

Canyon: Permanent Storage Layer for Limitless Scalability

Liu-Cheng Xu

Abstract

Canyon is a permanent storage network built on Substrate, which records the hashes of files on chain and stores the files off-chain. By blending PoS and a probabilistic proof-of-storage scheme inspired by Arweave, Canyon greatly reduces the barriers to entry for miners, incentivizing them to store as much data as possible for the block rewards.

Contents

1	Introduction	2
1.1	Motivation	2
2	Background	2
2.1	Filecoin	2
2.2	Crust	2
2.3	Arweave	2
2.4	Design purposes	3
2.4.1	Security and privacy	3
3	System design	3
3.1	Consensus	3
3.1.1	Proof of Access	3
3.1.2	Proof of Stake	4
3.2	Economy Model	4
3.2.1	Transaction Fee	4
3.2.2	Data Oblivion	4
3.2.3	Payments	4
4	Roadmap	4
5	Conclusion	4
5.1	Future work	4
A	Appendix	5

1 Introduction

1.1 Motivation

TODO

2 Background

2.1 Filecoin

Filecoin[1] proposes a sophisticated cryptographic solution based on zero-knowledge proofs (ZKPs) to prevent the common attacks on the decentralized storage verification, which uses pure mathematical methods and is able to achieve high security guarantees. The whole mining process consumes too much computing power, making the actual storage cost prohibitively expensive. Furthermore, the high hardware requirements largely raise the threshold for small miners, preventing these with pure common commodity hardwares from joining the network, the storage distribution today becomes centralized.

Due to the lack of authentic storage needs and system design of Filecoin, the miners themselves are economically impelled to store tons of garbage data in the network, increasing the storage power and earning more mining rewards. Despite the Filecoin team has proposed the off-chain governance approach like Filecoin Plus¹, it does not mitigate this problem substantially, leaving the promotion of the useful storage still a huge unresolved challenge in the Filecoin network.

2.2 Crust

To solve the problem of decentralized storage verification, Crust[2] adopts a hardware-based solution Trusted Environment Execution (TEE) to make sure the miners store a specific number of data copies as promised. Each of the storage nodes is required to enroll on Crust chain through TEE before it's allowed to deal with the storage orders from client. The TEE module of the nodes will periodically check and report whether the files are properly stored in the local storage space in a trusted way.

There are three major TEE providers with different implementations: Software Guard Extensions(SGX) on the Intel platform, Secure Encrypted Virtualization (SEV) on the AMD platform, and TrustZone on the ARM platform. The SGX of Intel is the mostly widely used TEE platform. Although Crust has a relatively low hardware demand for the mining compared with Filecoin, the miners are heavily dependent on a fairly narrow range of hardwares that support TEE, thus being greatly affected by the hardware manufacturers.

2.3 Arweave

Unlike the ephemeral storage services such as Filecoin, Crust, Arweave[3] serves as a permanent storage layer using a probabilistic and incentive-driven approach to maximize the number of redundant copies of any individual piece of data in the network, which fills in a crucial aspect of decentralized storage need in Web 3.0. The natural feature of perpetual storage is perfectly suitable for the NFTs, and is the cornerstone of new storage-based computation paradigm like everFinance².

Without the periodical audit of the data replications, Arweave is much effortless than the other solutions in terms of the constraint and cost of storage consensus. Nevertheless, the deficiency

¹<https://docs.filecoin.io/store/filecoin-plus>

²<https://medium.com/everfinance/a-storage-based-computation-paradigm-enabled-by-arweave-de799ae8c424>

of this scheme is the potential risk of data centralization that ultimately there possibly will be a single storage provider serves the whole data. The Arweave team has introduced and deloppyed an improved version of consensus Succinct Proofs of Random Access, attempting to disincentivize miners from retrieving data on demand from the network.

Canyon network is profoundly inspired by Arweave, aiming to be a permanent decentralized storage network using Substrate framework. The key contribution of Canyon is that it adapts PoS into the storage consensus of Arweave, whereas Arweave uses PoW, making Canyon network more scalable and environment-friendly.

2.4 Design purposes

2.4.1 Security and privacy

3 System design

3.1 Consensus

3.1.1 Proof of Access

$$P(\text{win}) = P(\text{has recall block}) * P(\text{finds hash first}) \quad (1)$$

$$P(\text{win}) = P(\text{has recall block}) * P(\text{claims slot}) \quad (2)$$

Algorithm 1: Generation of POA

Input :

The random seed S ;

The weave size W ;

Output:

The proof of accessing the recall block POA ;

1 Initialize the number of repeats x with 1;

2 **repeat**

3 Draw a random byte B with $\text{MULTIHASH}(S, x) \bmod W$;

4 Find the TX in which the random byte B is included;

5 $x \leftarrow x + 1$;

6 **until** *The data of TX is available*;

7 $POA \leftarrow \text{CONSTRUCTPOA}(TX)$;

8 **return** POA ;

$$\hat{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

$$R = \frac{1}{\hat{x}_{N \rightarrow +\infty}}$$

3.1.2 Proof of Stake

3.1.2.1 Staking Rewards

3.1.2.2 Stake

3.2 Economy Model

3.2.1 Transaction Fee

The fee of data storage transaction is composed of the perpetual storage cost and one-shot bandwidth cost. The perpetual storage cost is basically modeled after Arweave.

3.2.2 Data Oblivion

TODO

3.2.3 Payments

Canyon has an inherent advantage of easier interoperability with the Polkadot ecosystem by using Substrate, which is the same framework Polkadot is built on. Canyon will support the users pay the storage fee using the stable coins like USDT.

4 Roadmap

TODO

5 Conclusion

5.1 Future work

References

- [1] Protocol Labs, Filecoin: A Decentralized Storage Network, <https://filecoin.io/filecoin.pdf>.
- [2] <https://crust.network/>
- [3] Sam Williams, Viktor Diordiiev, Lev Berman, India Raybould, Ivan Uemlianin. Arweave: A Protocol for Economically Sustainable Information Permanence.
- [4] Gavid wood. Polkadot: Vision for A Heterogeneous Multi-chain Framework. <https://polkadot.network/PolkaDotPaper.pdf>
- [5] <https://github.com/paritytech/substrate>
- [6] Vitalik buterin. An Incomplete Guide to Rollups, <https://vitalik.ca/general/2021/01/05/rollup.html>
- [7] A Storage-based Computation Paradigm Enabled by Arweave, <https://medium.com/everfinance/a-storage-based-computation-paradigm-enabled-by-arweave-de799ae8c424>
- [8] <https://ever.finance/>

A Appendix