GDPR Compliant Blockchain

Ankita Sharma, Privacy-Conscious Computer Systems

1. Purpose

The blockchain has caught the attention of many. By design, it is resistant to modification, tamper proof, and transparent. It uses decentralized consensus to maintain an append-only ledger that everyone can see. While the blockchain has made great strides integrating into our lives and enabling the existence of cryptocurrency, its fundamental properties may prevent its usage with policies like the GDPR coming into action that mandate that data stored by the data controller and processor should be deleted as part of a 'right to erasure' request. As a result, it is important to think about the future and coexistence of the blockchain and GDPR.

2. Idea

In order for the blockchain to accommodate requests like the 'right to erasure', private data should not be stored on the blockchain. Instead, this data should be stored off the chain. For applications that store user data, each block on the main chain would correspond to a new user that is using the service for the first time. The block would contain metadata associated with the "transaction" and a link to where user data will be stored for the remainder of time the user continues to use the application. By storing this data off the chain, we retain the fundamental properties of the blockchain by never running into a situation where a block needs to be removed or the entire blockchain reconstructed. However, to further extend this idea and to apply the blockchain's tamper proof properties to user data, which seems desirable, the user data that is stored off the chain can *also* be stored in a blockchain. This multi-dimensional design satisfies blockchain properties at the meta and granular level per user. To prove GDPR compliance, one would just need to delete the chain storing user data that is stored off the chain. The main blockchain can continue to operate and function as if nothing happened.

3. Premise

It is helpful to think about this idea with a scenario and how one might envision its adoption in the real world.

Hospitals are required to be GDPR compliant. Just last year, a hospital in Portugal was fined 400,000 euros for not being GDPR compliant. As a result, we can think about a hospital adopting a GDPR compliant blockchain as a mechanism for storing patient medical records.

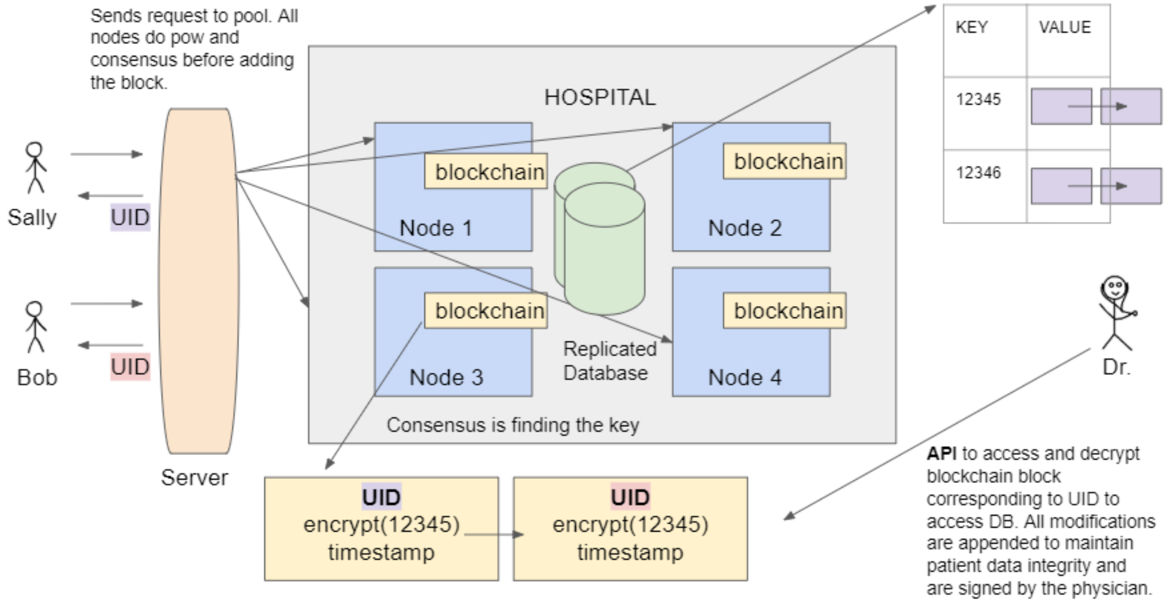
In this setting, the hospital would employ a private blockchain that is not open to the public, but is transparent and made available to hospital workers. When a patient visits the hospital for the first time, they are provided with a unique id. Their visit is recorded in the hospitals blockchain, maintained by a private decentralized cluster of nodes inside of the hospital. The block recording their first visit to the hospital contains an encrypted unused key to the hospital's central database that will store the patient's medical record data in an entirely separate blockchain. When a physician needs to access a patient's medical record, using the patient's unique id, will have the permission to decrypt the block, lookup the key in the database, and get the last block for the patient (a block is created if non exists). Every time a patient's block is accessed or modified, this is recorded as a new "transaction" on the blockchain.

4. Investigation

I plan to prototype this system end to end from a user first "opting" into the service, using the service, and finally "opting" out of the service in which case their data is removed. I will start small, making this system work with 4 distributed nodes that can each process client requests and then use a consensus algorithm[i] to agree on the "transaction" before adding it to their own internal representations of the blockchain. I will build an API that will allow people with the correct permissions to access sensitive data (i.e. the last block stored for that user in the db), modify it, and update the value stored in the central data base with these updates.

5. Assumptions

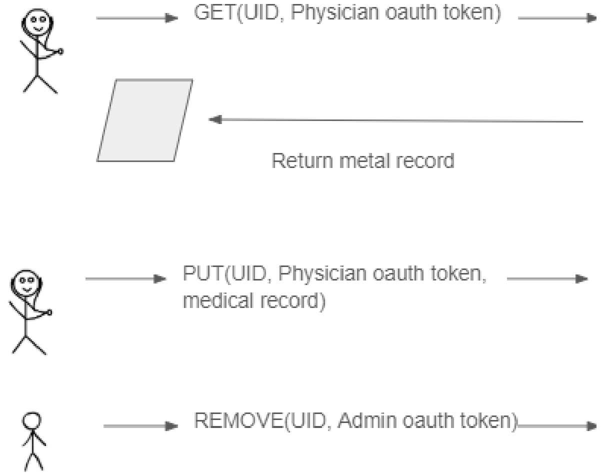
Every entry in the primary block chain will contain a unique id which will need to be provided for future transactions. The authentication and authorization for accessing the system could be enforced using mechanisms such as 2 factor authentication. The credential checking and enforcement aspects are beyond the scope of this project.



Sends request to pool. All nodes do pow and consensus before adding the block.

KEY	VALUE
12345	
12346	

API to access and decrypt blockchain block corresponding to UID to access DB. All modifications are appended to maintain patient data integrity and are signed by the physician.



GET(UID, Physician oauth token)
Return metal record
PUT(UID, Physician oauth token, medical record)
REMOVE(UID, Admin oauth token)

[i] I will be using the XRP leger as a reference.