When Blockchain meets Distributed File Systems: An Overview, Challenges and Open Issues [1]

Ajinkya Rajguru

San Jose State University

Department of Computer Science

CS 297 : Prof. Chris Pollett

Agenda

- Introduction

- Introduction to IPFS

- Introduction to SWARM

- Layered Architecture

- Advantages and Disadvantages

- Cutting edge studies

- Conclusion

Introduction

- This paper gives an overview of the layered structure, cutting-edge studies and challenges of IPFS and SWARM – blockchain based distributed filesystems.

IPFS

- Interplanetary File System is a peer-to-peer file system for storing and accessing files, websites, applications and data [3].

SWARM

- SWARM is a decentralized data storage and distribution technology [2].

- It has storage capability that work with blockchain enabling Ethereum based smart contracts [2].

Components of the Layered Architecture

Layers	IPFS	SWARM
Identity Layer	Encrypted hash of Node ID	Keccak hash of public key of Ethereum account
Routing Layer	DSHT (Distributed sloppy Hash table)	Distributed Preimage Archive (DPA)
Network Layer	libP2P	DevP2P
Data Layer	Block – Arbitrary sized data List – collection of blocks Tree – collection of blocks, lists, trees Commit -Snapshot	Chunk – fixed-size piece of data (4 KB max) File – complete set of chunks Manifest – mapping between paths and files

Components of the Layered Architecture

Layers	IPFS	SWARM
Incentive Layer	Filecoin – blockchain based digital payment system	Ethereum token – Makes use of SWARM Account Protocol (SWAP)
Consensus Layer	Expected consensus – Transaction Proof, Proof of Replication, Proof of Spacetime	Proof of work and Proof of stages
Data-swap Layer	BitSwap – based on bitTorrent protocol	Nodes store chunks for selling to get profits on data retrieve request

Issues and Challenges

- Scalability performance –

  - Limited bandwidth of each instance of IPFS could cause low scalability.

  - Bottlenecks of resolving and downloading while accessing remote nodes.

  - Efficient updating technique (change in hash address as the file changes)

  - Need better performance measurement methodologies and system measurement standards.

Issues and Challenges

- Privacy and Security Issues –

  - Some current versions of DFS like IPFS do not tolerate Byzantine attacks.

  - Every peer can access every file in IPFS.

Applications

- IPFS provides business solutions to enterprise
  - Storing data
  - Opus, a music sharing platform makes uses of both IPFS and Ethereum.
  - IPSE proposed a new search engine built on top of IPFS and blockchain.

- IPFS and SWARM can be combined with big data applications
  - Can be used for secure storage in IoT
  - Uses in Data Analytics

Cutting edge studies of blockchain

- For IPFS average download time increases with increase in cluster size and growth in replication.

- Storage optimization – Can explore Erasure codes and storing data off-chain

- Transparent transactions call for improved privacy – Access control and Peer Anonymity.

Conclusion

- The new generation of blockchain based distributed filesystems have great potentials with key characteristics

  - Novel solution of incentive

  - Low-latency data retrieval

  - Automated auditing

  - Censorship-resistant

Thank You

References

[1] [2020] When Blockchain Meets Distributed File Systems: An Overview, Challenges, and Open Issues. H. Huang, J. Lin, B. Zheng, Z. Zheng and J. Bian. IEEE Access, vol. 8, pp. 50574-50586, 2020, doi: 10.1109/ACCESS.2020.2979881

[2] https://www.ethswarm.org

[3] https://en.wikipedia.org/wiki/InterPlanetary_File_System