DataChain: Uma Ferramenta para Assegurar a Propriedade e Imutabilidade de Documentos Digitais

Gabriel O. Mendanha [1], **Lívia A. Cruz** [1], **Regis P. Magalhães**[1]

[1]Universidade Federal do Ceará – Campus de Quixadá
Quixadá – CE – Brasil

gabrielmendanha@alu.ufc.br, {livia.almada,regismagalhaes}@ufc.br

Abstract. *After the creation of the Bitcoin cryptocurrency with the study and research of the technology that made it possible – the blockchain – some initiatives take the opportunity to use it in other use cases not related to financial transactions or digital money. Motivated by a reality that people easily adulterate, copy and corrupt digital information, this paper presents a proof of concept using blockchain, taking advantage of the characteristics of immutability and ownership of the data. This article focuses on digital documents and opens up a range of use cases and possibilities for a variety of consumers who want a way to prove the authenticity and ownership of documents as well as transfer them to others.*

Resumo. *Após a criação da criptomoeda Bitcoin, com o estudo e pesquisa da tecnologia que a tornou possível, a blockchain, viu-se a possibilidade de utiliza-lá para outros casos de uso não relacionados a transações financeiras ou dinheiro digital. Motivado por uma realidade que pessoas facilmente adulteram, copiam e corrompem informações digitais, este trabalho apresenta uma prova de conceito utilizando a blockchain, tirando proveito das características de imutabilidade e propriedade dos dados. Voltado para documentos digitais, o presente trabalho também abre uma gama de casos de uso e possibilidades para uma variedade de consumidores que desejam uma maneira de provar a autenticidade e posse de documentos, assim como transferí-los a outras pessoas.*

1. Introdução

Em 2009 foi publicada por Satoshi Nakamoto[1] a primeira criptomoeda digital bem sucedida, o Bitcoin [Nakamoto 2008]. Uma combinação de fatores permitiu essa moeda ser difundida. Além de ser um ativo digital, ela também é um sistema descentralizado que permite transações financeiras seguras em uma rede *peer-to-peer* com participantes não-confiáveis, sem depender de um instituição financeira. A descentralização é uma dentre muitas características que foram incorporadas, graças a tecnologia *blockchain*, que é fundamentalmente um banco de dados transparente e descentralizado que contém o registro de todas as transações.

Este trabalho propõe uma ferramenta[2] que, utilizando a tecnologia de *blockchain* pública, permite às pessoas uma maneira de provar a posse e autenticidade de um documento digital, assim como transferi-lo a outra pessoa. O sistema também oferece um

[1]Satoshi Nakamoto é um pseudônimo. Até a conclusão deste trabalho não se sabe a identidade da pessoa ou organização responsável pela criação do Bitcoin.

[2]https://vimeo.com/225034651

meio para armazenar o documento com a garantia de que o mesmo não poderá ser modificado ou removido pelo dono, por outra pessoa, ou pelo administrador do sistema. A plataforma proposta tem em seu núcleo o BigchainDB, um banco de dados distribuído e descentralizado que incorporou as características da *blockchain* sem perda de escalabilidade [McConaghy et al. 2016].

A principal contribuição deste trabalho é a proposta de uma arquitetura descentralizada para armazenamento de documentos que integra a *blockchain* com um sistema de arquivos distribuído para garantir a posse, a integridade e a imutabilidade de documentos digitais. Vale ressaltar que a *blockchain* é uma tecnologia emergente e que seu uso fora do contexto de criptomoedas, como por exemplo, na prova de existência de documentos legais traz novos desafios. [Crosby et al. 2016].

2. Estrutura e características da *Blockchain*

A *blockchain* é uma lista encadeada e ordenada de blocos que contém transações como conteúdo principal. Cada bloco referencia o seu anterior e os mesmos são identificados pelo resultado de uma função *hash* criptográfica, que mapeia um dado de tamanho variável para um dado de tamanho fixo. Na *blockchain* é desejável que a função *hash* produza dois resultados iguais se, e somente se, as entradas forem as mesmas. Portanto, deve existir um valor *hash* diferente para cada bloco [Antonopoulos 2014].

A Figura 1 ilustra a estrutura da *blockchain*. Observe que os blocos são representados por retângulos. H representa uma função *hash* que recebe como parâmetro o conteúdo do bloco anterior, indicado pela seta. Logo, cada bloco sabe quem é o seu anterior através do valor *hash* obtido como resultado obtido da chamada da função H.

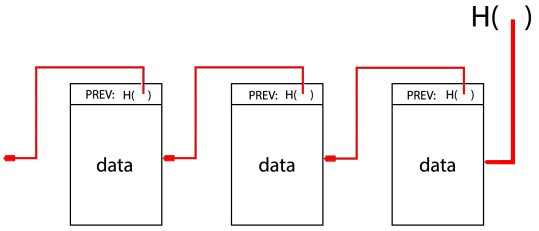

Figura 1. Estrutura da *blockchain*

2.1. Distribuição, integridade e segurança na *Blockchain*

Considerando o contexto da tecnologia da informação, um livro-razão é um mecanismo de armazenamento que permite apenas inserções [Peter Evans-Greenwood 2016]. Em um livro-razão, as informações são imutáveis e podem conter dados genéricos, portanto a *blockchain* pode ser vista como uma tecnologia de livro-razão. Na *blockchain*, não existe um agente central responsável pelo gerenciamento do sistema, ou seja, o controle é descentralizado. Este controle é feito através um conjunto de nós em uma rede *peer-to-peer* [McConaghy et al. 2016]. Assim, a responsabilidade de decidir o que incluir, em que ordem incluir e de garantir que um registro não seja alterado após sua inclusão é distribuída. Um grupo de nós, através de um algoritmo de consenso, divide essa responsabilidade.

Assim como em qualquer outro sistema distribuído, a *blockchain* possui um problema de resolução de conflitos. Se dois fatos incompatíveis chegarem no mesmo instante,

o sistema deve possuir regras que determinem qual dos fatos será considerado válido. A Figura 2 exemplifica o problema de resolução de conflitos. Nesta Figura, Alice envia $10 para Bob e os mesmos $10 para Charlie. O problema está no fato de que Alice possui somente $10 e está tentando gastar duas vezes este valor. Uma maneira de resolver este problema é ordenando os fatos, o primeiro que for registrado é o vencedor.

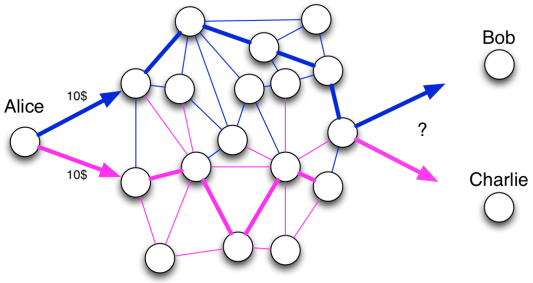

Figura 2. O problema do gasto duplo [Zaninotto 2016]

Porém, ambos os fatos podem aparecer em ordens diferentes em nós distantes um do outro. Para que toda a rede concorde na ordem dos fatos e preserve sua integridade é necessário um sistema de sincronização de dados, um algoritmo de consenso [Zaninotto 2016].

No quesito segurança na *blockchain*, a criptografia de chave pública/privada, que é um dos fundamentos da segurança moderna é usada para possibilitar que as pessoas assinem digitalmente documentos como: arquivos de texto, imagens, etc. A *blockchain* utiliza-se deste sistema de criptografia para assinar digitalmente as transações. O acesso e a utilização dos ativos digitais não é possível sem o conhecimento da chave privada do dono atual. [Peters and Panayi 2016]

3. O Sistema DataChain

O sistema proposto, DataChain[3], armazena os documentos e estabelece um vínculo de posse, que pode ser transferível a outra pessoa através de um par de chaves criptográficas, além de também permitir a validação do título de posse de documentos entre os indivíduos que utilizam a plataforma. O público-alvo são pessoas que querem uma forma de garantir o reconhecimento como autor de uma obra intelectual, como: artistas, compositores, pesquisadores, etc. Assim como pessoas de negócio que podem transferir, por exemplo, a escritura de uma casa para outra pessoa de forma ágil e segura, ou até mesmo por entidades que desejam de alguma forma combater fraude de documentos sensíveis, como: documentos de identidade pessoal, prontuários médicos, diplomas, etc. O sistema proposto, além de utilizar tecnologias inovadoras, também serve como base ou inspiração para o construção de aplicações mais robustas e especializadas.

3.1. Arquitetura

O sistema possui como principais componentes de sua arquitetura o banco de dados BigchainDB [McConaghy et al. 2016] e um sistema de arquivos distribuído denominado *InterPlanetary File System* (IPFS) [Benet 2014].

[3]Código-fonte disponível em: https://github.com/gabrielmendanha/tcc2

O BigchainDB é um banco de dados *open-source*, descentralizado e distribuído e incorpora as melhores características da *blockchain* e de bancos de dados distribuídos (BDD) [McConaghy et al. 2016]. No BigchainDB, a posse de determinado ativo digital é garantida através do par de chaves pública e privada. Para todas as transações é necessário gerar uma assinatura digital que é calculada com base na chave privada e na chave pública da pessoa. Logo, para uma pessoa expressar o seu consentimento em querer transferir um ativo digital para outra pessoa, ela deve informar ao sistema a sua chave privada.

O IPFS é um sistema de arquivos que implementa o modelo *peer-to-peer*. Ele tem como objetivo conectar vários dispositivos ao mesmo sistema de arquivos, provendo um modelo de armazenamento endereçado ao conteúdo, ao mesmo tempo em que os nós não precisam confiar uns nos outros e todos possuem a mesma influência na rede. Os nós se conectam entre si para transferir os arquivos.

O sistema de arquivos gera um valor *hash* único e imutável para cada documento digital, além de possibilitar encontrar o arquivo na rede *peer-to-peer* a partir deste mesmo valor *hash*. Este sistema replica o arquivo nos nós que o requisitam, diminuindo assim as chances de determinado arquivo ficar indisponível temporariamente ou permanentemente em caso de catástrofe (falha no disco rígido, quedas de energia ou interrupção do serviço de internet do servidor que provê os documentos, por exemplo).

A utilização do IPFS com o BigchainDB torna desnecessário o armazenamento dos documentos diretamente no banco de dados, o que implica ganhos de desempenho uma vez que o sistema não está mais limitado ao tipo e tamanho dos dados aceitos pelo BigchainDB. Portando, não é necessário que sejam processadas conversões toda vez que um documento for consultado ou inserido.

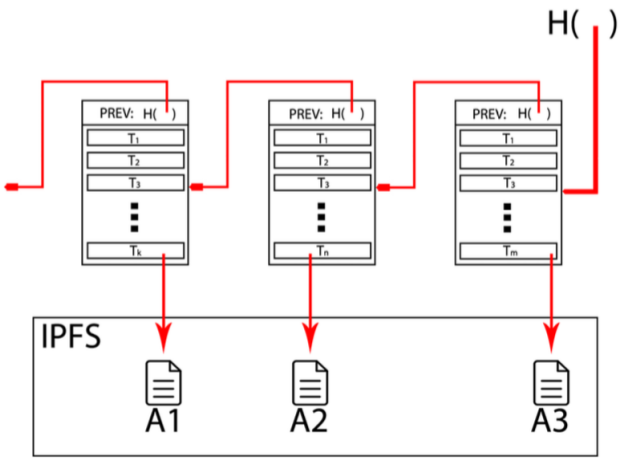

Figura 3. Integração entre a *blockchain* e o IPFS

A Figura 3 ilustra como a *blockchain* provida pelo BigchainDB integra-se aos objetos no IPFS. Como o valor *hash* dos objetos no IPFS é único, imutável e basta para localizar determinado arquivo no sistema, é viável e seguro armazená-lo como uma referência para o documento binário. Embora omitido na imagem, cada transação dentro de um bloco armazena o valor *hash* do nó inicial. Quaisquer modificações em documentos no IPFS geram um novo objeto que não está incluso na *blockchain*. Entretanto, o

objeto original permanece intacto e seu valor *hash* continua vinculado à *blockchain*. O novo objeto, com as modificações não é reconhecido pelo sistema, garantindo assim a imutabilidade do documento original.

4. Visão Geral da Demonstração

A submissão, recuperação e transferência de um documento de um usuário são realizadas através dos passos definidos a seguir.

1. Realização do download do par de chaves criptográficas, se necessário.
2. Escolha do documento que deseja submeter à *blockchain*.
3. Realização do download do comprovante
4. Com os dados fornecidos no comprovante, o usuário pode consultar determinado documento, além de fornecer uma assinatura pública (opcional) para verificar se o documento pertence a assinatura fornecida.
5. Em posse da assinatura privada e da referência do documento, ela pode transferir a posse a quem desejar, representado pela assinatura pública.

(a) Upload de documento

(b) Detalhes da transação submetida

(c) Consulta de documento

(d) Transferência de posse

5. Considerações Finais

Neste trabalho, foi apresentado o DataChain, um sistema baseado na *blockchain* que integrada ao sistema de arquivos IPFS possibilita prover funcionalidades interessantes ao usufruir das características de imutabilidade e posse da *blockchain* de maneira escalável. Voltado para pessoas e entidades que de alguma forma desejam uma maneira de provar o título de posse de determinado documento digital ou combater fraudes de documento sensíveis a adulteração.

Referências

Antonopoulos, A. M. (2014). *Mastering Bitcoin: unlocking digital cryptocurrencies*. "O'Reilly Media, Inc.".

Benet, J. (2014). Ipfs-content addressed, versioned, p2p file system. *arXiv preprint arXiv:1407.3561*.

Crosby, M., Pattanayak, P., Verma, S., and Kalyanaraman, V. (2016). Blockchain technology: Beyond bitcoin. *Applied Innovation*, 2:6–10.

McConaghy, T., Marques, R., Müller, A., De Jonghe, D., McConaghy, T., McMullen, G., Henderson, R., Bellemare, S., and Granzotto, A. (2016). Bigchaindb: a scalable blockchain database. *white paper, BigChainDB*.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

Peter Evans-Greenwood, Robert Hillard, I. H. P. W. (2016). Bitcoin, blockchain and distributed ledgers: Caught between promise and reality.

Peters, G. W. and Panayi, E. (2016). Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money. In *Banking Beyond Banks and Money*, pages 239–278. Springer.

Zaninotto, F. (2016). The blockchain explained to web developers. `http://marmelab.com/blog/2016/04/28/blockchain-for-web-developers-the-theory.html`. Acesso em: 21 jun. 2016.