

La supuesta inmutabilidad del blockchain...

...y el derecho al olvido





Contenidos

1. Blockchain y GDPR
2. Equívocos
3. Amenazas y riesgos para la protección de datos
4. Retos y PoC



1. Blockchain y GDPR

Red de participantes (personas físicas o jurídicas) que comparten un conjunto de datos de forma distribuida (cada uno mantiene su propia copia), en el que se anota quién posee qué (activos en forma de datos), donde se negocia con quién se intercambian dichos activos (transacciones) y con medidas para gestionar la consistencia e integridad de los datos



Blockchain es un sistema de registro distribuido (distributed ledger), peer-to-peer, criptográficamente protegido y actualizable mediante consenso entre pares

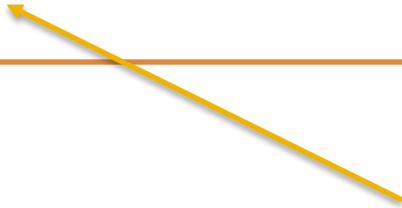


1. Blockchain y GDPR

La dirección es el identificador único obtenido a partir de la clave pública del participante que permite realizar actividad y efectuar transacciones.

La transacción es una operación que queda registrada en la Blockchain y modifica la información en ella, transfiriendo datos o valor entre el emisor y el receptor.

El bloque es la unidad fundamental de información que contiene un conjunto de transacciones validadas y otros datos relevantes, vinculado mediante hash con los bloques anteriores para gestionar la integridad.



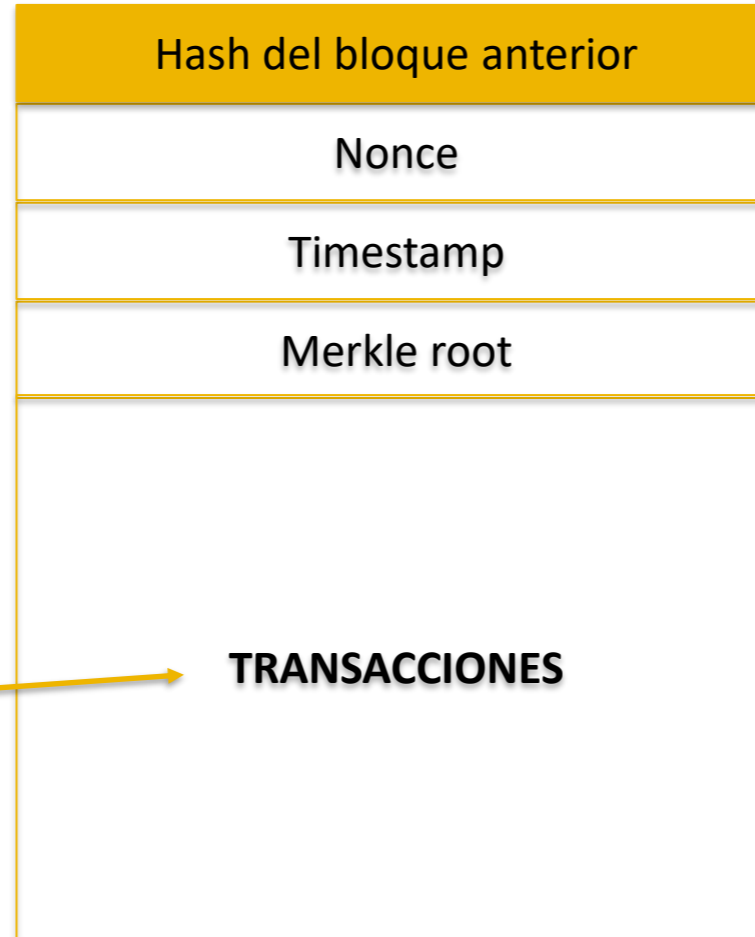
La transacción va firmada digitalmente por el emisor y ha de ser validada por los nodos para su inclusión en un bloque de la Blockchain



1. Blockchain y GDPR

BLOQUE

Para cada transacción: Lógica de transferencia de datos o valor, reglas relevantes, direcciones de origen y destino, y otra información de validación

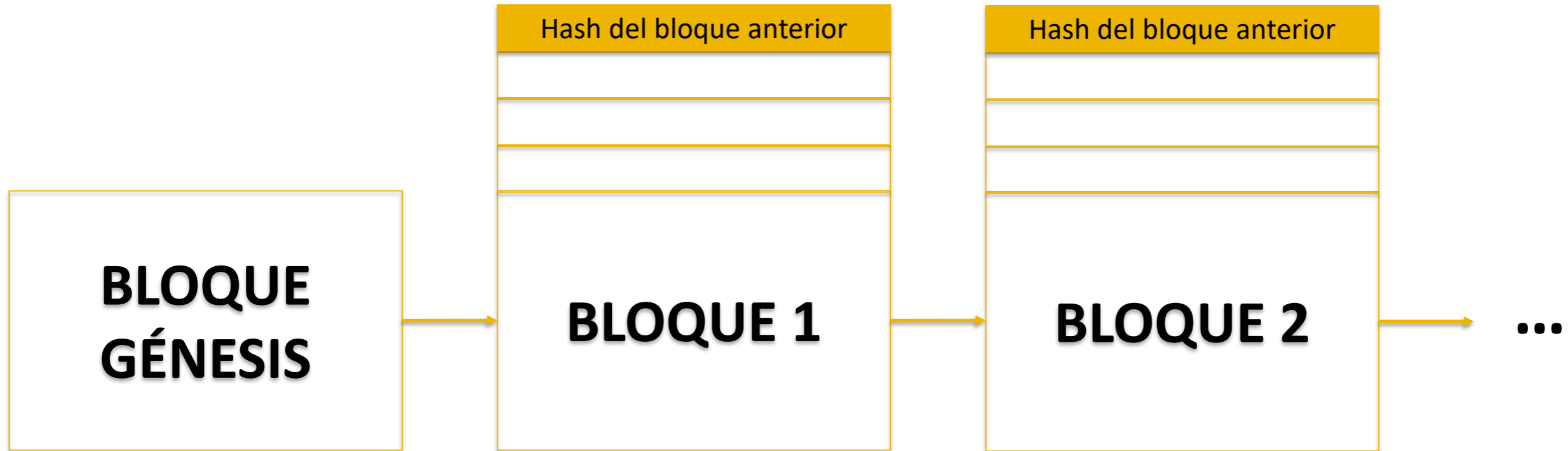


Cabecera

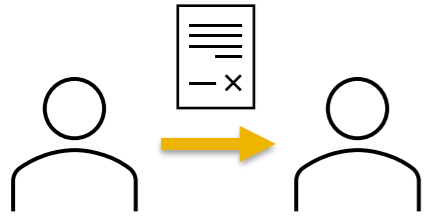
Cuerpo



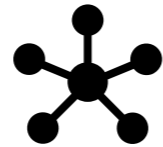
1. Blockchain y GDPR



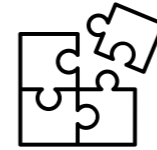
1. Blockchain y GDPR



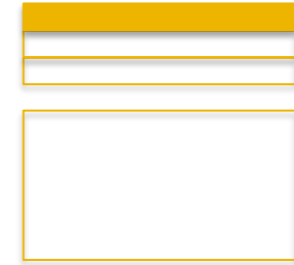
Se inicia la transacción



Se valida y se difunde



Se busca un bloque nuevo (minería)



Se encuentra un bloque nuevo

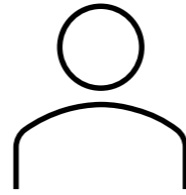


Se agrega un nuevo bloque a la cadena



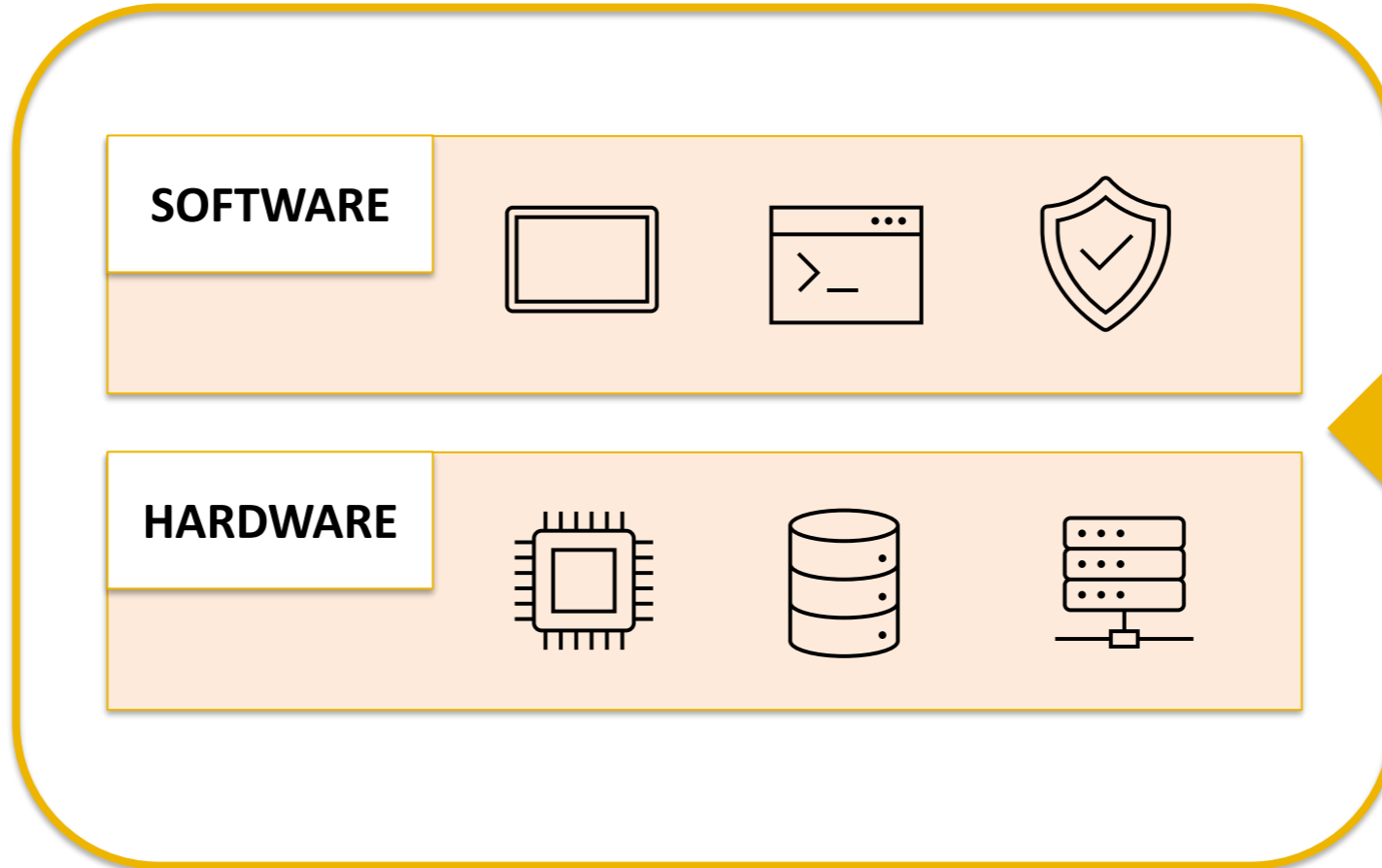
PARTICIPANTE

Persona física o jurídica



NODO

- Nodo ligero
- Nodo completo
- Nodo archivo
- Nodo minero
- Nodo validador



1. Blockchain y GDPR



Aplicaciones (Smart Contracts, apps descentralizadas, agentes autónomos)

Ejecución (bloques, transacciones, minado, incentivos)

Consenso (ejecución, tolerancia a fallos)

Criptografía (funciones hash, clave privada/pública, firmas digitales)

P2P (enrutado, difusión)

Red (pila de protocolos)



¿Dónde se pueden encontrar almacenados datos personales?

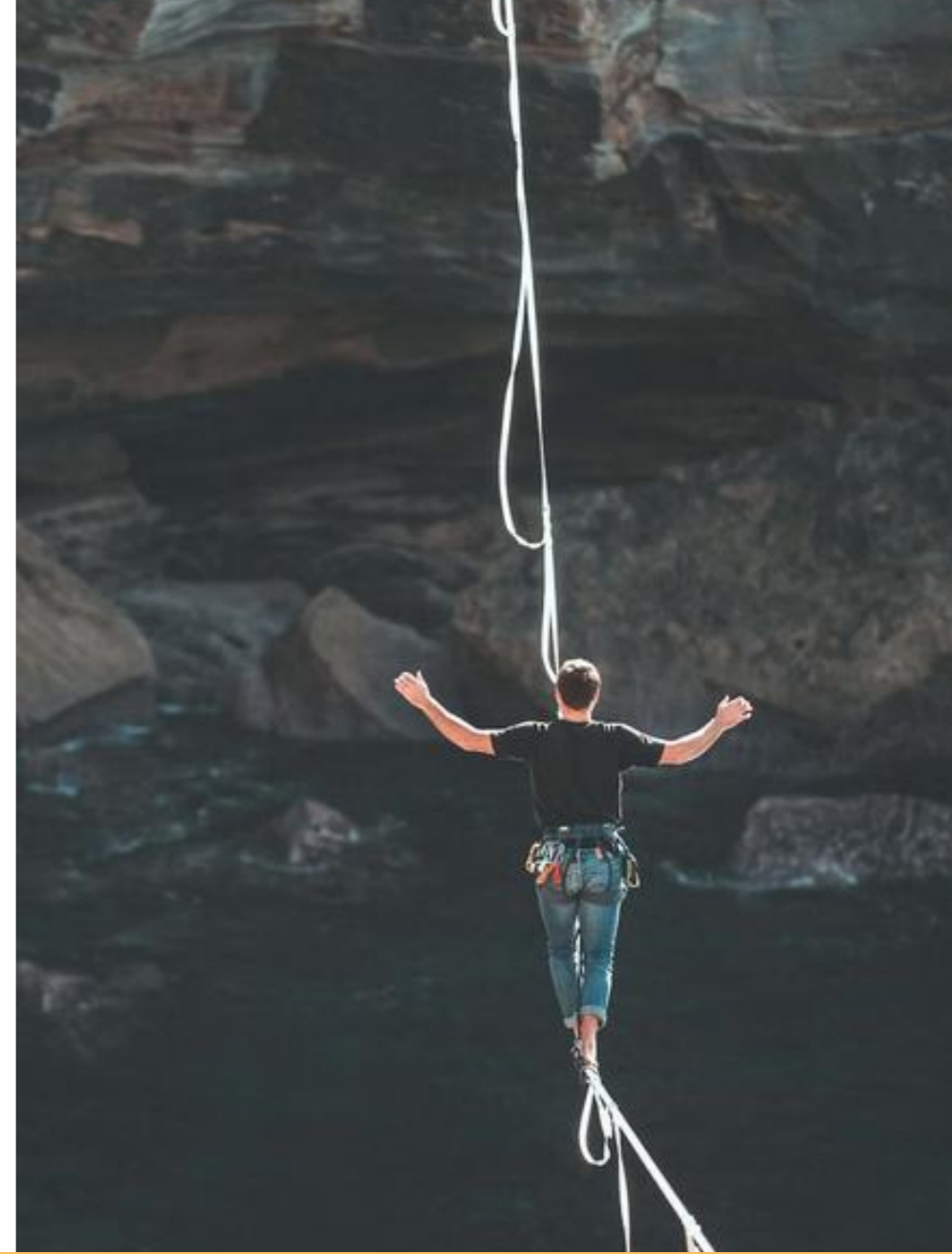
Transacciones (dirección origen, dirección destino y datos de la transacción)

Balances de cuentas

Recibos de transacciones

Almacenamiento de los Smart Contracts

Almacenamiento off-chain



Artículo 5

Principios relativos al tratamiento

Artículo 25

Protección de datos desde el diseño y por defecto

REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO

de 27 de abril de 2016

relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos)

GDPR

CAPÍTULO III

Derechos del interesado

Artículo 32

Seguridad del tratamiento



2. Equívocos

“Está implementado con blockchain, es inmutable”

Que sea complicado no significa que los datos no puedan modificarse o borrarse: Ethereum DAO Fork (2016), ataques 51% Ethereum Classic (2019-2020), Axie Infinity (2022), Bitcoin Gold (2018-2020), etc.

“Es completamente descentralizado”

Pero incluso en infraestructuras públicas y no permissionadas hay una gran concentración de poder en unos pocos intervinientes...



Aponte-Novoa, F. A., Orozco, A. L. S., Villanueva-Polanco, R., & Wightman, P. (2021). The 51% attack on blockchains: A mining behavior study. *IEEE access*, 9, 140549-140564.



2. Equívocos

“Gobernanza informal y automatizada”

Gobernanza “de facto” ejercida por la comunidad. Incompleta o improvisada, con desequilibrios entre participantes, pero existe, los nodos no operan de manera autónoma.

“El código es ley” (y se asume que es inteligente)

Los Smart Contracts son programas desarrollados por personas, para cumplir con sus objetivos, su ejecución la lanzan personas.

MEV: <https://ethereum.org/es/developers/docs/mev/>

➤ Daian, P., et al. (2020, May). Flash boys 2.0: Frontrunning in decentralized exchanges, miner extractable value, and consensus instability. In 2020 IEEE symposium on security and privacy (SP) (pp. 910-927).

<https://ethereumclassic.org/es/why-classic/code-is-law>



2. Equívocos

“Los usuarios tienen el control”

¿Sobre qué exactamente? ¿Cómo y ante qué entidad lo pueden ejercer?

“Sólo se almacenan datos en las transacciones y en los bloques”

Los nodos, además de las cadenas/tablas de bloques, necesitan almacenar una gran variedad de datos adicionales.



3. Amenazas y riesgos para la protección de datos

Tengo los malentendidos superados, sé lo que me hago.

Además, tengo claro que la mejor manera de diseñar/implementar mi tratamiento de datos personales es con una infraestructura Blockchain.

¡Manos a la obra!: Modelado de amenazas y gestión del riesgo





Vinculación y detección

A través de identificadores únicos

Por ejemplo, si sólo se utiliza un par de claves por participante, se puede saber qué transacciones pertenecen al mismo.

Mediante combinación o inferencia

Por ejemplo, de valores asociados a las transacciones, metadatos, huella de red, etc.



Identificación

2019 IEEE European Symposium on Security and Privacy (EuroS&P)

Deanonimization and linkability of cryptocurrency transactions based on network analysis

Alex Biryukov
University of Luxembourg
alex.biryukov@uni.lu

Sergei Tikhomirov
University of Luxembourg
sergey.s.tikhomirov@gmail.com

Abstract—Bitcoin, introduced in 2008 and launched in 2009, is the first digital currency to solve the double spending problem without relying on a trusted third party. Bitcoin provides a way to transact without any trusted intermediary, but its privacy guarantees are questionable. Despite the fact that Bitcoin addresses are not linked to any identity, multiple deanonimization attacks have been proposed. Alternative cryptocurrencies such as Dash, Monero, and Zcash aim to provide stronger privacy by using sophisticated cryptographic techniques to obfuscate transaction data.

Previous work in cryptocurrency privacy mostly focused on applying data mining algorithms to the transaction graph extracted from the blockchain. We focus on a less well researched vector for privacy attacks: network analysis. We argue that timing of transaction messages leak information about their origin, which can be exploited by a well connected adversarial node. For the first time, network level attacks on Bitcoin and the three major privacy-focused cryptocurrencies have been examined. We describe the message propagation mechanics and privacy guarantees in Bitcoin, Dash, Monero, and Zcash. We propose a novel technique for linking transactions based on transaction propagation analysis. We also unpack address advertisement messages (ADRs), which under certain assumptions may help in linking transaction clusters to IP addresses of nodes. We implement and evaluate our method, deanonimizing our own transactions in Bitcoin and Zcash with a high level of accuracy. We also show that our technique is applicable to Dash and Monero. We estimate the cost of a full-scale attack on the Bitcoin mainnet at hundreds of US dollars, feasible even for a low budget adversary.

1. INTRODUCTION

Bitcoin was, and still to some extent is, mistakenly referred to as an anonymous currency [42]. Indeed, unlike traditional financial systems, Bitcoin addresses are not tied to any real-world identity at the protocol level, but this fact alone does not guarantee strong privacy. Bitcoin transactions are broadcast through a peer-to-peer network in plaintext, after being verified by miners they are stored in a massively replicated shared database (the blockchain). A common technique to improve privacy in Bitcoin is to use a fresh address for every transaction (generating addresses is only limited by the size of the 256-bit key space). This piece of advice, often implemented in wallets, is no panacea, as the relationships between transactions can be inferred through blockchain analysis.

Multiple cryptographic techniques have been proposed to address the Bitcoin privacy problem, from services on top of the original protocols such as mixers to new alternative cryptocurrencies such as Dash, Monero, and Zcash. Dash

relies on built-in background mixing powered by the so-called masternode network. Monero implements ring signatures and confidential transactions. Zcash uses zero-knowledge proofs, namely, zk-SNARKs (though the majority of transactions do not take advantage of them due to heavy performance cost). Zcash and Dash are based on a fork of the Bitcoin Core codebase, while Monero is not.

Previous attacks on the privacy of cryptocurrency transactions mostly employed some form of data analysis on the transaction graph. We take another approach and analyze propagation times of protocol messages to infer relationships between transactions.

The ultimate goal of deanonimization is to reveal the relationship between cryptocurrency transactions (or addresses) and real-world identifiers, such as IP addresses. In our model, the goal of the adversary in our model is to infer a connection between a cryptocurrency transaction and the IP address of a node which was the first to introduce it into the network.¹ We rely on the core observation that a node can be uniquely identified by its set of connected peers (*entry nodes*). Earlier network-based deanonimization attacks [16] and [30] only took into account the first node to propagate a given transaction to the adversary. Our approach is more sophisticated. We apply carefully chosen weight functions to message timing information. This allows us to link transactions broadcast from one node, even if all addresses involved are unrelated (consequently, blockchain analysis would gain no insight).

Instead of associating transactions with IP addresses directly, we first cluster the transactions, and then try to assign IP addresses to clusters. Even if the latter step gains no insight, the clustering data used in combination with information from other sources is useful for the attacker. Moreover, our technique does not simply produce a binary decision (whether two transactions are related), but also allows for manual visual inspection of transaction clusters using heatmaps.

The rest of the paper is organized as follows. Section II provides an overview of the propagation mechanisms in various cryptocurrencies. Section III describes our approach to transaction clustering based on propagation timing. We implement and evaluate our technique on real-world cryptocurrencies. We were able to cluster our own transactions in Bitcoin and

¹Even though an IP address is not linked to a physical person, it can be used to determine a relatively precise location of the device involved, and can be linked to a real-world identity if the responsible ISP is compromised.

©2019, Alex Biryukov. Under license to IEEE.

DOI: 10.1109/EuroS&P.2019.00022

172

Authorized licensed use limited to: IEEE Xplore. Downloaded on February 19, 2025 at 08:45:05 UTC from IEEE Xplore. Restrictions apply.

IEEE
COMPUTER
SOCIETY

Deanonimizing Ethereum Validators: The P2P Network Has a Privacy Issue

Lioba Heimbach*
ETH Zurich
hlioba@ethz.ch

Yann Vonlanthen*
ETH Zurich
yvonlanthen@ethz.ch

Juan Villacis
University of Bern
juan.villacis@unibe.ch

Lucianna Kiffer
IMDEA Networks
lucianna.kiffer@imdea.org

Roger Wattenhofer
ETH Zurich
wattenhofer@ethz.ch

Abstract

Many blockchain networks aim to preserve the anonymity of validators in the *peer-to-peer* (P2P) network, ensuring that no adversary can link a validator's identifier to the IP address of a peer due to associated privacy and security concerns. This work demonstrates that the Ethereum P2P network does not offer this anonymity. We present a methodology that enables any node in the network to identify validators hosted on connected peers and empirically verify the feasibility of our proposed method. Using data collected from four nodes over three days, we locate more than 15% of Ethereum validators in the P2P network. The insights gained from our deanonimization technique provide valuable information on the distribution of validators across peers, their geographic locations, and hosting organizations. We further discuss the implications and risks associated with the lack of anonymity in the P2P network and propose methods to help validators protect their privacy. The Ethereum Foundation has awarded us a bug bounty, acknowledging the impact of our results.

1 Introduction

Ethereum is a blockchain that emphasizes decentralization, aiming to keep its consensus mechanism accessible to many participants, which contributes significantly to the complexity of its protocol. In particular, Ethereum faces challenges in scaling its consensus protocol while remaining accessible to smaller participants. The large number of validators involved in the consensus process and their extensive message exchanges lead to unprecedented complications. To address this challenge, innovative scaling solutions for the *peer-to-peer* (P2P) network have been proposed and implemented [77].

Our work demonstrates the impact of these scaling solutions on the privacy and security of the Ethereum P2P network and blockchain. We outline how to deanonimize validators in the P2P network by mapping a validator's identifier to the IP address of the machine it is hosted on. Our technique relies

*These authors contributed equally to this work.

solely on observing *attestation* (i.e., consensus layer vote) messages received from peers (i.e., nodes with established TCP connections). By analyzing messages from a peer p , we can infer whether a validator v is hosted on this peer.

Concretely, the main vulnerability stems from the current broadcast implementation, in which nodes are only responsible for propagating a pre-determined subset of all attestations. Thus, when a peer p sends an attestation created by validator v that falls outside their broadcasting responsibility, we can infer that the attestation was produced by p itself. If we observe this behavior repeatedly, we demonstrate that with high confidence the attesting validator v is connected to the peer p .

The Ethereum P2P network's privacy issue poses a major security risk, allowing attackers to identify nodes associated with validators set to create new blocks. This could lead to (D)DoS attacks, halting chain progress, or more targeted attacks on nodes associated with validators handling high value blocks, letting a subsequent malicious validator scoop these profits. We hope our work highlights this lack of privacy and informs future privacy-enhancing solutions.

Contributions. We summarize our main contributions:

- We propose a simple and low-cost technique for a node in the network to deanonimize its peers, i.e., infer which validators they host.
- We perform a measurement study to demonstrate the feasibility of the deanonimization. Using four nodes in just three days, we can locate more than 15% of validators in the P2P network.
- We outline the implication of the lack of anonymity in Ethereum's P2P network (e.g., fairness, liveness and safety concerns) and discuss possible mitigations.
- Finally, we expose novel security risks in the P2P network, highlighting how validators concentrate on certain peers (e.g., we locate over 19,000 validators on a single peer) and how they are spread globally and across organizations (i.e., cloud service and internet service providers). We also discover that operators for different staking pools run multiple pools' validators on the same machine, creating undesirable dependencies.

arXiv:2409.04366v2 [cs.CR] 1 Feb 2025

Biryukov, A., & Tikhomirov, S. (2019, June). Deanonimization and linkability of cryptocurrency transactions based on network analysis. In *2019 IEEE European symposium on security and privacy (EuroS&P)* (pp. 172-184).

Heimbach, L., Vonlanthen, Y., Villacis, J., Kiffer, L., & Wattenhofer, R. (2024). Deanonimizing ethereum validators: The P2P network has a privacy issue. *arXiv preprint arXiv:2409.04366*.

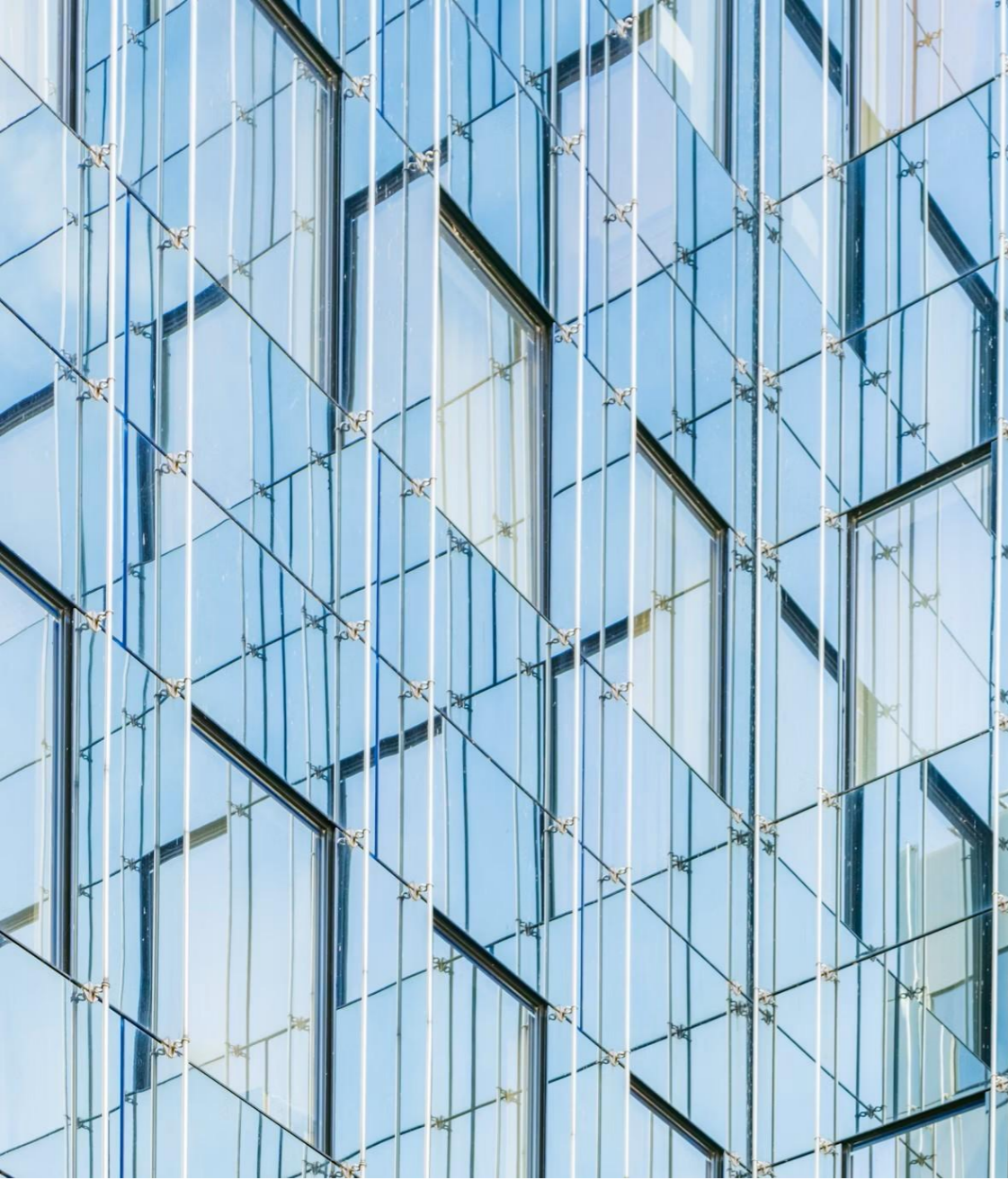


Divulgación

Falta de confidencialidad de las transacciones

Las transacciones revelan información sobre los participantes y su actividad.

La mayor parte de las implementaciones de los mecanismos de validación y consenso exigen acceso a todos los datos de la transacción, al código del Smart Contract, etc.



Inexactitud y no repudio

Si se gobierna la blockchain como si fuera inmutable

Las transacciones quedan registradas y un participante no podrá rectificar sus datos personales. O negar haber participado en una transacción.



Falta de transparencia y de capacidad para intervenir

¡EQUÍVOCOS! -> INCUMPLIMIENTOS

Se consideran características intrínsecas de la tecnología aspectos que tienen que ver en realidad con su diseño o su gobierno y esto hace que no se cumplan muchos de los principios y requisitos establecidos por el RGPD: transparencia, limitación de la finalidad, minimización de datos, limitación del plazo de conservación, derecho de supresión, derecho a la limitación del tratamiento, etc.



4. Retos y PoC

No puede invocarse la imposibilidad técnica para justificar el incumplimiento de los principios y requisitos del RGPD

Artículo 25

Protección de datos desde el diseño y por defecto

1. Teniendo en cuenta el estado de la técnica, el coste de la aplicación y la naturaleza, ámbito, contexto y fines del tratamiento, así como los riesgos de diversa probabilidad y gravedad que entraña el tratamiento para los derechos y libertades de las personas físicas, el responsable del tratamiento aplicará, tanto en el momento de determinar los medios de tratamiento como en el momento del propio tratamiento, medidas técnicas y organizativas apropiadas, como la seudonimización, concebidas para aplicar de forma efectiva los principios de protección de datos, como la minimización de datos, e integrar las garantías necesarias en el tratamiento, a fin de cumplir los requisitos del presente Reglamento y proteger los derechos de los interesados.



4. Retos y PoC

La elección por parte de un responsable de una infraestructura Blockchain concreta como elemento de su tratamiento podría provocar incumplimientos:

1. Rediseño del tratamiento: sin Blockchain, con otra infraestructura, etc.
2. Rediseño de la infraestructura.



¿Cómo rediseñamos las infraestructuras?

Mixing protocols

Blind/Anonymous signatures

Indistinguishability obfuscation (IO)

Homomorphic encryption

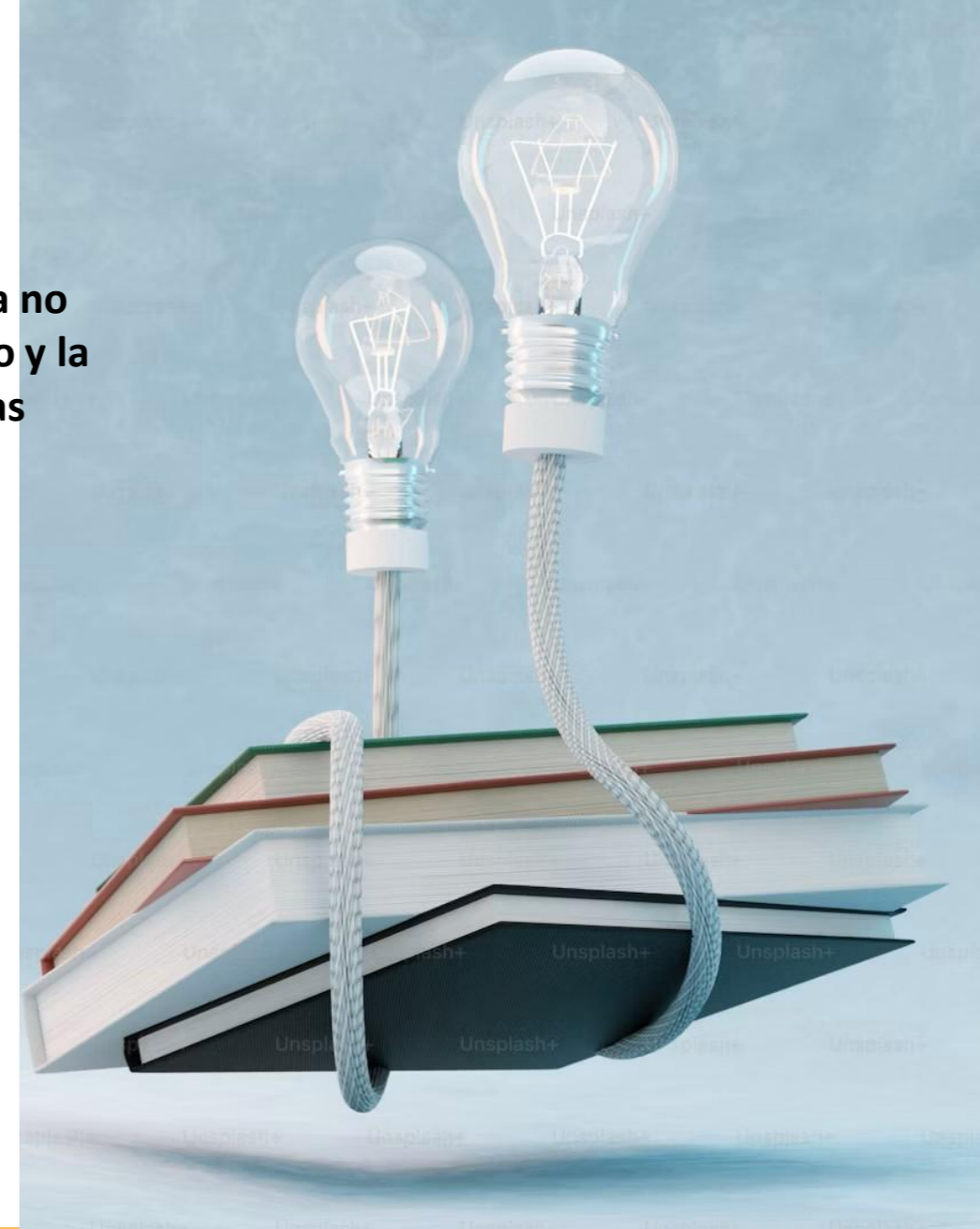
Trusted hardware-assisted confidentiality

Commitment schemes

Zero Knowledge Proofs

Todo muy orientado a la no vinculación, el anonimato y la confidencialidad de las transacciones

Ejemplos: CoinJoin, CoinSwap, TumbleBit, Dandelion, Pedersen commitments, MimbleWimble



Peng, L., Feng, W., Yan, Z., Li, Y., Zhou, X., & Shimizu, S. (2021). Privacy preservation in permissionless blockchain: A survey. *Digital Communications and Networks*, 7(3), 295-307.



¿Qué pasa con las amenazas relativas a la Inexactitud, al No repudio o a la Falta de transparencia y de capacidad para intervenir?

<https://www.aepd.es/guias/nota-tecnica-blockchain.pdf>

**Prueba de concepto
Blockchain y el derecho
de supresión**



Prueba de Concepto: derecho al olvido

Trabajamos sobre Ethereum (objetivo: probar, no diseñar una nueva infraestructura Blockchain desde cero)

- ❑ Se han creado varias cuentas de usuario y dos nodos validadores, se ha efectuado el despliegue de dos Smart Contracts sencillos, se han efectuado transacciones que los invocan. Adicionalmente, se han efectuado varias transacciones de traspaso de Ether entre usuarios.
- ❑ Solicitud de supresión de un usuario: origen y destino de transacciones de traspaso de Ether, destino de una transferencia de tokens en un Smart Contract, creador del otro Smart Contract y emisor de una transacción de creación de tokens.



Medidas técnicas y organizativas

1. Detección de los registros afectados en los distintos nodos

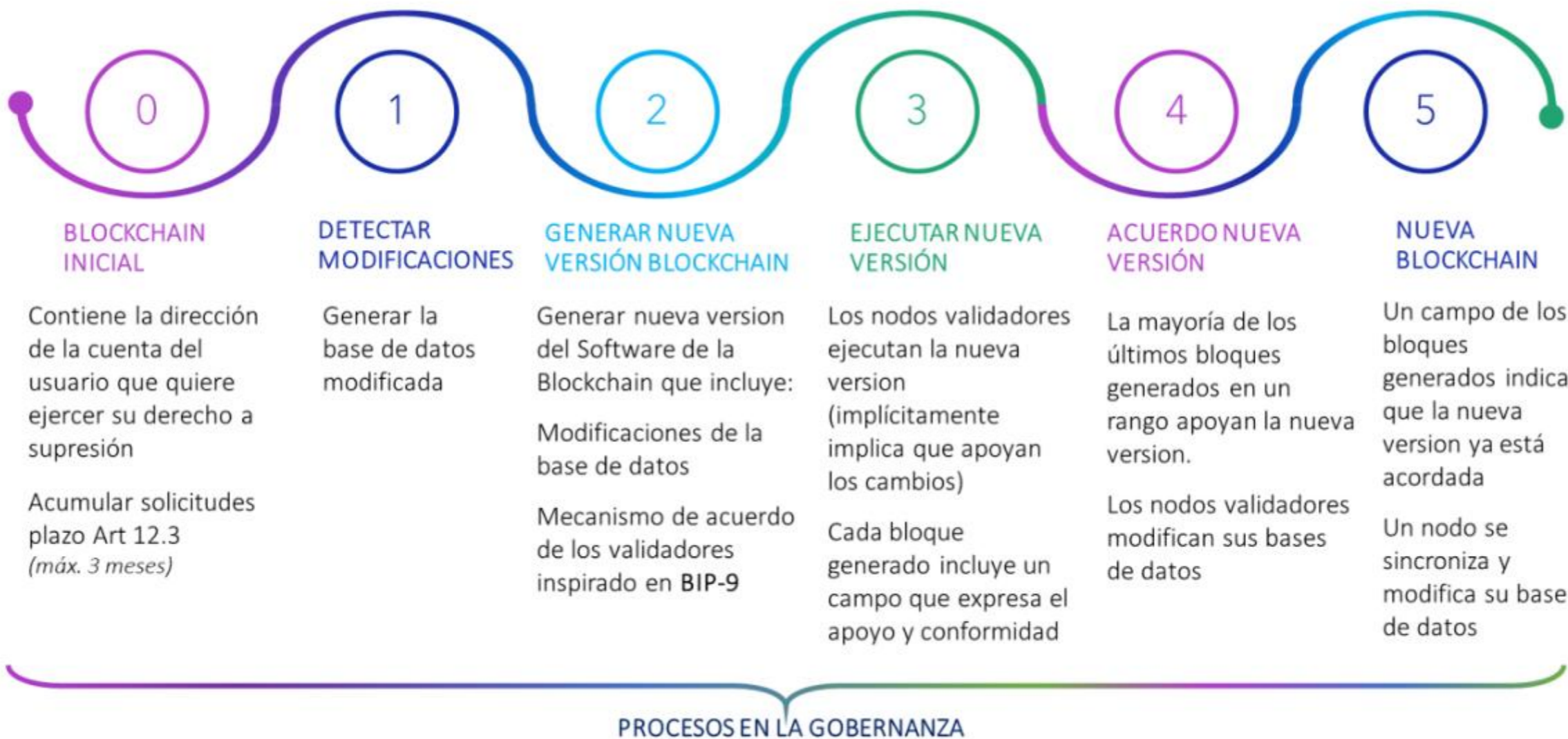
2. Generación de una nueva versión de software de la infraestructura Blockchain, implementando un Hard Fork

3. Distribución de la nueva versión de software y ejecución de este en los nodos

4. Mecanismo de consenso en la nueva versión de la infraestructura Blockchain

5. Medidas organizativas: gobernanza





¡Gracias! ¿Alguna pregunta?

La supuesta inmutabilidad del blockchain y el derecho al olvido

Marta Beltrán | mbeltran@aepd.es y Arturo Brazal | abrazal@aepd.es

[@mbeltranpardo.bsky.social](https://bsky.social/@mbeltranpardo)



**Reconocimiento-CompartirIgual 3.0
España (CC BY-SA 3.0 ES)**

©2025 AEPD

Algunos derechos reservados.

Este documento se distribuye bajo la licencia “Reconocimiento-CompartirIgual 3.0 España” de Creative Commons, disponible en <https://creativecommons.org/licenses/by-sa/3.0/es/>

Presentación creada con Visme (<https://www.visme.co/es/>)

Fotografías: <https://unsplash.com>

Iconos: <https://www.flaticon.es/>