MINISTERUL EDUCAȚIEI NAȚIONALE

UNIVERSITATEA TEHNICĂ
DIN CLUJ-NAPOCA

FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE
DEPARTAMENTUL CALCULATOARE

SUMMARY
of the License Thesis entitled:

DATA RELIABILITY USING DECENTRALIZED SYSTEMS

Author: **Ionuț-Dan MATIȘ**
Advisor: **S. L. Dr. Eng. Claudia Daniela POP**

1. Requirements

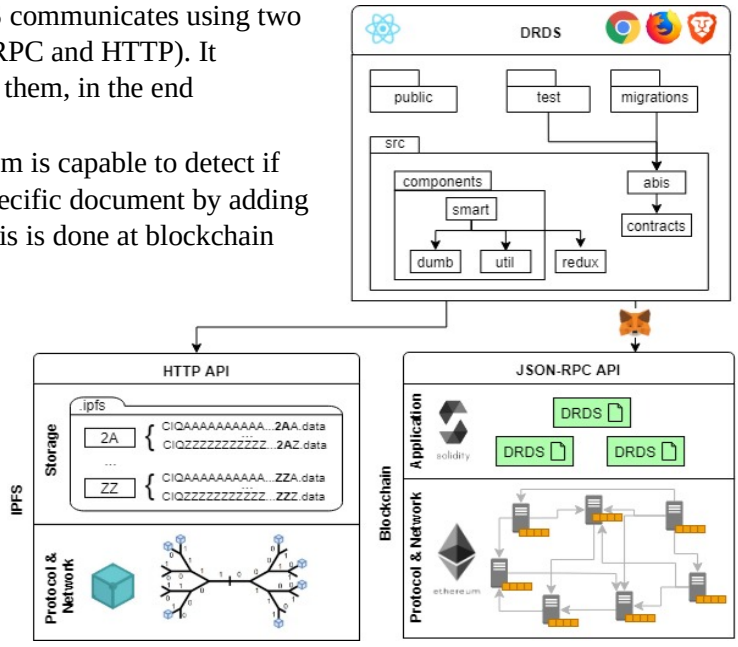
Develop a platform that is able to detect data plagiarism and data corruption while providing its users the ability to version their already uploaded data. These features will be developed using two widely known decentralized systems, Ethereum blockchain and IPFS data storage, as no single central authority has the decisional power. Moreover, the cost of storage should be taken into account.

2. Proposed solution

Storing large amounts of data solely on blockchain is very unpractical. expensive and it defeats its purpose. The solution consists of storing the data on separate specialized storage system (IPFS) and link it in the blockchain (ETHEREUM), in the end combining the best of both worlds.

DRDS is built on top of Ethereum and IPFS communicates using two protocols specified on the diagram (JSON-RPC and HTTP). It leverages their functionalities and enhances them, in the end performing the following:

- Data plagiarism detection: the system is capable to detect if someone is trying to plagiarize a specific document by adding it in the system the second time. This is done at blockchain level, in the DRDS Smart contract.
- Data corruption detection: as data is stored in a remote, trustless server, the system is able to detect alterations of uploaded data and prompts its users of such actions. This is done with repetitive calls to the IPFS nodes that store the files or parts of them



- Data versioning: as data is continuously evolving, the system allows the editing of it. This can be done only by the users that uploaded it in the first place. Moreover, a history of versions is kept and users can navigate between them. The previous discussed features apply to this last one also. The versioning is done through the aid of the smart contract and by emitting events from it, which are then queried are processed in DRDS.

3. Experimental results

The purpose of this system was to find a solution for storing data in a secure, reliable manner which is acceptable in terms of cost. Two secure and reliable ways of storing data were compared: one fully on Ethereum blockchain and this solution, which combines IPFS data storage with Ethereum blockchain. Three usual files have been used for the testing.

Table 6.1: Results for the three files when using only Ethereum

File name	File size	Transactions required	Ether required	Price ($)	Average tx per block
test-file.txt	47 bytes	1	0,000604	0,135	140
fibonacci.txt	949 kilobytes	6	1,614	362,875	15.16
test-image.jpg	8,62 megabytes	52	14,658	3.295,558	1.19

Table 6.2: Results for the three files when combining systems

File name	File size	Transactions required	Ether required	Price ($)	Average tx per block
test-file.txt	47 bytes	1	0,000695	0,156	140
fibonacci.txt	949 kilobytes	1	0,000695	0,156	140
test-image.jpg	8,62 megabytes	1	0,000695	0,156	140

We can clearly see how expensive is to store even a relatively small amount of information (a 8 MB image) in the blockchain. This is due to the fact that the information is replicated to all nodes within the network. The solution that combines the two decentralized systems is far less expensive, having a constant cost, and does not affect the overall blockchain system with very large transactions. This is because the reference to the uploaded file is stored in the blockchain.

4. Personal contributions

The main contributions of this thesis are the following:

- studied Ethereum blockchain and IPFS data storage
- analyzed similar solutions in terms of decentralized data storage
- allowed the two decentralized systems to work together
- added three important features (data plagiarism detection, data corruption detection and data versioning) on top of the two working systems

5. Documentation sources

[1] Nishara Nizamuddin, Haya Hasan, Khaled Salah, *Ipfs-blockchain-based authenticity of online publications*, 2018.
[2] Gavin Wood, Andreas M. Antonopoulos, *Mastering Ethereum: Building Smart Contracts and DApps* O'Reilly Media, 2018, vol. 1.
[3] Juan Benet, *Ipfs - content addressed, versioned, p2p file system*, draft 3

Date: **09.07.2020**

Author: **Matiș Ionuț-Dan**
Coordonator: **S.L. Pop Claudia-Daniela**