Identified Problem

Today's way for white hat to alert a protocol of a bug is to use traditional communication tools. The first problem is that the bug information is critical and they have to trust the security of the used communication tool. Secondly, there is no public fingerprint that could be used to certify the skills and contribution of the white hat (especially in the case where bug bounties are involved and that the white hat expects to be paid!). Thus, the white hack has no choice but to trust and rely on the company to assess their contribution. This also tends to increase the shortage of white hats and cybersecurity contributions because of that unbalanced relationships between white hats and protocols.

Solution

I developed Safe Alert, a dApp that combines Privy, IPFS, Sismo, Kleros, Polygon and Alchemy technologies to create a secure way for white hats to create bug alerts and to have a public certification of their skills and contribution. We have thus the three following properties:

1. The bug information is secured by open source code so the integrity of the company is preserved
2. The dApp provide the white hats with public assets (NFTs) to prove with Kleros their contributions in case of dishonest behavior by the company
3. Those public assets are a useful tool for white hack to showcase them and to establish their public and professional reputation

How the dapp works

<u>1) Sending a secured Bug alert through Privy and IPFS</u>

To secure the bug alert we integrated Privy and IPFS to Safe Alert so that the information filled by the white hat is encrypted and hosted on IPFS, and is only readable by the involved protocol.

At the same time, we use NFT technology to mint a ERC-721 that contains within its metadata the link toward the encrypted data, a hash that acts as a proof that the encrypted data is unaltered, the blockchain address of the white hat alert creator, and a certification status that can take two values "certified" and "not certified".

<u>2) Using Kleros as a Bug Certifier to defend White Hat's interests</u>

To enforce justice, if the white hat encounters a problematic situation with the involved protocol (protocol didn't pay the bug bounty, protocol didn't certify the NFT related to the bug detection) the white hat can open a Kleros dispute where jurists will be granted access to the encrypted data to investigate and give their verdict.

The irrefutable asset the jurists will mainly use is the "proof of bug", filled initially by the white hat and contained in the encrypted data, that represents a set of actions that exploit the bug. Thus, that asset is providing an irrefutable and easy way for Kleros' jurists to give their verdict. Kleros' verdict is able to change the certification status of the NFT to "certified", thus publicly providing the white hacker with a proof of their skills and contributions.

3) Providing White Hats with Sismo Certification Badges “Proof of Hat”

With those certified NFTs, white hats are able to use Sismo to mint a badge, on the blockchain address of their choice (because they might want to separate their address on which they receive bug bounties and the address that acts as a public professional profile). That badge can have three attributes (gold, silver, bronze) depending on the importance and on the quality of the work of the white hat. That badge assess publicly the skills and contribution of the white hat and is used to build a professional reputation.

The dApp is built on top of Polygon, and we used the Alchemy API to interact easily and seamlessly with the Polygon blockchain to deploy our NFT.