# [NOT UPDATED]

# What are Storage Proofs and how can they improve Oracles?

## Introduction

Storage proofs are a cryptographic way to store blockchain information so that it can be shared across chains. Similar to oracles, they provide proof that the information is true. But unlike oracles, they do *not* require trust in a third-party for this proof—with storage proofs, the trust is built into the storage.

Storage proofs can, in some cases, replace oracles, and in others can enhance them. And they open up new blockchain use cases that weren’t possible before.

So let’s look in detail at storage proofs—what they are, how they work, their use cases, and how they can enhance (and sometimes replace) oracles

## What are Storage Proofs

**Storage proofs** are cryptographic commitments of state generated using zero-knowledge mathematics (such as SNARKs, STARKs, etc). These zk-proofs prove that a **particular state existed and was valid existed at a particular block in the past**.

Fundamentally, blockchains are databases that contain data cryptographically committed using Merkle trees, Merkle Patricia trees, Verkle trees, etc.). Since all the data is committed, we can prove that some information is encapsulated in a given state. However, with simple commitment schemes, the size of this proof becomes more prominent as the size of the data it includes becomes larger. Verifying such proofs on-chain becomes too expensive to be practical.

While fundamentally, blockchains can already do this (blockchains are databases that contain cryptographically committed data using Merkle trees, Merkle Patricia trees, Verkle trees, etc.) the simple commitment schemes of the chains cause the size of the proof to become more prominent as the size of the data it includes becomes larger. Verifying such proofs on-chain becomes too expensive to be practical.

Storage proofs, on the other hand, are relatively small, and allow you to verify a specific piece of state, at a specific point in time, on any domain—without having to trust a third party. Instead of third parties, they rely on the security of the underlying chain itself.

Why is this important? Ethereum today is not the simple monolithic chain (L1) from a couple of years ago. With the advent of layer two blockchains and rollups, the data is now spread across multiple chains.

Synchronous assumptions about the state of the chain can no longer be made. Many solutions for sharing data are now live such as L1->L2 messaging systems, cross-chain bridges, and oracles. But the issue with all these current solutions is that they include trust in a third party. Storage proofs allow us to validate the state of a blockchain at any point in time using cryptographic commitments assuming no trust in a third party.

## The Use Cases for Storage Proofs

Since storage proofs allow us to efficiently “compress” a blockchain and transmit the data elsewhere, they have quite a few applications. The cheap cost of verification, an integral property of storage proofs, allows the proof to be validated on the *destination* chain, eliminating the need to develop cross-chain messaging systems.

Potential use cases include:

* **General information transfer** from one chain to another about state and transactions on the blockchain.
* **Simplified cross-chain voting systems**. Frequently users hold their assets on a slow but more secure chain A, but some token-based voting occurs on a chain B that has cheaper transactions. This forces the user to either skip their vote or pay huge transaction fees to bridge their assets over from A to B, cast their vote, then bridge them back to A. In such cases, storage proofs enable users to prove their token balance on chain A at a given block and seamlessly cast their vote on chain B.
* **Alternative to cross-chain bridges**. Currently, cross-chain bridges assume a level of trust in a third party because they typically involve an intermediary, such as a custodian or a decentralized autonomous organization (DAO). This intermediary is responsible for ensuring that a certain amount of tokens are received on the source chain by the intermediary and for holding the assets on the source chain. Afterward, the corresponding tokens are minted on the destination chain. Storage proofs can enable trustless cross-chain bridges since a smart contract application on the destination chain could validate a transaction where assets were transferred to the bridge smart contract on the source chain and mint the bridged assets.
* **Enhanced UX for Account Abstraction (AA) use cases**. AA has been implemented in different chains and is considered a crucial innovation in onboarding the first billion users to the blockchain space. With storage proofs, wallets could include the additional functionality of restoring access only if the wallet did not send any transactions over a long duration. Additional checks that require some data to be used from other chains could also be enforced.

## What are oracles and why do we care about them ?

So how do storage proofs compare and contrast with oracles?

By design, blockchains cannot retrieve off-chain data. This keeps a blockchain trustless but also introduces limits on a smart contract’s ability to make decisions based on real-world events.

To solve this problem, special entities named **oracles** have been created to retrieve this off-chain data (or retrieve results from some heavy off-chain computation). Currently, these oracles require a third party (it could be an institution or a decentralised network of node operators) to submit data on-chain that becomes public to users and smart contracts. This assumption of trust is inevitable at the moment, yet not ideal.

[**Chainlink**](https://chain.link/)is an example of a blockchain oracle that provides a wide variety of real-world data (stock prices, weather data, etc.), off-chain computation services (to minimize the cost of heavy computations on-chain), and cross-chain services (that read and write information between different blockchains).

Since smart contracts have no other way of knowing what is happening in the real world except for using oracles, oracles have become an indispensable part of the blockchain ecosystem.

## The state of oracles on Starknet

On Starknet testnet, the previously mentioned Chainlink currently provides price data feeds for seven pairs of cryptocurrencies and has [partnered](https://www.coindesk.com/business/2023/02/06/starkware-partnering-with-chainlink-for-starknet-growth/) with the Starkware team to "further accelerate app development and general growth for the StarkNet ecosystem." Chainlink tries to minimise the assumption of trust by having a decentralised network of nodes that provide data from off-chain sources but the aggregation of this data happens off-chain.

[**Empiric**](https://www.empiric.network/)and [**Stork Network**](https://www.stork.network/) are two of the largest oracle providers on Starknet and live on mainnet and testnet. Along with the price tickers for multiple cryptocurrency pairs, the team at Empiric is working on implementing a verifiable randomness feed on mainnet that would allow protocols to request secure randomness on-chain. Price feeds on Empiric are based on price submissions by large institutions and market makers and the aggregation of prices happens on-chain leveraging efficient ZK technology.

## Can oracles be replaced or improved by storage proofs ?

In some cases, yes, a storage proof can replace an oracle.

Not all data that oracles provide actually needs to be supplied by a third party. In some cases, the data an oracle provides was already available on chain (in the form of on-chain storage, or a transaction) and can be retrieved by taking a peek at a past state of the blockchain. In these cases, a storage proof can replace the need for trust in a third party, and the oracle, and allow smart contracts to rely completely on the security of cryptographic commitments.

In other cases, where storage proofs can’t completely replace an oracle, they can often still *enhance* them with additional functionality.

* As discussed, oracles transmit information from data providers to data consumers. However, not all the consumers of this data are on the same chain. With the help of storage proofs, it is possible to complete some **computation on the data from different sources and export the result to other chains**. The source chain for such data is preferred to be the one with cheap computation, and validation of the proof can be cheaply done on other destination chains. [Herodotus](https://www.herodotus.dev) is one of the research leaders in this domain and they offer cross-domain data access across different Ethereum chains using storage proofs and zero knowledge maths. Empiric Network is also partnering with Herodotus to enable cross-chain oracle support in the near future.
* Storage proofs can **unify the state of multiple rollups**, and even allow synchronous reads *between* Ethereum layers.
* Another enhancement is a **trustless retrieval of historical data published on-chain**. By leveraging the fact that stateful blockchains such as Ethereum and Starknet record and cryptographically preserve their state through specialized data structures (Merkle/Verkle trees, MPTs), it is possible to prove the inclusion of any data stored in these structures. Hence, any past data published on-chain can also be trustable retrieved and used in other applications (not even necessarily on the same chain).
* Empiric Network is researching the viability of developing an oracle as an **L3** on Starknet from where data can be “pulled” on other chains and verified using storage proofs.. The benefits of having the oracle in a different domain (on top of a computationally cheap network like Starknet) are:
* Since the L3 could be a highly customizable chain, various parameters can be tweaked to **achieve consensus much quicker on the blocks**, greatly reducing data latency for the oracle.
* In combination with storage proofs, the **low-latency data can be asynchronously transferred to other chains** (as soon as consensus is reached in the source chain).
* The possibility to **enhance trust in data** by developing an inbuilt system in the L3 to slash dishonest data providers. If given appropriate incentives, the data providers on the L3 could stake their assets as a guarantee of publishing correct data. Since the consensus of the whole network on L3 is needed before other chains can utilize the data, the data provided by the oracle can be considered to be secured by the validators' stake on the L3.

## Conclusion

Over the past few months, the growing use of L2s on Ethereum has given us a clearer view of the industry's future. The L2 narrative has been gaining traction (with networks like [Starknet](https://www.starknet.io/en), [Optimism](https://www.optimism.io/), and [Arbitrum](https://arbitrum.io/). However, one of the primary backstops for its growth has been implementing a decentralized cross-chain messaging system. Storage proofs, though at a nascent stage, promise incredible improvements to this problem.