# Morphée: a coordinated and automated brokerage layer on top of Filecoin

Polyphene

contact@polyphene.io

June 22, 2022

**Abstract**

*While the Filecoin network has proven to successfully operate since 2020 and the launch of its mainnet, the realization of its full potential is still in its early stages. Among the leads for the near future, the configurability of its virtual machine for instance offers the promise of automated brokering systems. This light paper tries to first show why the functioning of such a system cannot be considered as technologically nor politically neutral and starts to lay down the first elements of the design of a coordinated and automated brokerage layer on top of Filecoin. After synthesizing the elements that can reasonably be fixed at the time of writing, it provides a quick overview of the remaining limitations and open discussions before such a project can eventually be realized.*

## I. Introduction — Of the importance of digging the Web3 channel

Since the launch of Bitcoin in 2009, the types, promises, and uses of blockchains have diversified widely. The emergence of the Ethereum blockchain and the Solidity smart contract language in 2015 for example undeniably marked a major milestone.

In parallel and for various motives such as the search for improved scalability, different variants of these technologies have continually surfaced, making the deliberate choice to cut back definitively on the openness and decentralization of the network, yet one of the elements of the so-called blockchain trilemma[1]. Today this dichotomy is accentuated around the definitions and soundness of the notions of *closed metaverse* and *open metaverse*. It is however undeniable that in the continuity of the development of the most open blockchain networks, the potential of Web3 technologies is far from being fully explored. Subjects of money, finance, or digital objects — with the advent non-fungible tokens (NFTs) — have probably already been scratched. But if, as it is now often the case in the literature, the spectrum of their use is broadened to a much wider range of types of data, by starting to consider NFTs as *Web3 backpacks* that may be tied to any data stream, then it becomes conceivable, and even necessary, to invest more seriously in the subjects of identity, self-sovereign identity (SSI), and the storage of data associated with the experiences of these identities.

Alongside other protocols such as Arweave[2], new networks that can be described as blockchain-enabled cooperative storage clouds have emerged, most notably the Filecoin blockchain, in conjunction with the InterPlanetary File System (IPFS). This undeniable technological breakthrough has now proven sound operation on its mainnet for more than a year and is storing over 115 PiB of data[3]. While the current version of the protocol constitutes in itself a major achievement and as attention could hazardily start to focus on relatively siloed

---

[1]The blockchain trilemma refers to the practical problem met when looking for a balance between the three elements expected to characterize a perfect blockchain: decentralization, security, and scalability.

[2]Arweave is a protocol focused on providing permanent storage.

[3]Numbers extracted by file.app analytics, on June 19th, 2022.

uses, exploiting above all a competitive advantage on the competitive pricing of large datasets archiving, it seems at least equally important to continue to develop the potential of this network in more purely Web3 oriented uses. In that sense, focusing on coordination functionalities in the network or interactions with other Web3 protocols represents both a challenge and an opportunity. While since the end of 2021, the development of the new Filecoin Virtual Machine (FVM) is pushing positively in this direction, the present Morphée project aims at digging more precisely into the coordination topic and at crystallizing compelling ideas around one dedicated project.

## II. The hows and whys of Morphée

### i. Initial rationale for a brokerage layer

The Filecoin blockchain is a great tool to instill trust, without a third party, between a user willing to store data on the IPFS network and a provider capable of fulfilling that request. The blockchain allows a storage provider to prove that it initially received the data to be stored and to then continuously demonstrate, throughout the duration of the contract, that it maintains the storage of this data at the expected location. Otherwise, the machine would inevitably inflict an economic penalty. This is essentially the real technological breakthrough that Filecoin has already achieved.

At present, therefore, buyers and providers maintain a large part of their exchanges off-chain. The encounter of the storage buyer and the provider, the majority of the discussions before signing any deal, are indeed done off-chain. However, the more the quality of service of the suppliers will be attested directly on-chain, and the more in particular the reputation systems of the providers will be decentralized, the easier it will become to increase the level of responsibility and intelligence available directly on the blockchain. We could indeed relieve the customer of the burden of choosing their supplier manually, and let them dictate this choice according to reputation metrics, dictating the storage parameters with an adequate language and *profiles* rather than directly specifying the precise list of their providers. This should realize another step towards a truly open place of storage supply across the planet (or planets).

In short, this gain in expressiveness should initially serve clients to dictate their storage requirements and let the trustless machine, over the course of time, choose the supplier(s) best able to fulfill them, rather than having to make their choices themselves.

### ii. Primitives of the brokerage layer

Formalizing this system requires some new lexical items. We will say that providers describe and propose *racks* while clients handle *buckets*. A bucket is characterized by its content, a rack by its capacity, and both by their configuration.

A configuration, real for a rack and expected for a bucket, is described according to a set of shared primitives calibrated on the following characteristics:

- geographic location
- latency and throughput performance, tied to the ability or will to store unsealed copies
- availability
- durability
- greenness, i.e. the ability to store data with a low carbon footprint
- providing an HTTP gateway
- pricing conditions

While these characteristics are fixed on the provider side, they pave the way on the storage buyer's for more programmability of their storage configurations, with higher-level storage *profiles*. This gain in expressiveness should translate into plasticity of storage arrangements, as for instance the set of providers meeting a customer's particular expectations may change over time in a way ultimately transparent to the customer. The notion of storage *profiles* also makes it possible to consider certain new use cases — such as setting geolocalized performance requests that would result

in requirements for both performance and the geographic distribution of replicas.

### iii.  Embracing the non-neutreality of the reconciliation algorithm

Matching buyers' expectations with providers' supply over time is the logical responsibility of the reconciliation algorithm. By introducing this component, the second main rationale behind Morphée is touched upon: the motives of this algorithm are not neutral. Should it aim at minimizing customers' expenses? Or on the contrary to maximize the global profits of suppliers? Should it favor suppliers using decarbonized energy and if so to what extent? Should it even economically favor the storage of certain data sets over others?

These are just the premises of real questions that need to be asked and which answers could potentially apply to the entire network. Answers that could be found, calculated by an evolutionary but unique network-wide reconciliation algorithm.

Morphée therefore proposes to embrace the necessary non-neutrality of such an algorithm, which on the one hand makes possible a gain in expressiveness and intelligence useful to individuals' use of Filecoin, and on the other hand allows to guide at the global scale the objectives of development and use of storage resources. Therefore, Morphée chooses, from the outset, to make this algorithm configurable and to entrust its governance to the most legitimate community for this purpose. This is the objective of Morphée, by providing the protocol, the network and the tools that unlock collaboration in order to jointly achieve both individual and collective objectives.

While this section has laid the groundwork for the design of Morphée, the remainder of this document provides an attempt to describe the first technical elements that may now be outlined.

## III.  EARLY TECHNICAL DESIGN CONSIDERATIONS

### i.  Optimizations based on the content of buckets

In the chosen design, the very content of the buckets can be taken into consideration during the execution of the reconciliation algorithm. This makes possible, for example, to globally favor the storage of certain datasets deemed to have higher priority, or to factor out the storage costs of datasets or portions of datasets targeted by several clients at the same time. While standards such as Content Identifiers (CIDs) and Interplanetary Linked Data (IPLD) should increase the potential impact on this part of the protocol, the implementation details are yet to settle concerning in particular the availability of this metadata, i.e. the availability of the information allowing to efficiently identify common parts in the content of different buckets.

Nevertheless, this problem can be considered as logically totally distinct from other optimization problems. It can be considered as a pre-processing step to transform an initial set of buckets into a new set, optimized with respect to their content.

### ii.  Setting the reconciliation problem

The issue of reconciling buyers' expectations and providers' capabilities is written as a combinatorial problem. More precisely, it is written as an assignment problem which tasks and agents are parameterized from different sources:

- the internal state of the machine, made of solutions of past iterations of the problem
- contextual information coming from the Filecoin mainnet, e.g. the status of tracked storage deals
- bucket profiles
- and rack profiles

### iii.   Around a solution to the reconciliation problem

The reconciliation algorithm is in charge of solving a maximum generalized assignment problem.

It is known that a good approximation of the solution can be found in polynomial time. [Fleischer, Lisa, et al., 2006]

Initially, this algorithm should be able to rationally run on one (or more) sidechain(s) connected to the Filecoin mainnet via optimistic rollups.

## IV.   Conclusion

The Morphée light paper lays the groundwork for what a coordinated and automated brokerage layer could look like. While the Filecoin blockchain has proven to work in production for more than a year, it would now seem essential to look beyond its often siloed applications and to work on more collaborative uses of the network.

At the same time, many complementary projects are advancing and enriching an ecosystem that is increasingly capable of realizing this complete vision, with inter-chain communication, layer-2 solutions, decentralized reputation integrating first and foremost energy and ecological data, etc.

The mission of this light paper and the team originally behind it would be realized if it could help to structure the debates that would lead to the emergence of a real coordinated and automated brokerage layer. The mission would be fulfilled if one day such a layer arises that continuously and simultaneously meets the objectives of each individual and those shared by the global network.

## References

[Fleischer, Lisa, et al., 2006] Fleischer, Lisa, et al. (2006). Tight approximation algorithms for maximum general assignment problems. *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm.*