Blockchain-based reputation system for peer reviewing

Master's Thesis

Viktor Jacynycz García

Directed by

Samer Hassan Collado and Antonio Sánchez Ruiz-Granados

Máster en ingeniería informática
Masters degree in software engeenering
Facultad de Informática
Universidad Complutense de Madrid

Curso 2017/2018

Document made with TEXIS v.1.1+. modified by Viktor Jacynycz García

This document is prepared for duble-side printing.

Blockchain-based reputation system for peer reviewing

Memoria que presenta para optar al master en ingeniería informática

Viktor Jacynycz García

Directed by

Samer Hassan Collado and Antonio Sánchez Ruiz-Granados

Máster en ingeniería informática
Masters degree in software engeenering
Facultad de Informática
Universidad Complutense de Madrid

Curso 2017/2018

El/la abajo firmante, matriculado/a en el Máster en Ingeniería Informática de la Facultad de Informática, autoriza a la Universidad Complutense de Madrid (UCM) a difundir y utilizar con fines académicos, no comerciales y mencionando expresamente a su autor el presente Trabajo Fin de Máster: "TÍTULO", realizado durante el curso académico 2017-2018 bajo la dirección de Samer Hassan Collado y Antonio Sanchez Ruiz-Granados en el Departamento de ISIA, y a la Biblioteca de la UCM a depositarlo en el Archivo Institucional E-Prints Complutense con el objeto de incrementar la difusión, uso e impacto del trabajo en internet y garantizar su preservación y acceso a largo plazo.

A mi madre, porque sí

A mi hermano, porque también

A jenny, poque tl

A mi padre, porque yokse

Agradecimientos

I find your lack of faith disturbing
Darth Vader, Star Wars: A New Hope.

I thank you.

Resumen

...
...

... Resumen chulo

Índice

Índice de figuras

Índice de Tablas

Capítulo 1

Introduction

Frasemolona

Vik

1.1. Publications systems

Citamos algo para que aparezca en la bibliografía...

Y también ponemos el acrónimo CVS para que no cruja.

Ten en cuenta que si no quieres acrónimos (o no quieres que te falle la compilaciín en "release" mientras no tengas ninguno) basta con que no definas la constante `\acronimosEnRelease` (en `config.tex`).

En el próximo capítulo...

...

Capítulo 2

Technology

The needs of the many outweigh the needs of the few

Spock - The Wrath of Khan

2.1. IPFS

IPFS stands for Interplanetary File System. It is a peer-to-peer file-sharing protocol that uses a cryptographic hashes to store files in a distributed network. IPFS works very similar to HTTP protocol but in a BitTorrent way. It can be seen as a giant git repository where everyone can store, share and exchange files[1].

IPFS merges three main ideas: Distributed Hash Tables, BitTorrent, Git and Self-Certified Systems.

2.1.1. Distributed Hash Tables

A distributed hash table(DHT) is a decentralized structure that works very similar to a hash table. Hash tables are used to identify items in a database. The table performs simple mathematical operations generating a random string called hash. The hash acts as a pointer that directs to the data, this allows the user to find data directly instead of looking through the entire database[2].

In a distributed hash table, any node can use a hash as a key to retrieve data. This system includes a data structure called "keyspace" that is a set of all possible keys, which is split up across the nodes in the system. The mapping of the keys is made by another function that describes the distance from one key to another. All the nodes have and identifier and a set of identifiers pointing to all its neighbours nodes. If a node is removed from the network, only a small portion of the data must be recovered by other nodes[2].

This system makes *DHTs* scalable, fast and robust. It is used by frameworks such as Tapestry [3], Chord [4], Kelips [5], Kademlia [6] and IPFS [1]. These platforms are similar in cost and performance if they are tested in a large enough network. They behave very

fast when it comes to searching for a key through massive networks of nodes[7], that's why it is used by IPFS to create its distributed file system.

2.1.2. BitTorrent - File sharing

BitTorrent [8] is a P2P file sharing system used worldwide. In this system, files are divides into very small chucks of data, and are shared in a peer-to-peer network. Each peer aims to maximize its download rate by connecting to low latency peers. In BitTorrent's network, peers with high upload rate will get higher download rate, so the key is balancing the network bandwidth between downloading and uploading files[9].

IPFS uses three main features from BitTorrent's protocol[1]:

- BitTorrent's data exchange protocol rewards nodes who contribute to the network, and punishes the ones who don't.
- BitTorrent tracks the availability of file chunks, sending the rarest first rather than sending the most common ones.
- IPFS uses PropShare[10] bandwidth allocation strategy to improve BitTorrent's behaviour facing exploitable scenarios.

2.1.3. Git - Version control system

Git is a distributed version control system (*DVCS*)[11]. Git was born in 2005 when the development process of the Linux kernel lost its version control system. The Linux kernel is one of the biggest free software projects nowadays, it has a great team of developers behind and the code usually changes very frequently. In 2002 the team used BitKeeper as VCS since they had a free license. But in 2005 when this license was over, Linus Torvalds decided to develop his own VCS[12].

Git was designed to be scalable and distributed, and the most important factors that IPFS inherits from Git are: [1]:

- Git implements a Merkle Directed Acyclic Graph [13], an object that reflects changes in a file system in a distributed way.
- Objects are identified by the cryptographic hash of their contents.

Parte I

Apéndices

Apéndice A

Anexo 1: Reuniones del equipo

...

...

Resumen: ...

A.1. Reunión del 07 de Septiembre de 2017

En la primera reunión del equipo se hicieron las presentaciones de los integrantes, y se discutieron las posibles ideas que se podrían implementar como proyecto en la *hackathon*. El tema principal sobre el que se discutía era el impacto social del proyecto, y que las métricas de la *hackathon* así lo exigían.

Realizamos una tormenta de ideas en la que surgieron las siguientes:

- **Plataforma de publicación de artículos académicos distribuida**: Implementar un sistema de publicación de artículos para compartir a través de la comunidad científica utilizando IPFS. Esta plataforma pretende eliminar los costes para el acceso a los artículos que imponen las empresas que se encargan de publicarlos y se benefician por ello. Implementar una plataforma totalmente descentralizada para compartir los artículos de divulgación científica conseguirá que el conocimiento de la investigación académica sea público y accesible por todos.
- **Wikipedia distribuida con modelos de gobernanza**: La idea de este proyecto inicialmente era descentralizar la plataforma de Wikipedia a través de IFPS y añadir algún modelo de gobernanza y de reputación para las revisiones de los artículos. EL problema es que es un proyecto muy complejo para implementarlo en sólo un mes, y haría falta un equipo bastante grande y la colaboración de la propia Wikipedia para llevarlo a cabo.

- **Aplicación de contactos para homosexuales en países donde son colectivos reprimidos**: En países como Rusia, los colectivos LGBT son reprimidos hasta el punto de que expresar su sexualidad puede ser un peligro para su seguridad personal. Esta idea trataba de poner en contacto de la manera más anónima y discreta posible a esas personas sin exponerse a los riesgos que ello conlleva.
- **ONG distribuida**: Esta plataforma pretendía ofrecer una bolsa de dinero en la que las personas iban realizando donaciones. Cada semana los donantes votaban dónde se iban a invertir el dinero mediante un sistema de votos.
- **Plataforma de intercambio de conocimientos de programación distribuida**: Stack Exchange es una de las web más importantes en la comunidad informática. Esta solución propone una altenativa totalmente distribuida mediante blockchain.
- **Plataforma de toma de decisiones distribuida**: La toma de decisiones en comunidades reprimidas es bastante dificil. Mediante una aplicación de toma de decisiones en blockchain (como la que tiene Loomio), se pueden ofrecer una herramienta para que estas personas en riesgo de exclusión se hagan oir.
- **Plataforma de crowdfunding para *wistleblowers***: El problema de las plataformas de crowdfunding es que una vez que se financia el proyecto, el usuario sólo puede ver el final del producto esperando que lo que ha financiado sea como promenten los desarrolladores. Esta plataforma propondría una alternativa con varios entregables en función del dinero que se vaya consiguiendo.

...

A.2. Reunión del 08 de Septiembre de 2017

Una vez que el equipo ha decidido el proyecto que vamos a afrontar, nos reunimos para ir decidiendo poco a poco las funcionalidades que habría de tener nuestra plataforma. Algunas de ellas son:

Bibliografía

Y así, del mucho sleer y del poco dormir, se le secó el celebro de manera que vino a perder el juicio.

Miguel de Cervantes Saavedra

[1] J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561*, 2014.

[2] A. Kaluszka, "Distributed hash tables," 2010.

[3] B. Y. Zhao, L. Huang, J. Stribling, S. C. Rhea, A. D. Joseph, and J. D. Kubiatowicz, "Tapestry: A resilient global-scale overlay for service deployment," *IEEE Journal on selected areas in communications*, vol. 22, no. 1, pp. 41–53, 2004.

[4] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 4, pp. 149–160, 2001.

[5] I. Gupta, K. Birman, P. Linga, A. Demers, and R. Van Renesse, "Kelips: Building an efficient and stable p2p dht through increased memory and background overhead," *Peer-to-Peer Systems II*, pp. 160–169, 2003.

[6] P. Maymounkov and D. Mazieres, "Kademlia: A peer-to-peer information system based on the xor metric," in *International Workshop on Peer-to-Peer Systems*. Springer, 2002, pp. 53–65.

[7] J. Li, J. Stribling, T. M. Gil, R. Morris, and M. F. Kaashoek, "Comparing the performance of distributed hash tables under churn." in *Iptps*, vol. 4. Springer, 2004, pp. 87–99.

[8] B. Cohen, "Incentives build robustness in bittorrent," in *Workshop on Economics of Peer-to-Peer systems*, vol. 6, 2003, pp. 68–72.

[9] J. Pouwelse, P. Garbacki, D. Epema, and H. Sips, "The bittorrent p2p file-sharing system: Measurements and analysis," in *IPTPS*, vol. 5. Springer, 2005, pp. 205–216.

[10] D. Levin, K. LaCurts, N. Spring, and B. Bhattacharjee, "Bittorrent is an auction: analyzing and improving bittorrent's incentives," in *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4. ACM, 2008, pp. 243–254.

[11] L. Torvalds and J. Hamano, "Git: Fast version control system," *URL http://git-scm.com*, 2010.

[12] D. Spinellis, "Version control systems," *IEEE Software*, vol. 22, no. 5, pp. 108–109, 2005.

[13] D. Bleichenbacher and U. M. Maurer, "Directed acyclic graphs, one-way functions and digital signatures," in *Annual International Cryptology Conference*. Springer, 1994, pp. 75–82.

–¿Qué te parece desto, Sancho? – Dijo Don Quijote –
Bien podrán los encantadores quitarme la ventura,
pero el esfuerzo y el ánimo, será imposible.

Segunda parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes

–Buena está – dijo Sancho –; fírmela vuestra merced.
–No es menester firmarla – dijo Don Quijote–,
sino solamente poner mi rúbrica.

Primera parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes