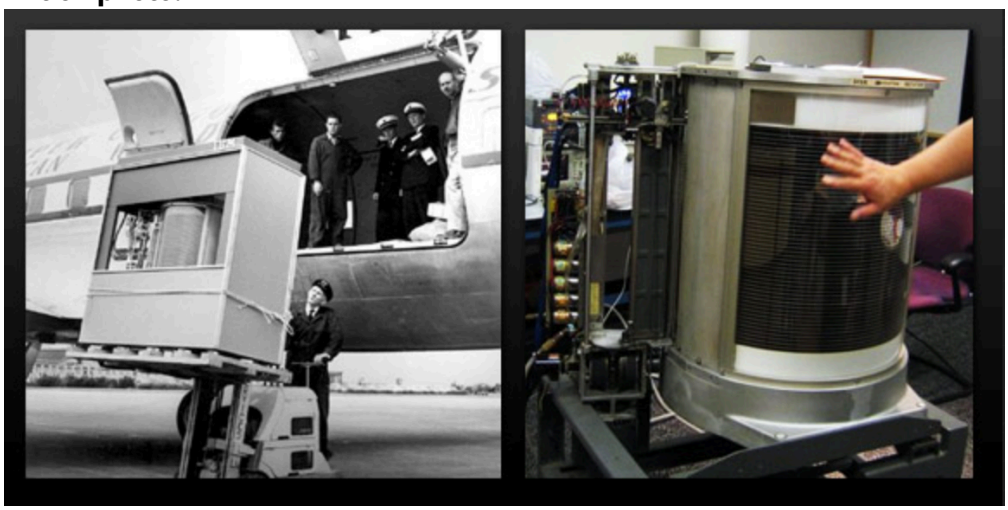
Arweave: Censorship-Resistant Permanent Storage

Arweave's blockweave is a set of data blocks that link to previous blockweave history. In order to mine a new block of data and earn AR tokens from the endowment, miners must provide "Succinct Proofs of Random Access" (SPoRA), – proof that they've been randomly verifying the data. The system rewards miners who provide the most data close to the CPU compute layer.

The minimum number of copies of the entire blockweave is set at ten, but is currently much higher given AR incentives. Miners can choose what data to store, allowing them to censor their own copy if required under local law without affecting the global network.

Arweave requires AR tokens to be added to the endowment of AR tokens to cover the incremental costs of 200 years of the added storage in advance. Even if some miners drop out of storing files, the ongoing endowment should entice new entrants to take over. The trick here is that storage costs have been falling by 30% per year, while the endowment model only assumes a 0.5% annual decline in costs.

IBM 350 hard drives ca. 1954, each able to store 3.7 MB, equivalent to a single high definition photo.



Source: Wikipedia via Creative Commons.

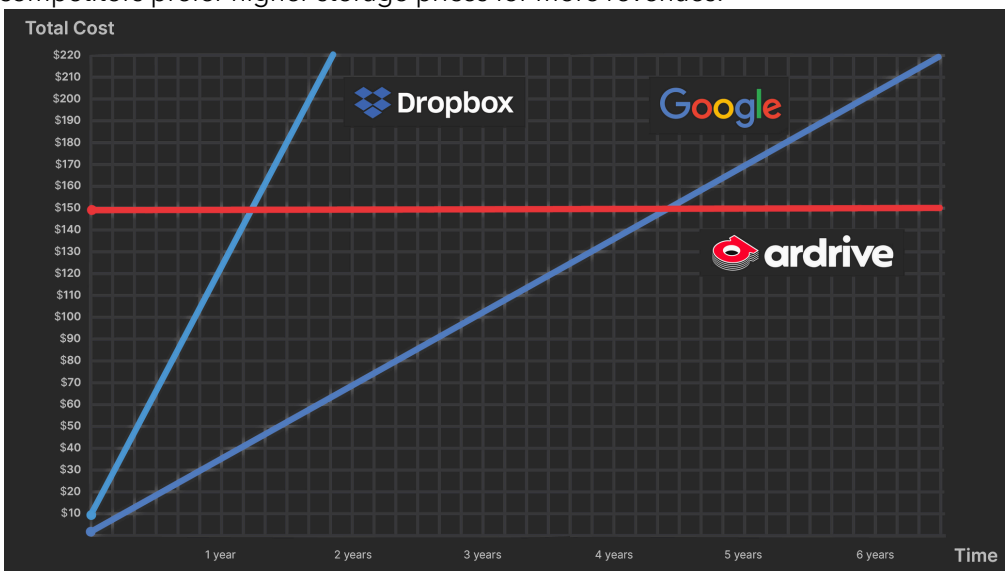
One need only think of the table-sized hard drives of the 1960s to realize how fast storage costs fall, with a (much smaller) PC drive today storing 1 million times more data than the 1960s version seen above.

Currently, the cost for permanent storage is 0.4 cents per megabyte. At the current AR price of around $10, the endowment holds $180,000, more than enough to pay for the current blockweave of 10 terabytes (TB) of data storage for the next 200 years.

Arweave is currently much more expensive than Web2 and even Web3 competitors and over time it is unclear if it will be cheaper than its decentralized competitors that focus on non-permanent storage. However, Arweave's endowment model stands out as a mechanism

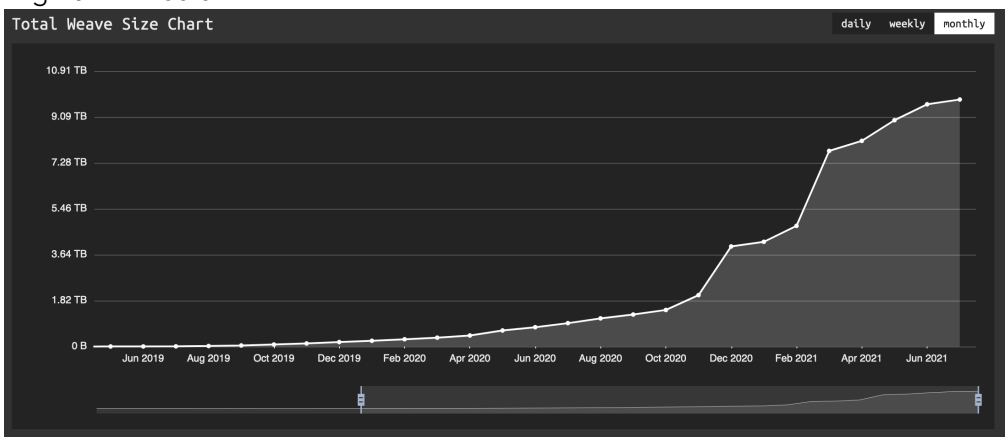
that could ensure storage truly is permanent. Of course, neither Dropbox, Google, or even Filecoin or Sia, are designed for permanent storage.

Due to the endowment model, Arweave actually wants the cost of storage to fall, whereas its competitors prefer higher storage prices for more revenues.



Source: Ardrive.

Arweave's storage capacity and volume have been rapidly increasing over the past year. There were 1.4 million transactions in June 2021, and the blockweave is currently permanently storing 10 TB in data.



Source: Viewblock.io.

Permanent Storage For DeFi & NFTs

While storing and providing access to data that has otherwise been censored is an admirable service, Arweave's volume to date, has been driven by less politically charged use cases: NFTs and the data storage requirements of other Layer-1 blockchains.

Blockchains have increasing data storage requirements and no ability or incentive to provide storage. Arweave's first archive bridge, Solar, was built to handle Solana's entire ledger

history. One potential issue with Arweave is that the protocol optimizes for permanent storage as opposed to the fast information retrieval times. Decentralized start-up, Kyve, allows the blockchains to pull data quickly off Arweave to sync their respective chains and clients now include Terra, Solana, Polkadot, and Avalanche.

The second driver of demand and less obvious use case has been using Arweave as the on-chain and permanent repository of NFT data, including images and music.

Source. Twitter.com

Because Ethereum wasn't designed for data storage, most NFTs hold only a small amount of data on-chain, and include a link, often using IPFS, to a centralized storage location for the actual item. Early NFT buyers have sometimes learned the hard way that, while IPFS can act as a decentralized router, and direct the user to the data if it exists, the data itself can easily disappear. IPFS didn't lose the above user's NFTs, but the pointer (reference to the data on IPFS) was no longer valid because it wasn't maintained. This is, in part, why companies like Filebase and Pinata have been created to provide users the ability to pay a provider to store information on IPFS.

Arweave solves these issues for NFTs by providing a permanent location for the metadata and the data, which can be accessed through IPFS or through Arweave directly. Owners of Hashmasks, for example, can be assured that the storage of the actual art, linked through IPFS, is unlikely to be lost anytime soon. Top NFT marketplace OpenSea allows NFT minters to link data through IPFS to both Filecoin and Arweave. Other platforms and creators are also using Arweave as their storage provider, including Beeple's NFT marketplace and Gary Vee's Veefriends.

Going forward, atomic NFT exchange on Arweave allows for *all* of the relevant data to reside inside the NFT as a smart contract, effectively tokenizing the data itself as the NFT.

The Permaweb And its Ecosystem

Storage is only one element of Web3 and, like Sia with Skynet – a decentralized file sharing and content distribution protocol – Arweave has incentivized protocol creators to build on its "Permaweb".

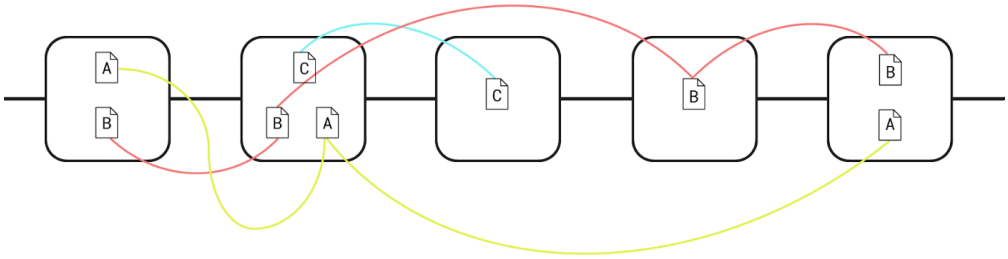
Developers have also created tools such as simple indexing services that allow individuals to query the Arweave network through browsers such as Brave.

Arweave's storage is only one element of its wider Web3 ecosystem. In order to use the permaweb to its fullest, Arweave has either built or is incentivizing others to build on top of the storage network.

Smartweave allows the blockweave to execute smart contracts by what is known as "lazy consensus". To verify a transaction in a smart contract, the user must verify the entire transaction history available inside the blockweave.

Arweave has adopted a unique approach to developing its permaweb, supporting and empowering DAO-like profit sharing communities by offering two years of AR token rewards.

Email (weavemail), Dropbox-like storage (Ardrive, considered one of the best blockchain use cases of 2020 by the World Economic Forum), and a decentralized Medium (Scribe) are just a few of the DApps that use Arweave's data storage chain at their center. Mirror.xyz, an upcoming decentralized and user-owned publishing platform, hosts all user blogs on the blockweave to ensure that censorship is near-impossible.



Source: Arweave.medium.com.

The AR Token

Since the relevant cost for miners is the cost per GB hour, AR token's value can be implied by the amount of tokens demanded by miners to add a block. Currently, the cost of a block is around $130 and revenue is running at $150 per block.

Currently, tokens are released to the network when the cost to miners is greater than mining rewards sourced from the endowment.

Amplify allows users to buy and sell bandwidth, or stake tokens to run a full node – requiring TBs of data storage capacity – or a dapp-specific node. Stakers receive rewards from the endowment and minting for storage, or fees from users for dapp-specific nodes. Staking is risky: nodes can have their staked AR slashed for failing to operate as expected.

55 million "Genesis" AR were created on June 8, 2018, and a further 11 million will be used as block mining incentives.

Conclusion

At first glance, Arweave is a decentralized storage protocol in the same vein as Filecoin, Storj, and Sia, but it sets itself apart by focusing on data permanence. This unique approach makes Arweave very attractive to emerging uses cases like NFTs as well as existing companies such as the Internet Archive that require perpetual storage.

Some key questions still remain. Is data permanence truly achievable? Will Arweave's model hold up against AR value volatility? While the endowment currently has enough AR for 1,700 years of storage, will their forecasts remain accurate to new changes in technology and an exponentially increasing amount of data?

It is also questionable just how available the data is to those facing government restrictions. BitTorrent, for example, has not been as anonymous a data provider as expected. For illegal or sensitive data, will miners police themselves properly? Or perhaps they will err on the side of caution, acting as censors themselves? Arweave is working with the Internet Watch Foundation to identify the best ways to filter out abusive content.

Can Storage Ever be Truly Decentralized?

Storage in Web3 appears unlikely to exist without a robust ecosystem, and therefore there are likely to be significant network effects. This is where the real battle for Web3 lies. It's a sort of recentralization: "come join my ecosystem and build on it". Sia has Skynet, Arweave has the Permaweb. IPFS is a growing standard.

The Web2 internet has converged to one set of standards with a few dominant (centralized) platforms. Is this the inevitable future for Web3?

If big is what currently matters, Filecoin should already be declared the winner, with almost 20 times the storage of its next biggest competitor, while accounting for a full 79% of the total storage solutions market capitalization.

Yet Arweave does have a unique offering, and market capitalization isn't the determining factor for network growth. Through its unique focus on data permanence, Arweave might be able to carve out its niche that potentially blossoms in the future.

Of course, having one dominant storage provider is perhaps suboptimal for the fully-decentralized Web3 of the future. If we end up with multiple storage solutions and a corresponding number of blockchains, Arweave's success to date is encouraging.