FileDES: A Secure, Scalable, and Succinct Blockchain-based Decentralized Encrypted Storage Network

Minghui Xu[1], Jiahao Zhang[1], Hechuan Guo[1], Xiuzhen Cheng[1], Dongxiao Yu[1], Qin Hu[2], Yijun Li[3], Yipu Wu[3]
[1]Shandong University
[2]Indiana University-Purdue University Indianapolis
[3]BaishanCloud

Presented by: Jiahao Zhang

2024/6/4

山东大学 SHANDONG UNIVERSITY

IUPUI

白山云科技 BAISHANCLOUD

Let’s Start with a Simple Question:

➢ If you have a 1TB hard drive gathering dust, what would you do with it?

❑ Do nothing, I will use it in the future. OK!

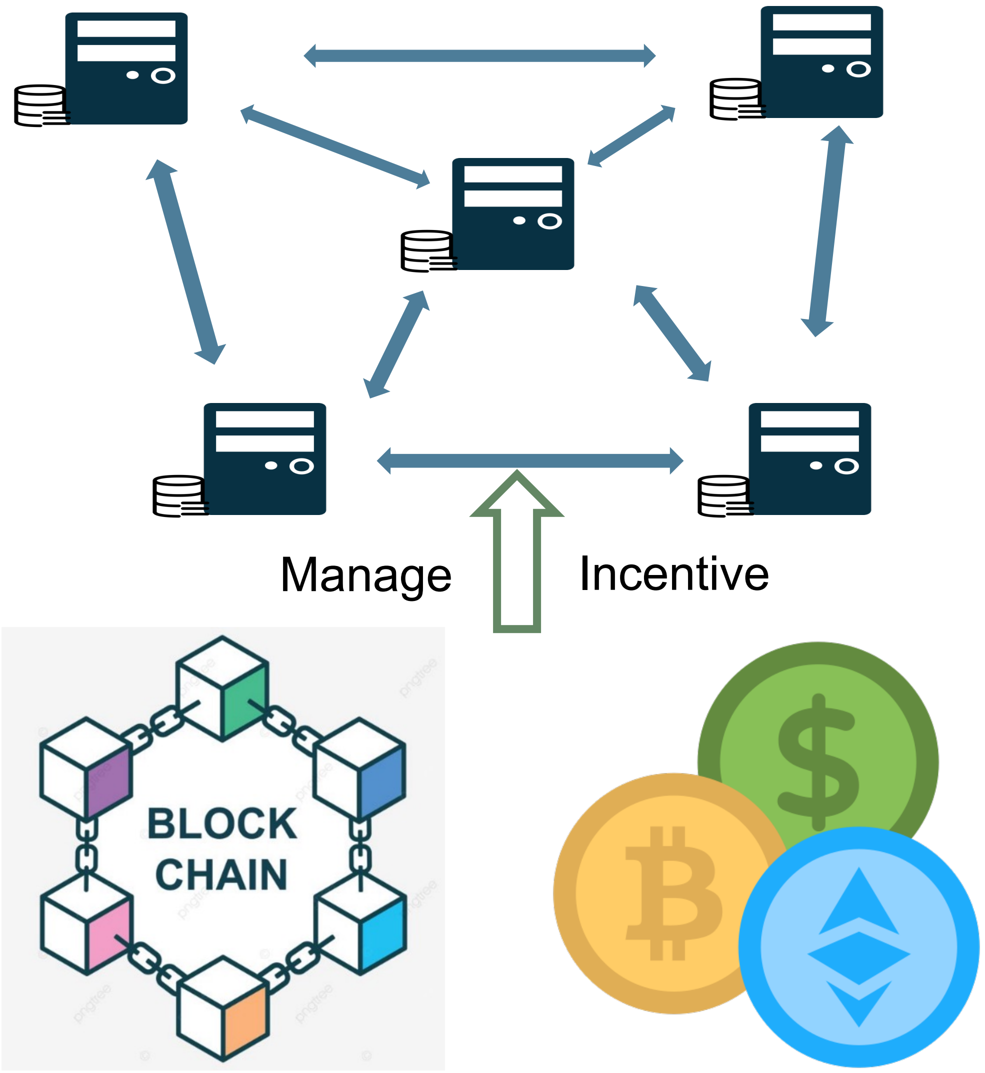
❑ Share it with my friends. Great!

❑ Share it with anyone in the world and get reward.

This is what decentralized storage networks (DSNs) want to do.

Background

➢ Decentralized Storage Network (DSN)

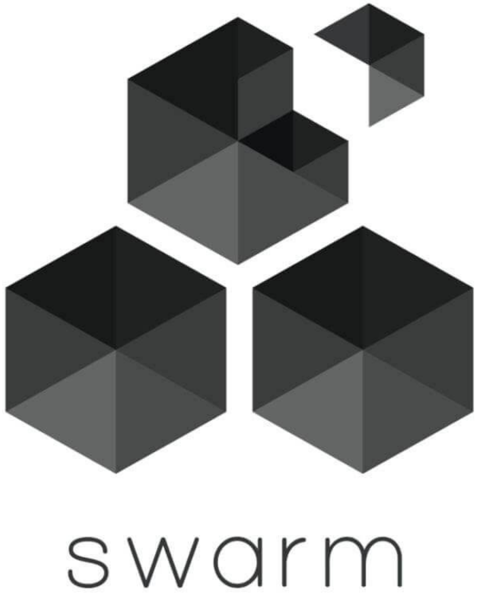


- A P2P network to aggregate available storage space from independent storage providers.

- Blockchain act as a manager and incentive layer to encourage storage providers to provide security and robust storage services.

Background

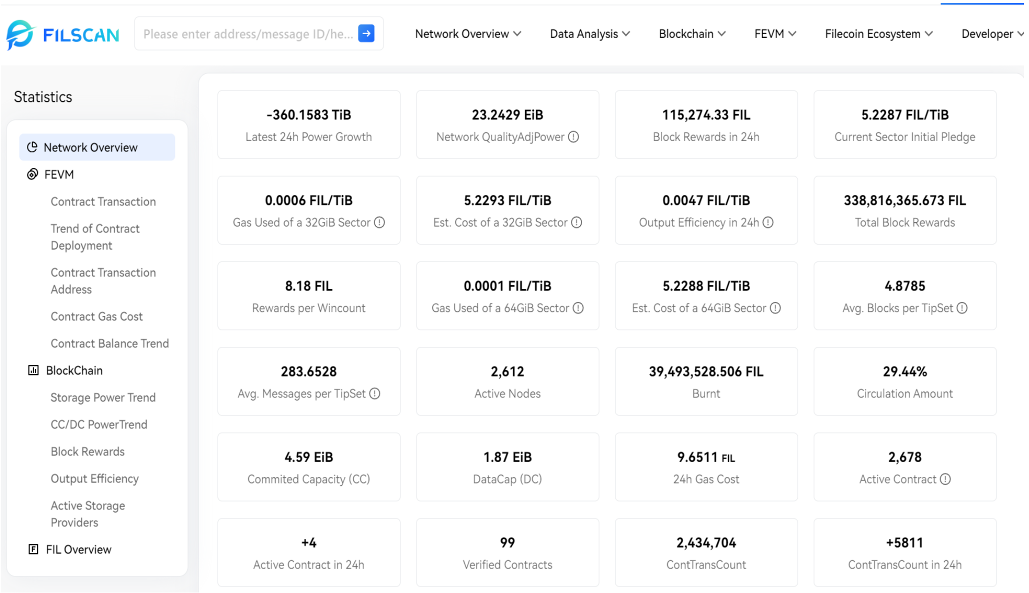
- **Popular DSN projects**



Background

➢ Current status of DSN

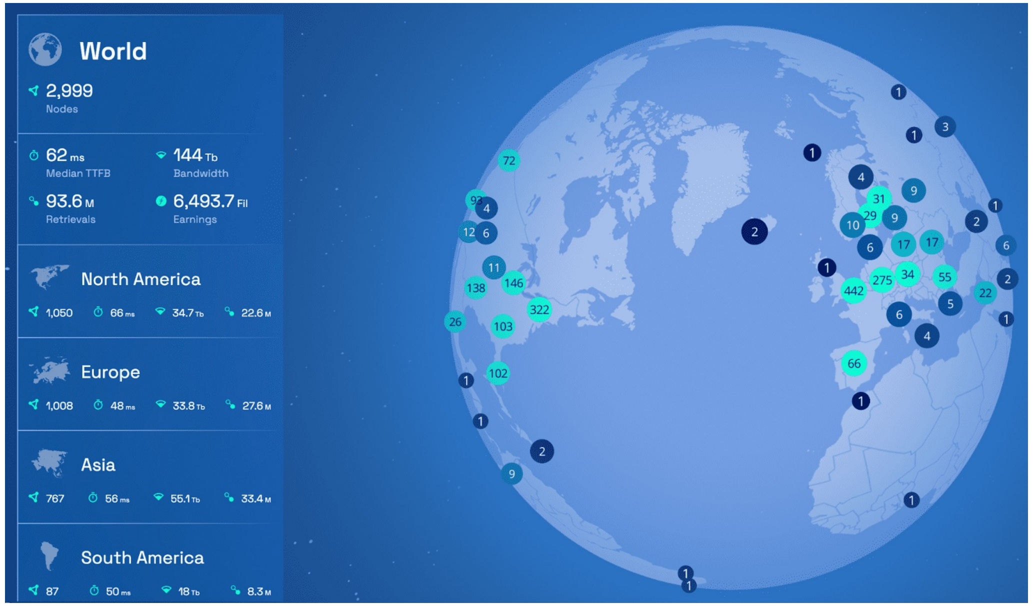
- ❑ The mainnet of Filecoin launched in 2020.
- ❑ More than 20 EiB of storage capacity.
- ❑ More than 2500 storage providers distributed around the world.



Statistics

- Network Overview
- FEVM
  - Contract Transaction
  - Trend of Contract Deployment
  - Contract Transaction Address
  - Contract Gas Cost
  - Contract Balance Trend
- BlockChain
  - Storage Power Trend
  - CC/DC PowerTrend
  - Block Rewards
  - Output Efficiency
  - Active Storage Providers
- FIL Overview

-360.1583 TiB Latest 24h Power Growth	23.2429 EiB Network QualityAdjPower ⓘ	115,274.33 FIL Block Rewards in 24h	5.2287 FIL/TiB Current Sector Initial Pledge
0.0006 FIL/TiB Gas Used of a 32GiB Sector ⓘ	5.2293 FIL/TiB Est. Cost of a 32GiB Sector ⓘ	0.0047 FIL/TiB Output Efficiency in 24h ⓘ	338,816,365.673 FIL Total Block Rewards
8.18 FIL Rewards per Wincount	0.0001 FIL/TiB Gas Used of a 64GiB Sector ⓘ	5.2288 FIL/TiB Est. Cost of a 64GiB Sector ⓘ	4.8785 Avg. Blocks per TipSet ⓘ
283.6528 Avg. Messages per TipSet ⓘ	2,612 Active Nodes	39,493,528.506 FIL Burnt	29.44% Circulation Amount
4.59 EiB Commited Capacity (CC)	1.87 EiB DataCap (DC)	9.6511 FIL 24h Gas Cost	2,678 Active Contract ⓘ
+4 Active Contract in 24h	99 Verified Contracts	2,434,704 ContTransCount	+5811 ContTransCount in 24h



World

2,999 Nodes

62 ms Median TTFB

144 Tb Bandwidth

93.6 M Retrievals

6,493.7 Fil Earnings

North America: 1,050 · 66 ms · 34.7 Tb · 22.6 M

Europe: 1,008 · 48 ms · 33.8 Tb · 27.6 M

Asia: 767 · 56 ms · 55.1 Tb · 33.4 M

South America: 87 · 50 ms · 18 Tb · 8.3 M

Background

➢ Current status of DSN

- ❑ A valuable storage infrastructure for application in Web3.
- ❑ Providing reliable decentralized storage for non-fungible tokens (NFTs).
- ❑ Providing decentralized, scalable and crypto token incentivized video streaming management.
- ❑ Metaverse, DeFi…

Background

➢ Components of DSN

Clients: Pay tokens to use storage services.

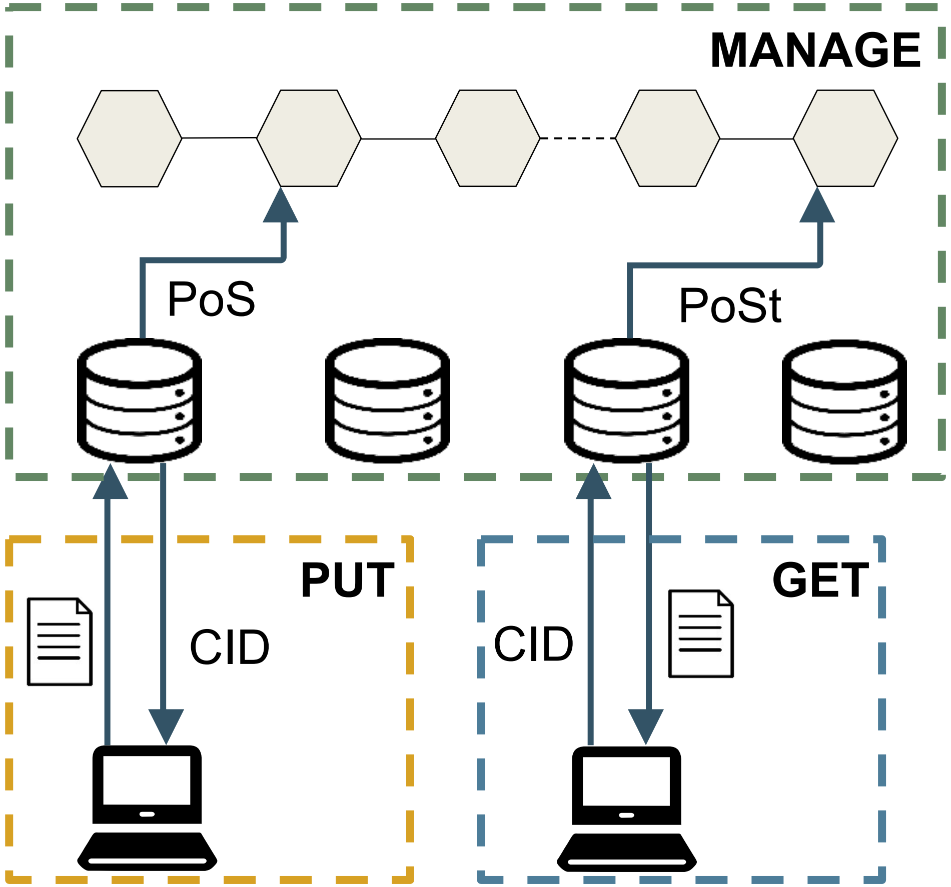
Miners: Earn tokens by keeping files safe and mining new blocks.

- ❑ **Content Identifier (CID):** an unique identifier that locates a file in the network.
- ❑ **Proof-of-Storage (PoS):** a cryptographic protocol to confirm a miner have correctly stored a file ones the miner have received a file uploaded by a client.
- ❑ **Proof-of-Spacetime (PoSt):** a cryptographic protocol to confirm a miner have correctly stored a file for a specific period.

Background

➢ Basic Protocols

- ❑ A DSN consists of three protocols:
  - **PUT:** Clients execute the PUT protocol to upload the file to a miner in a DSN, and obtain the CID of the file.
  - **MANAGE:** Miners execute the MANAGE protocol to make sure the files are stored correctly and to prevent any issues.
  - **GET:** Clients execute the GET protocol to send a CID to the DSN and retrieve the corresponding file from miners.



Challenge Statement

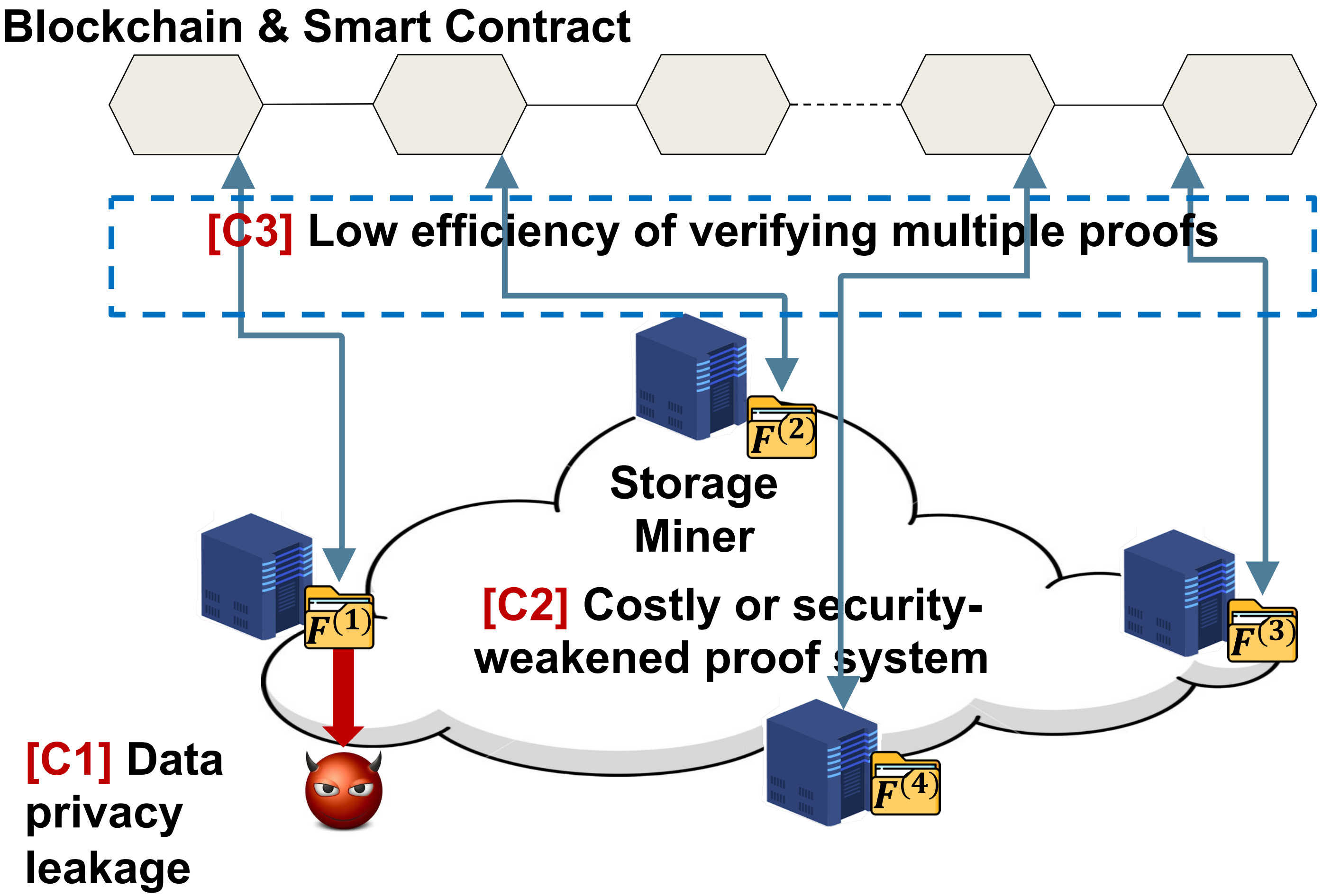
➢ **Three major challenges faced by DSNs**

❑ 1. Data privacy leakage

❑ 2. Costly or security-weakened proof system

❑ 3. Low efficiency of verifying multiple proofs

Affect the performance and security of DSNs.


Blockchain & Smart Contract
[C3] Low efficiency of verifying multiple proofs
$F^{(2)}$
Storage Miner
[C2] Costly or security-weakened proof system
$F^{(1)}$
$F^{(3)}$
$F^{(4)}$
[C1] Data privacy leakage

Challenges

➢ Data Privacy Leakage

❑ Plaintext storage

  ❑ Straightforward design

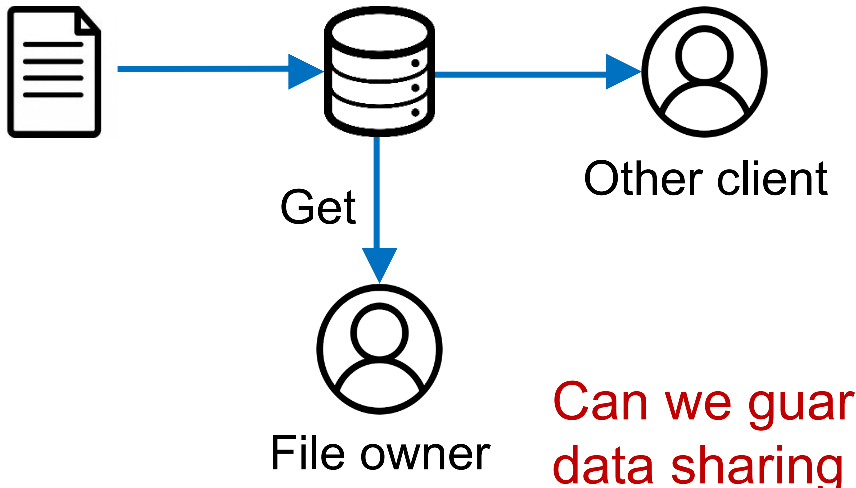
  ❑ Suitable for storing non-sensitive files

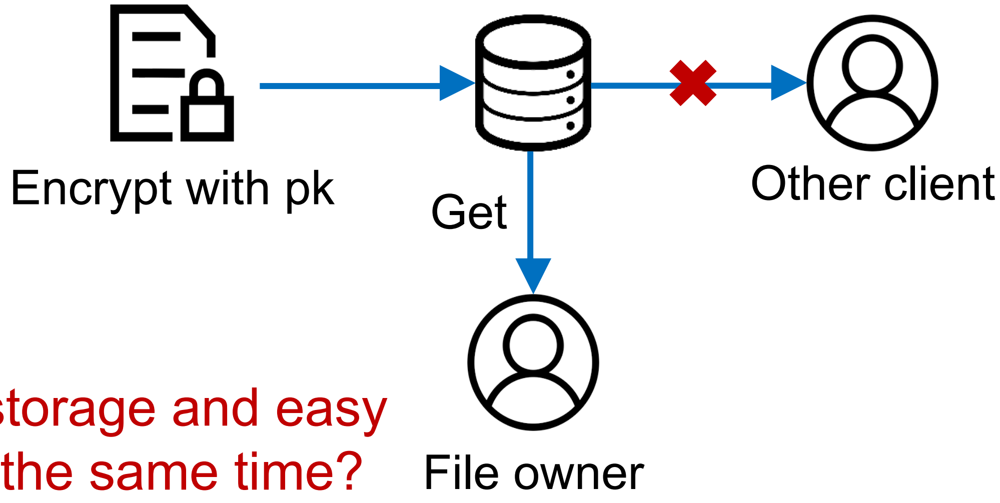
  ❑ Suffers data privacy leakage problem

❑ Simple encryption storage

  ❑ More secure than plaintext storage for sensitive files

  ❑ Harms the data availability and hinders data sharing



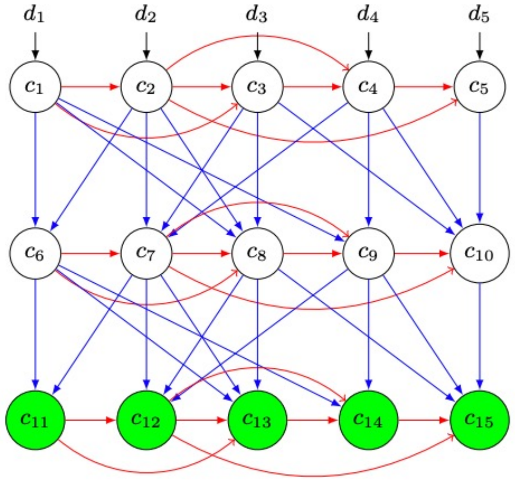


Can we guarantee secure storage and easy data sharing capabilities at the same time?

Challenges

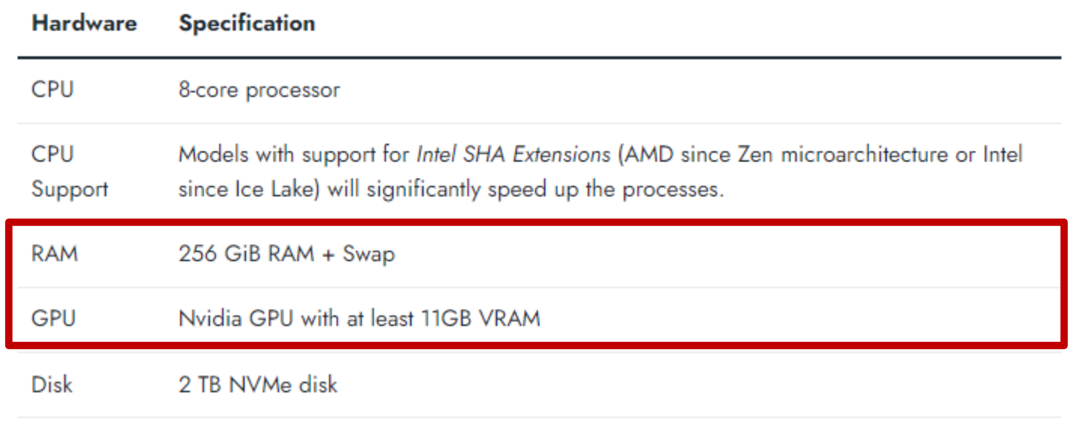
➢ Costly or security-weakened proof system

- For DSNs using plaintext storage, generating a PoS can be a time-intensive and hardware-demanding to prevent Sybil and Generation attacks.
- In Filecoin, PoS relies on complicated Stacked depth robust graph (SDRG).
- Miners in Filecoin needs ~4h to process a 32GB file with high hardware configuration.



Ben Fisch, “Tight proofs of space and replication”, EUROCRYPT’19

Stacked depth robust graph



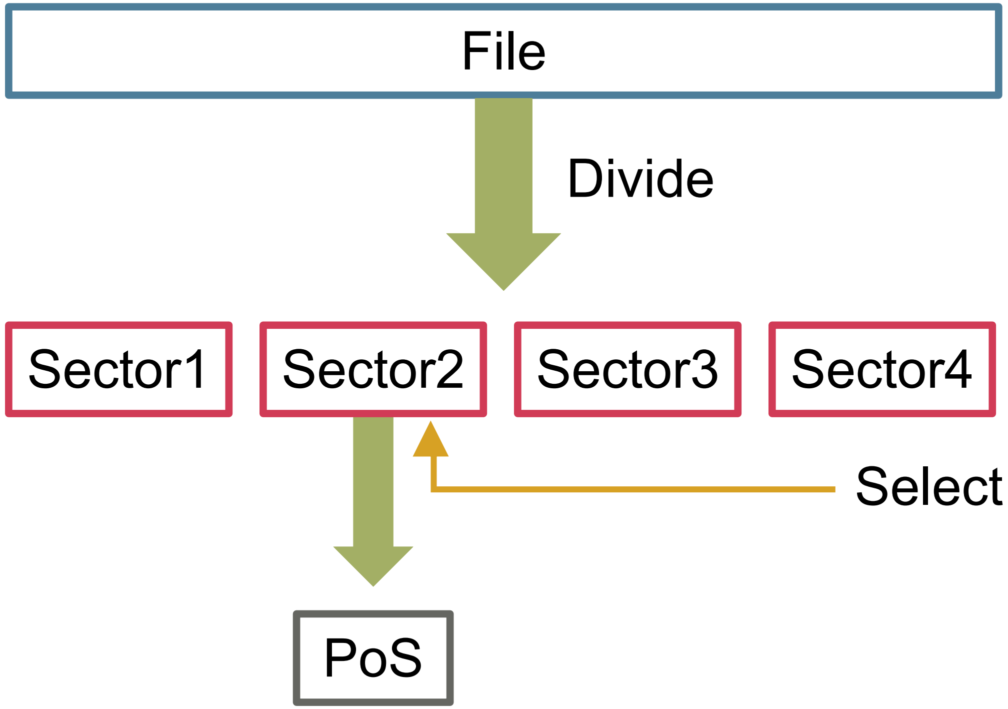
Hardware	Specification
CPU	8-core processor
CPU Support	Models with support for *Intel SHA Extensions* (AMD since Zen microarchitecture or Intel since Ice Lake) will significantly speed up the processes.
RAM	256 GiB RAM + Swap
GPU	Nvidia GPU with at least 11GB VRAM
Disk	2 TB NVMe disk

Hardware requirements of Filecoin

Challenges

➢ Costly or security-weakened proof system

- ❑ For DSNs using simple encryption storage, PoS generation process efficiency but sacrifice security.
- ❑ In Sia, files are divided into 256KB sectors and PoS of a file is only provided using the data in a randomly select sector.



PoS generation process of Sia

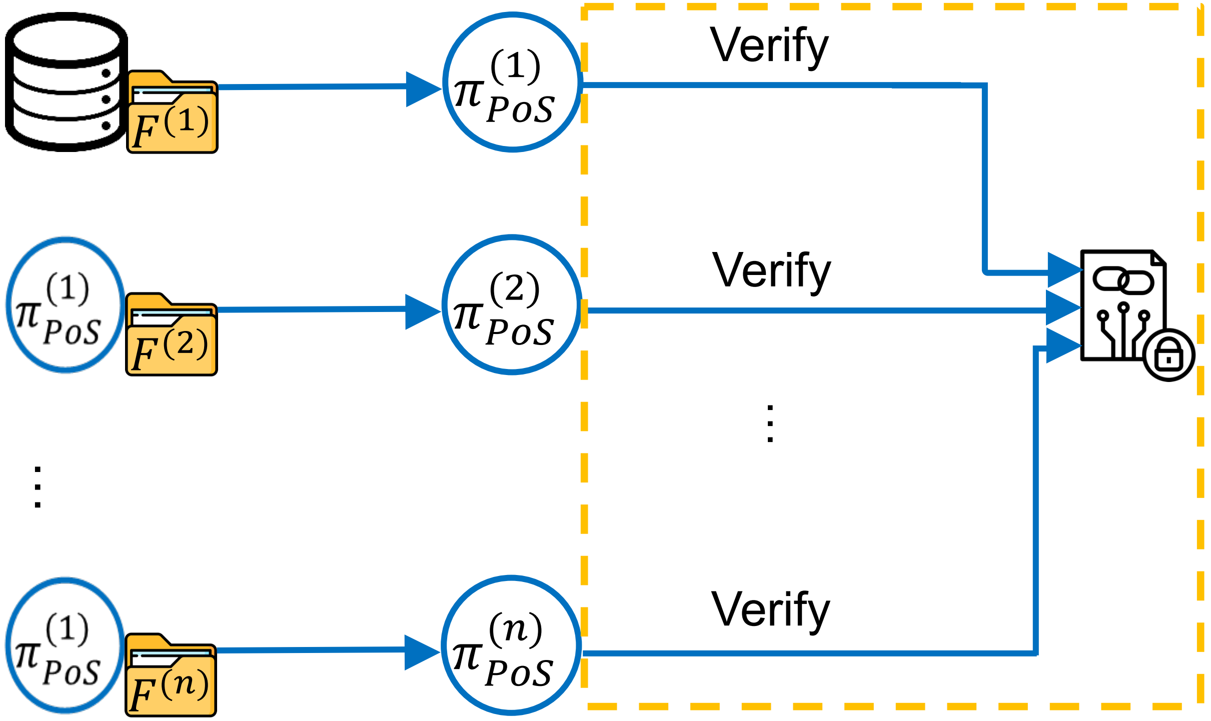
Can we make the proof system efficiency and secure at the same time?

Challenges

➢ Low efficiency of verifying multiple proofs

- ❑ DSNs needs to recurrently verify the PoS and PoSt of each file to ensure their correct storage.
- ❑ The number of PoS/PoSt needed to be verified increases linearly with total number of files in DSN, resulting in significant computational burden.
- ❑ Managing multi-version files in DSN is complex.

Verify n times for n files



Verification of multiple proofs in DSN

Can we reduce the computational and verification workload of proof system?

FileDES

➢ Technique 1:Encrypted Storage

- Technique Sketch:

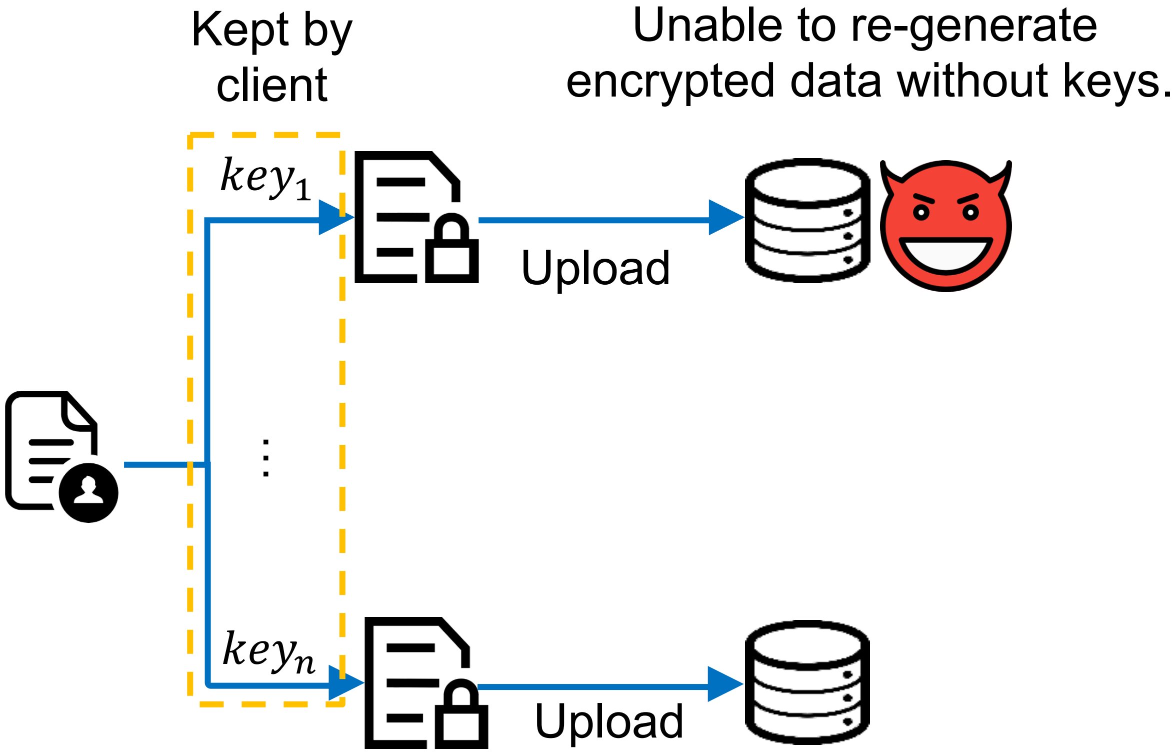
❑ Before upload, a client uses multiple keys to encrypt a file to create different replicas of the file.

❑ Encrypted files are uploaded to miners, and the keys used in encryption are kept by the client.

- Advantages:

❑ Encryption enhances data privacy.

❑ A malicious miner is unable to reproduce the replicas even with plaintext. Preventing Sybil and Generation attacks.



The sketch of encryption storage

FileDES

➢ Technique 1:Encrypted Storage

- For files that can be shared directly.

❑ RSA-based encryption

  ❑ A file is encrypted using secret key.

  ❑ Other clients can directly use the shared public key to get the file.

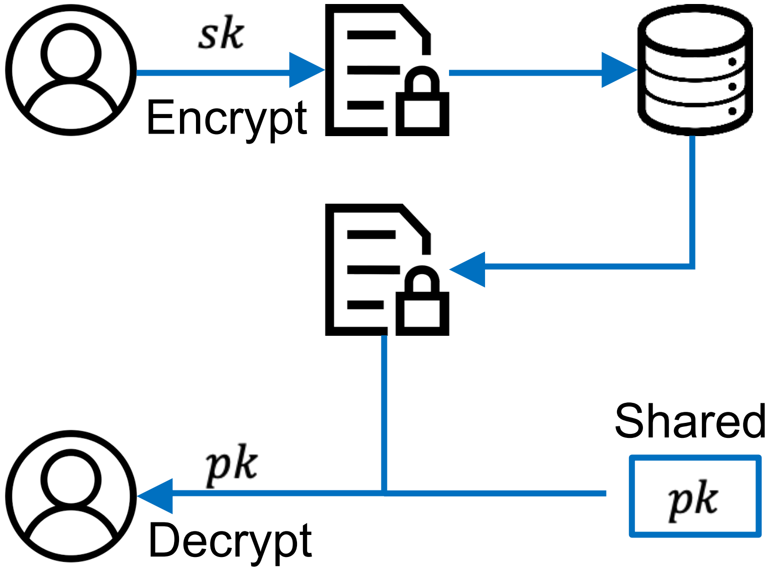
- For files with private data.

❑ Unidirectional Proxy Re-Encryption (PRE)

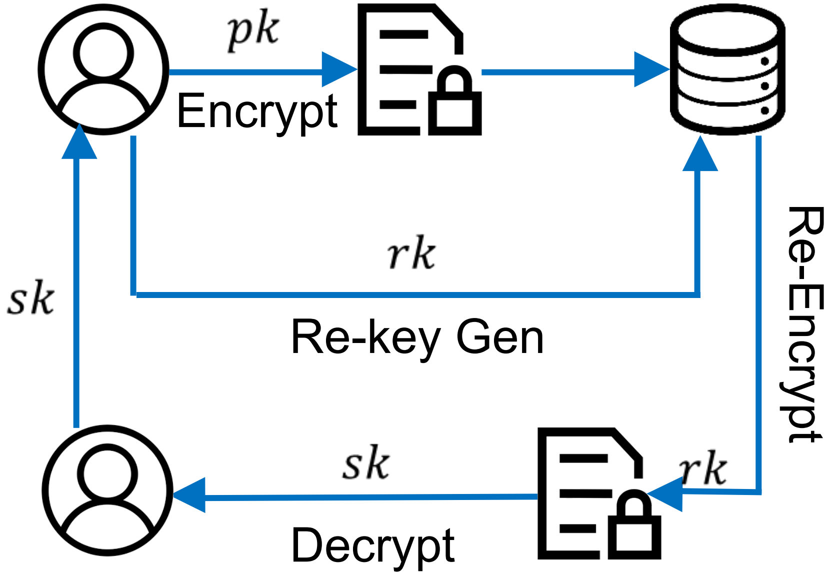
  ❑ A file is encrypted using public key.

  ❑ Miners act as proxies. PRE ensures the proxy can’t see plaintext.

  ❑ Other clients need to ask for permission to access the file.



RSA Overview



PRE Overview

FileDES

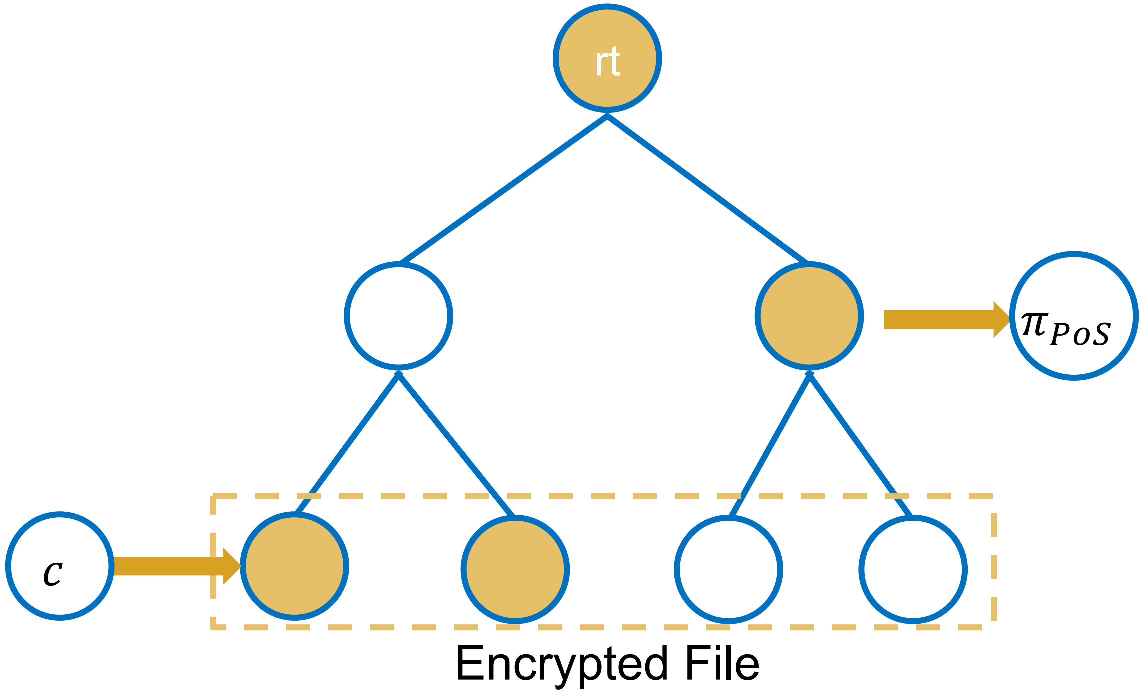
➢ Technique 2: Proof of Encrypted Storage (PoES)

- Generate PoS efficiently

❑ Technique 1 have already ensure the prevention of Sybil and Generation attacks.

❑ The time-intensive and hardware-demanding process of miners can be abandoned.

❑ To generate a PoS, the only thing a miner should do is to prove the integrity of an encrypted file (e.g. Merkle Proof).



The sketch of PoS generation in FileDES

FileDES

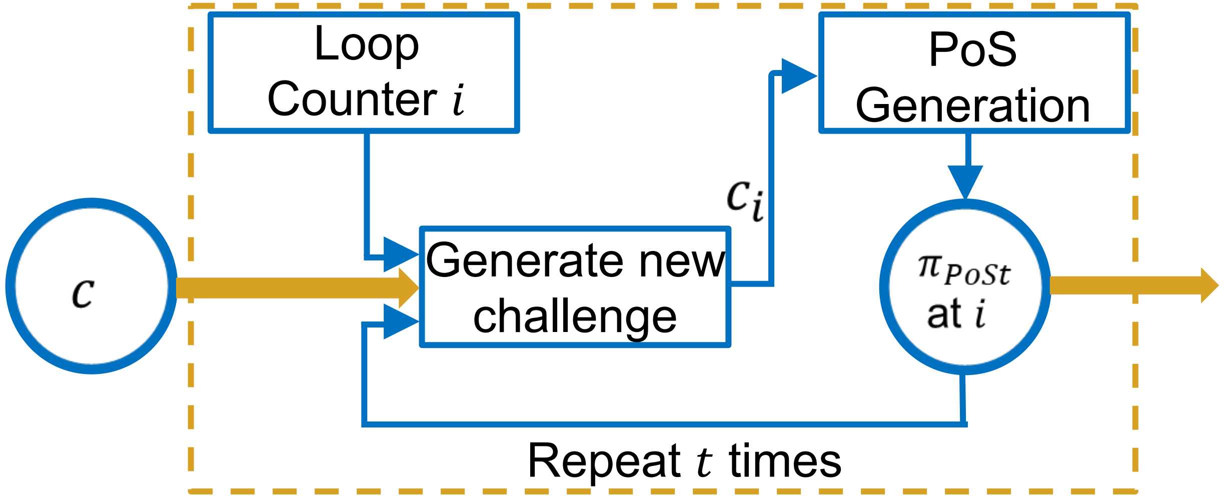
➢ Technique 2: Proof of Encrypted Storage (PoES)

- Generate PoSt efficiently

❑ Making a miner frequently generate PoS for a file can force the miner correctly store the file over a period of time.

❑ It is adequate to conduct periodic spot checks within short timeframes of this range.

❑ Our sketch: Recurrently generate sequential PoS.

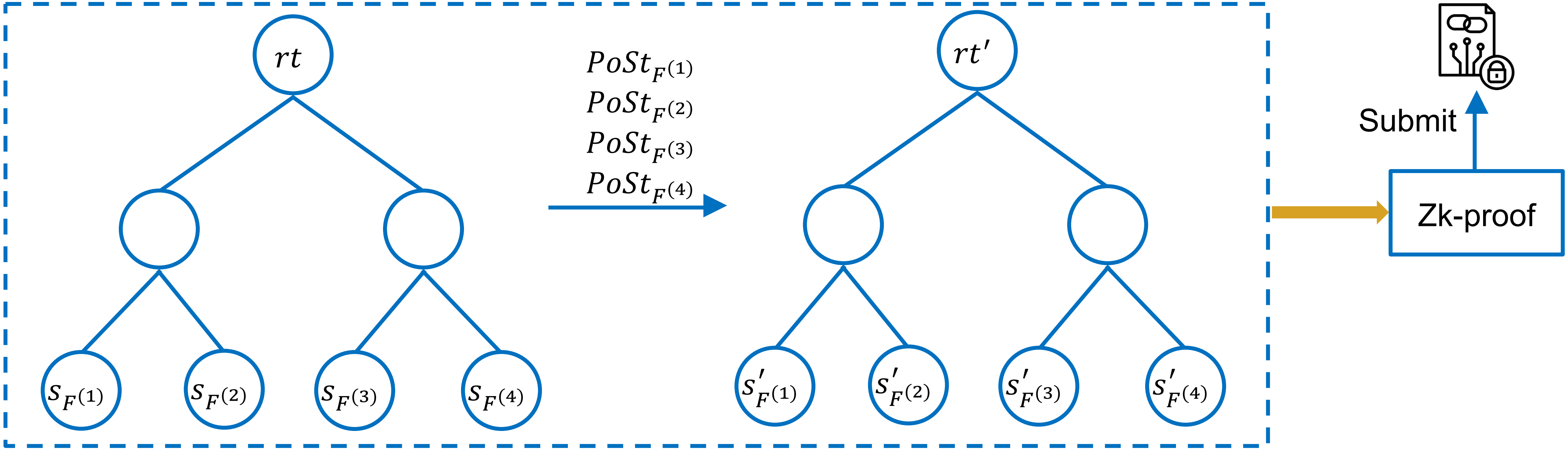


The sketch of PoSt generation in FileDES

FileDES

➢ Technique 3: Batch Verification of PoS and PoSt

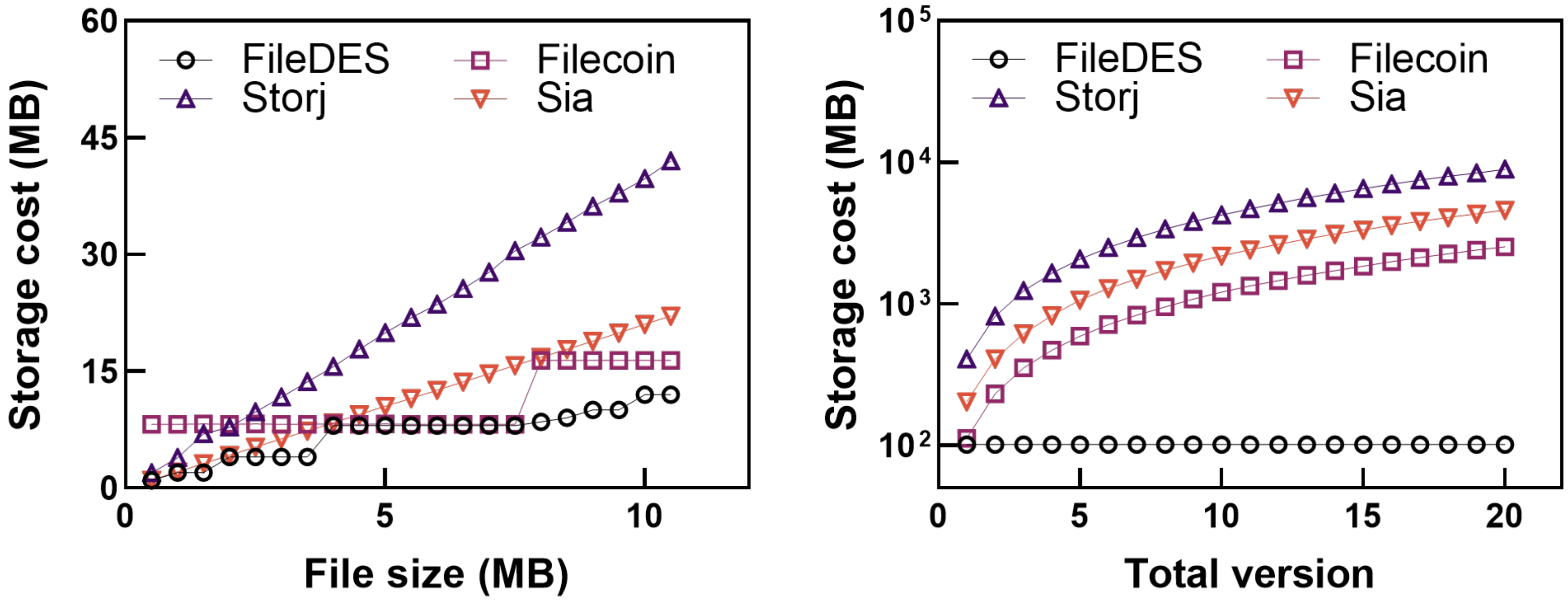
- ❑ Following the basic idea of rollup: transfer the verification of multiple proofs to an aggregated succinct proof.
- ❑ PoS and PoSt can be used as evidence to indicate the storage state change of a file.
- ❑ The whole process can be represented using a zk-circuit to generate a succinct proof following the sketch of zk-rollup.



Batch verification sketch in FileDES

FileDES

➢ Storage Cost of a Single File and Multi-Version Files



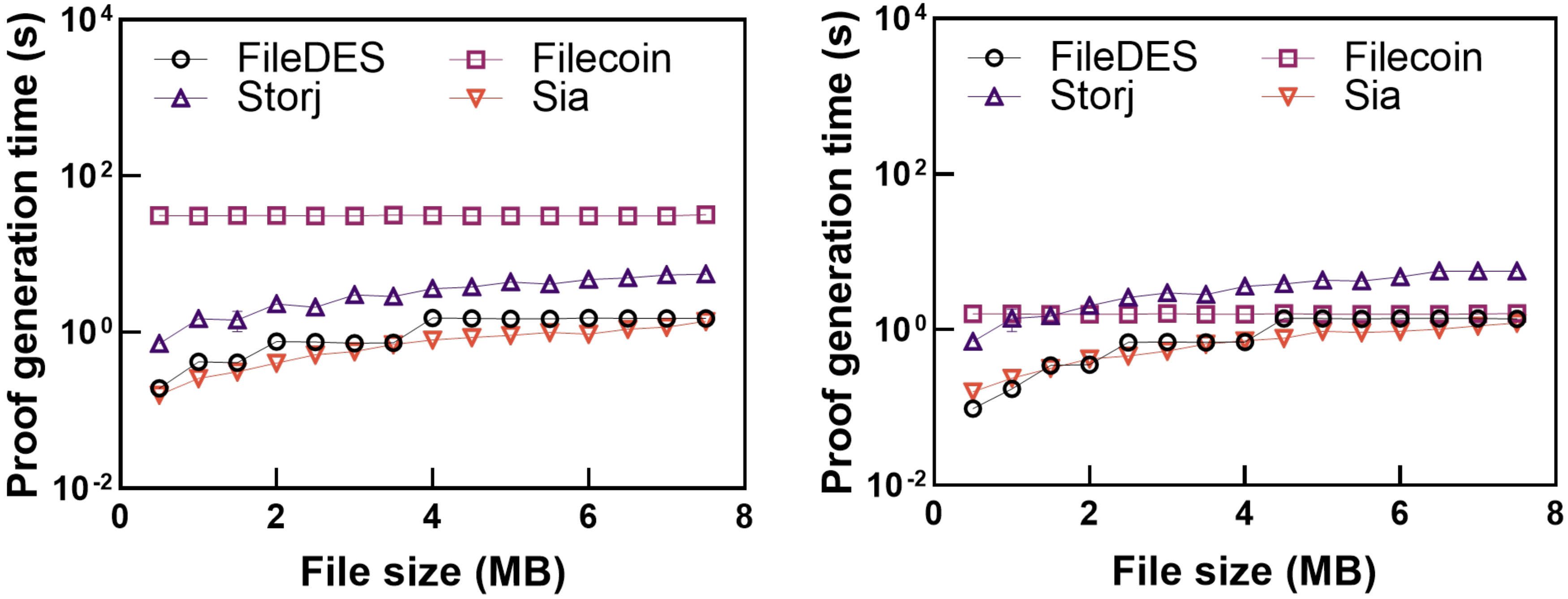
(a) Single Files

(b) Multi-Version File

FileDES has the lowest storage cost for storing single files.
FileDES has the lowest storage cost when saving multi-version files as we use file increment.

FileDES

➢ PoS/PoSt Generation time for Files with Different Sizes



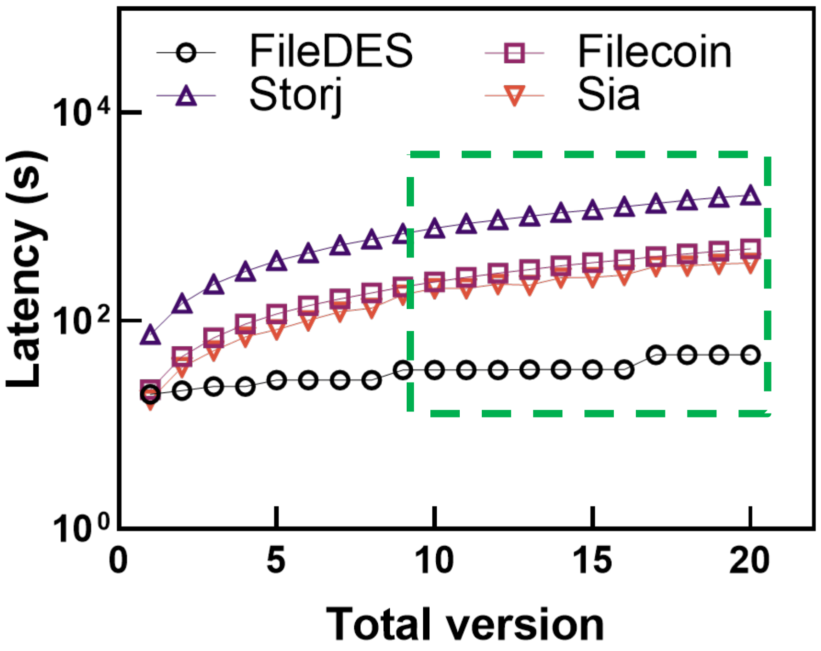
(a) PoS

(b) PoSt

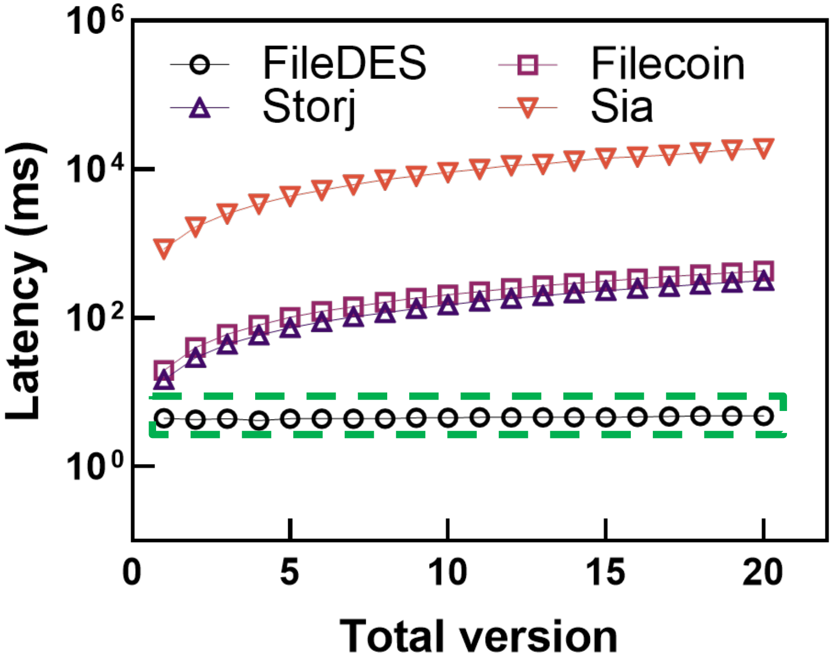
The PoS/PoSt generation times in FileDES are faster than Filecoin and Storj and comparable to Sia.

FileDES

➢ Proof Generation and Verification Time of a Multi-Version File



(a) Proof Generation

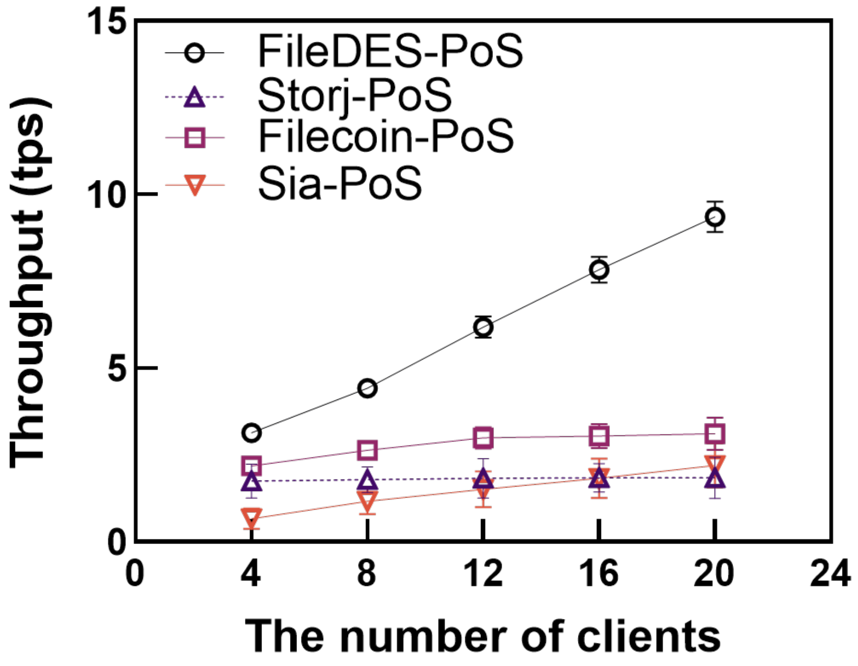


(b) Proof Verification

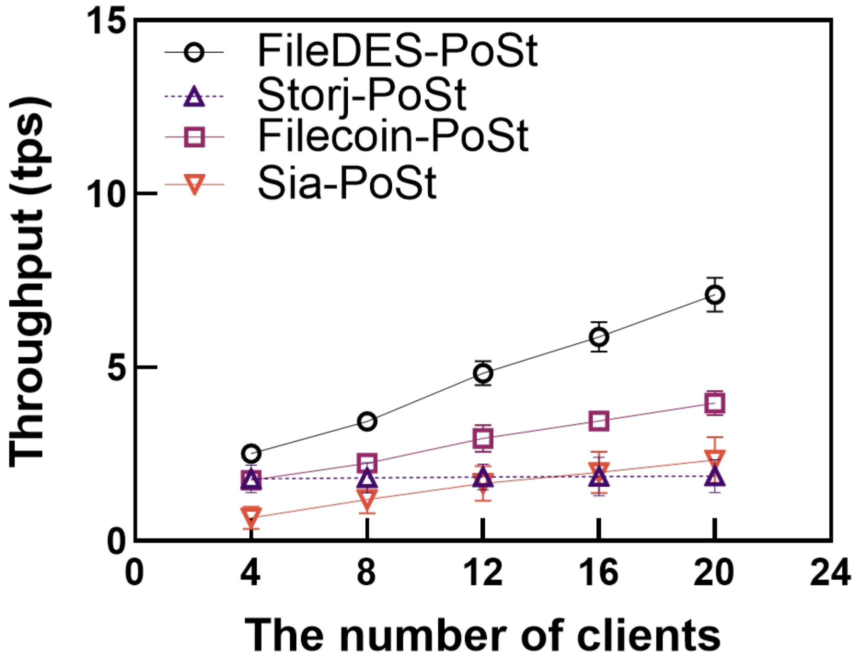
FileDES has the shortest proof generation time and a constant verification time.

FileDES

➢ Throughput and Latency of File Upload and Aggregation (PoS & PoSt)



(a) PoS Throughput



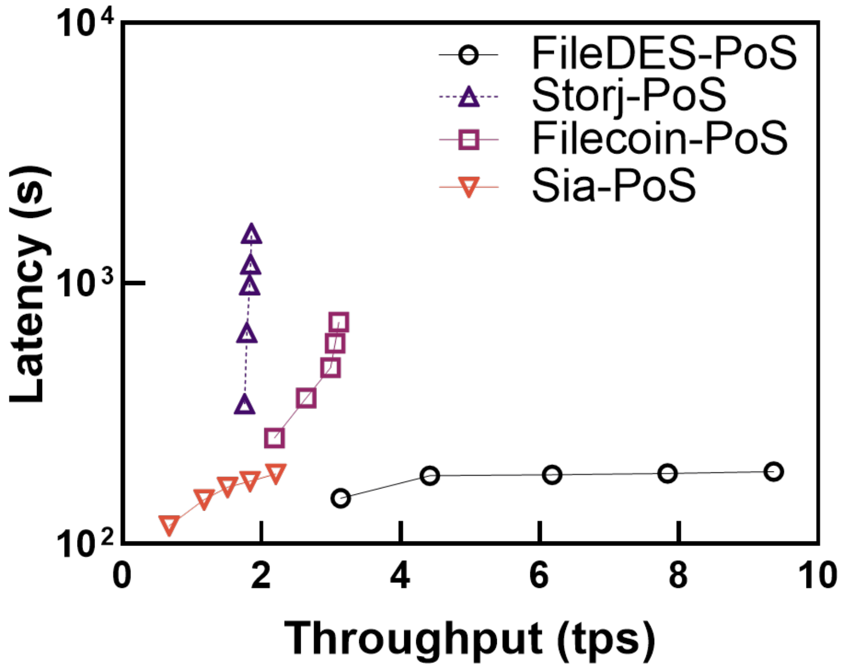
(b) PoSt Throughput

Note:
Tested with 120 ECS instances, 100 miners and 4-20 clients.

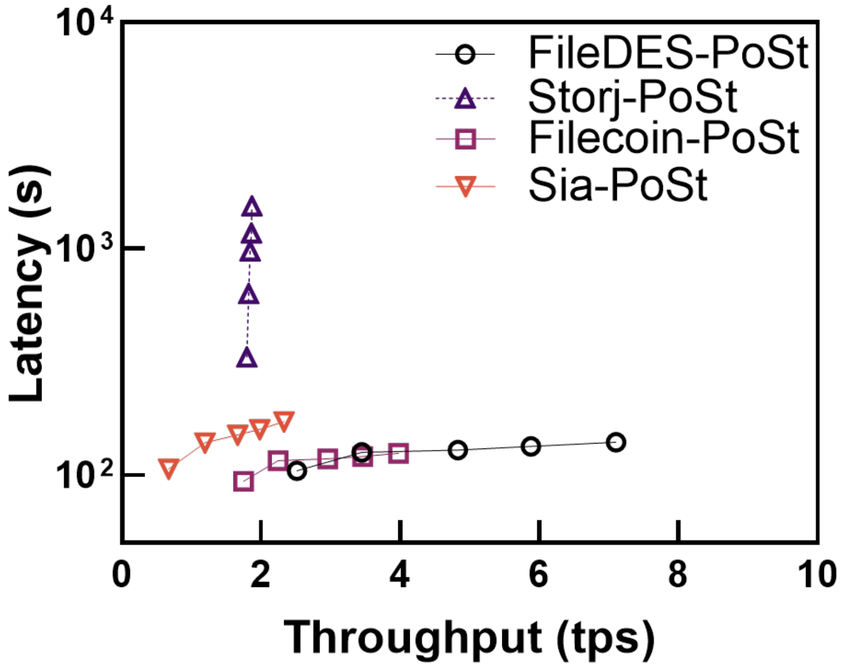
FileDES has the highest throughput in PoS and PoSt generation.

FileDES

➢ Throughput and Latency of File Upload and Aggregation (PoS & PoSt)



(a) PoS Latency



(b) PoSt Latency

Note:
Tested with 120 ECS instances, 100 miners and 4-20 clients.

The latency of PoS and PoSt in FileDES are stable, which shows FileDES has good system scalability.

FileDES

➢ Conclusion and Future Work

- Conclusion:
  - ❑ Encrypted storage to prevent privacy leakage and maintain high data availability.
  - ❑ PoES algorithm to generate PoS and PoSt efficiently.
  - ❑ Batch verification to reduce the computational and verification workload of multiple proofs and improve overall system performance.
- Future work:
  - ❑ Protect user privacy when retrieving files from DSNs.
  - ❑ Fine-grained access control for files in DSNs.

Thank you!
Q & A

Jiahao Zhang

jiahaozhang@mail.sdu.edu.cn

Shandong University