Degrees of Decentralized Freedom: Comparing Modern Decentralized Storage Platforms

Zhengyu Wu, Brian Lin, Nick Nikiforakis, Aruna Balasubramanian
Stony Brook University
Stony Brook, New York, United States
Emails:{zhenwu, brialin, nick, arunab}@cs.stonybrook.edu

***Abstract*—Decentralized storage platforms distribute control across individual peers, thus reducing reliance on a single entity and mitigating common vulnerabilities of centralized storage systems. In this paper, we compare the architecture and operation of four popular decentralized storage platforms: IPFS, Filecoin, Swarm, and Storj. Our study reveals significant implementation differences in four key aspects: data routing, data persistence, incentivization mechanisms, and resource requirements. These architectural decisions directly influence network characteristics, performance metrics, and economic sustainability.**

We collect comprehensive snapshots of the entire network to analyze network properties including peer uptime, geographical distribution, and network availability. Our analysis shows that while IPFS maintains the largest user base, it exhibits the lowest peer uptime due to lack of incentivization, with 50% of peers online for less than 4 days. In contrast, incentivized platforms exhibited median peer uptimes around 80-96% of the study period. We found considerable performance variations that directly correlate with implementation choices. Storj, with its centralized data routing architecture, achieves performance nearly on par with centralized solutions like Google Drive. In contrast, Swarm showed the slowest performance metrics with its full commitment to decentralization. Finally, our analysis reveals that cryptocurrency price fluctuations significantly influence participation and cost, suggesting potential sustainability challenges in these decentralized storage platforms.

***Index Terms*—content addressing, decentralized storage, peer-to-peer system, Interplanetary file system**

I. Introduction

The rapid increase in centralized digital infrastructure has raised significant concerns regarding data privacy, security, and user autonomy [1], [2]. According to Sandvine's 2023 Global Internet Phenomena Report, major corporations such as Google, Netflix, Facebook, Apple, Amazon, and Microsoft account for over 57% of global internet traffic [3]. This centralization leads to single points of failure [4], easier censorship [5], and increased risks of data breaches [6].

In response, decentralized storage platforms have gained increasing adoption in recent years. These platforms aim to distribute control and authority away from centralized entities, allowing individuals and communities to govern digital infrastructure. For instance, the Inter-Planetary File System (IPFS) provides storage services to more than 3 million web clients [7], with billions of files shared daily [8]. Similarly, Filecoin [9] is being used to store all data [10] for Internet Archive [11] and SETI Institute [12].

Despite their growing importance, there is a notable absence of comprehensive studies comparing these decentralized storage platforms. Existing works [7], [13], [14] typically focus on specific aspects of individual platforms without providing a holistic understanding of how design decisions influence network characteristics, performance, and long-term sustainability.

978-3-903176-74-4

We analyze four active decentralized storage platforms in this work: IPFS [15], Filecoin [9], Storj [16], and Swarm [17].

Although these platforms adopt fundamentally different incentive and economic models, ranging from free-use to token-based rewards, we treat them uniformly as decentralized storage systems because they share the core goal of enabling distributed, peer-operated data storage at scale. Our analysis reveals distinct design choices across four critical dimensions: data routing (how content is discovered and retrieved), data persistence (how content availability is maintained), incentivization (how participants are motivated to contribute), and resource requirements (hardware and computational demands for participation). Figure 1 showed each platform's position across the spectrum of each dimension. We discuss this more in our analysis §II-B. We conduct a comparative analysis of the four platforms and their specific design choices using three research questions:

1) **What are the adoption patterns of these systems, and who are their users?** We explore network size, peer uptime, geographical distribution, and infrastructure choices to understand participation patterns.
2) **How do these systems perform compared to centralized solutions?** We evaluate upload and download performance across various file sizes and geographic locations.
3) **How does incentivization influence adoption and long-term sustainability?** We analyze cryptocurrency price effects on network participation and examine token distribution patterns.

For the first question, we collected network data over a 50-day period, capturing 101 snapshots of each platform's complete network structure. Our analysis shows that resource requirements primarily determine network size, while incentivization significantly impacts availability. IPFS, with moderate resource requirements and free-to-use model, maintains the largest network (23,000 peers). In contrast, Filecoin's high resource/participation requirements limit participation to only 1,700 peers. In terms of network availability, cryptocurrency incentive plays a crucial role. Platforms without financial rewards showed 50% of peers online for less than 4 days, while incentivized platforms exhibited median uptimes of 40-48 days. We studied the availability patterns between 2024 and 2025; our data shows that IPFS's peer availability declined from 60% to 40%, highlighting the sustainability challenges of systems lacking economic incentives. Filecoin and Swarm also showed decreasing network sizes as cryptocurrency prices fell.

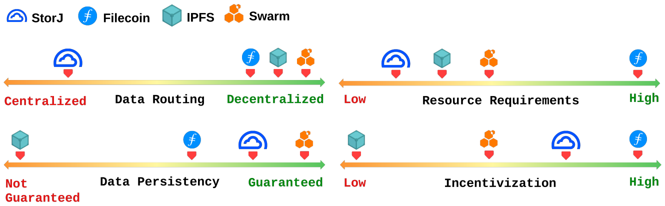

Fig. 1: *Design choices of IPFS, Storj, Swam, and Filecoin that falls on spectrum of centralized-to-decentralized data routing strategies, guaranteed-to-not guaranteed data persistency, high-to-low resource requirements, and high-to-low incentivization.*

We conducted extensive upload and download tests across multiple file sizes and geographic locations: data routing architecture and persistence strategies predominantly influence performance metrics. Storj's centralized satellite-based architecture achieves upload speeds (1.18s for 5MB files) nearly matching the centralized Google Drive performance (0.94s). In contrast, Swarm's full decentralization with small 4KB chunks results in dramatically slower uploads (1002s for 512MB files)—up to 100 times slower than centralized alternatives. Download performance follows similar patterns, with Storj offering the best throughput among decentralized solutions but still falling slightly short of the centralized Google Drive performance.

In terms of incentivization and sustainability, we analyzed how different cryptocurrency reward models and token price fluctuation impact network growth, stability, and pricing. We found that while systems like Filecoin and Storj offer significantly cheaper storage prices than traditional storage systems like Amazon AWS, but they also come with different sustainability challenges. Filecoin's blockchain-centric design requires substantial hardware investment, creating a direct relationship between token value and capacity. When FIL price fell from \$200 to \$6, storage capacity decreased as providers exited the market. Storj, however, leverages a low entry barrier model that produces the opposite effect. As STORJ token value declined from \$3 to \$0.5, network capacity grew from 20PB to 100PB as providers deployed more storage to maintain earnings. These economic structures translate directly to pricing differences. Filecoin offers extremely cheap storage at \$1/TB/month due to its low crypto coin price and high storage capacity. On the other hand, Swarm commands \$335/TB/month because of the low supply. This demonstrates how incentive design shapes both provider sustainability and client affordability.

Our findings suggest that there is no universally superior decentralized storage solution. Instead, each platform offers distinct trade-offs between decentralization, performance, and economic sustainability based on the design decisions on data routing, persistence strategies, incentive mechanisms, and resource requirements. This work provides users with a framework to evaluate platforms based on their specific needs while offering researchers insights to inform next-generation decentralized storage solutions. The dataset and code can be found at GitHub SBUNetSys/decentralized-storage-analysis

II. A Survey of Decentralized Storage Platforms

A. Decentralized storage platforms in the market

For the analyses presented in this paper, we first compiled a comprehensive list of existing decentralized storage platforms available in the market. Table VII (in the Appendix) shows the storage platforms including their current operational status. Two out of the twelve platforms are still in the development phase. Table I shows the network and storage capacity of the seven active platforms. We obtain the information in the table directly from the official website of each platform. Out of the ten active platforms, SAFE, Holochain, and Hypercore are just rolled out into production, and there is less information available. Internext operates solely under the control of a single company without public access to its operational data. Sia and Arweave are extremely small in terms of network size. Given our focus on the network performance of open decentralized storage solutions, we choose the remaining four platforms for our analysis—IPFS, Filecoin, Storj, and Swarm.

TABLE I: *Network and storage size of active storage platforms.*

Platforms	Network Size	Stored Storage Size
IPFS	~25k [18]	N/A
Filecoin	~1.7k [19]	1.64 EiB [20]
StorJ	~23.8k [21]	15.4 PB [22]
Swarm	~16k [17]	-
Sia	~500 [23]	1.86 PB [23]
Arweave	~90[24]	167.12 TB [24]

B. IPFS, Filecoin, Swarm, and Storj

Decentralized storage platforms have a common goal: to address the limitations of centralized storage platforms including the problems associated with single points of failure. Each platform makes design decisions on the following aspects: (i) data routing, including how to store and retrieve data in the decentralized network; (ii) ensuring data persistence without the need for a central authority; (iii) incentivizing storage participants and pricing strategy; (iv) the resource requirement to participate and use the system.

TABLE II: *Overview of the different design choices made by the four decentralized storage platforms.*

Platform	Data Persistency	Data Routing	Incentivization	Resource Requirements
IPFS	Not Guaranteed	Kademlia DHT	N/A (Free)	Medium
Filecoin	Guaranteed by Storage Provider	Kademlia DHT	Cryptocurrency (FIL)	High
Swarm	Duplication, Erasure Coding	Kademlia DHT	Cryptocurrency (BZZ)	Medium
Storj	Erasure Coding	Satellites	Cryptocurrency (StorJ)	Low

TABLE III: *Recommended resource requirement for IPFS, Swarm, Storj, and Filecoin*

	CPU	RAM	Storage	GPU
IPFS	2 core	6 GB	N/A	N/A
Swam (Light Mode)	Any	Any	N/A	N/A
Swam (Full Mode)	2 Core	8 GB	30 GB SSD	N/A
Storj	1 Core	2 GB	2 TiB	N/A
Filecoin	8 Core	256 GB	2 TiB Nvme SSD	Nvidia GPU with at least 11GB VRAM

We first provide an overview of the four platforms. Figure 2 shows the overall flow of IPFS, Filecoin, and Swarm which has a similar high level structure (but the design details vary). Storj uses a different structure captured in Figure 3.

We will also describe the design choices for each platform. Table II provides the details of the design choices, and Table III shows a

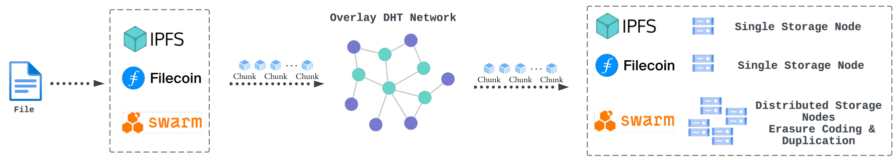

Fig. 2: The overlay network architecture used by Filecoin, IPFS, and Swarm.

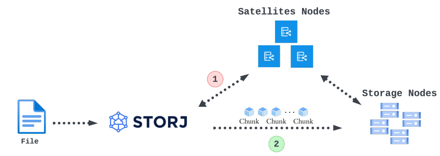

Fig. 3: The overlay network architecture used by Storj.

detailed resource requirement for each platform. We will categorize the design choice on a spectrum from centralized-to-decentralized data routing strategies, guaranteed-to-not guaranteed data persistency, high-to-low resource requirements, and high-to-low incentivization. The goal of this work is to study how the design choices influence the performance and scalability of decentralized storage systems.

a) The **Interplanetary File System (IPFS)** developed by Protocol Labs [25], uses a peer-to-peer network for decentralized storage and retrieval of data [26]. As shown in Table II, IPFS's data routing network is based on the Kademlia Distributed Hash Table (DHT). DHT network enables decentralized routing through peer communication and so can be considered to be more on the decentralized end of the data routing spectrum (Figure 1). IPFS promotes a free and open web by using a non-incentivized model, where participation is voluntary and storage is not rewarded. This lowers entry barriers but makes data persistence unreliable, as content can disappear when peers go offline. All uploads are publicly accessible by default, reflecting IPFS's openness-first design.

b) Filecoin, developed by Protocol Labs [25], builds on the IPFS architecture and shares its core use of the Kademlia DHT for decentralized routing. Unlike IPFS, Filecoin introduces a blockchain-based incentive layer, rewarding storage providers with its native cryptocurrency, FIL, to ensure long-term data persistence.

Filecoin operates on its own dedicated blockchain, where all transactions and storage proofs are recorded on-chain. This design provides strong reliability guarantees and verifiable storage but requires substantial computational resources. Participants must engage in consensus and maintain the blockchain state. Filecoin primarily targets enterprise and institutional clients with large-scale storage needs [27], [28]. Notable adopters include the Internet Archive, which stores over 1 petabyte of cultural heritage data [11].

c) Similar to IPFS and Filecoin, **Swarm** uses a peer-to-peer overlay network built on the Kademlia Distributed Hash Table for data routing. Initially developed within the Ethereum Foundation as part of the Ethereum ecosystem, Swarm targets decentralized applications (dApps) and Web3 services that require censorship-resistant storage [29]. Unlike IPFS and Filecoin, Swarm implements data persistence through both content duplication and erasure coding (Figure 2), allowing data reconstruction even if some fragments are lost.

Swarm uses the BZZ token, issued on Ethereum, to incentivize storage and retrieval, making it a fully token-driven system. This economic model enables stronger persistence guarantees and availability, but also introduces blockchain interaction overhead and moderate resource demands for participation. Swarm's design places it on the decentralized end of the routing spectrum while committing to long-term data retention through economic incentives (Figure 1).

d) Unlike the other platforms that use Kademlia DHT for routing, **Storj** uses centralized coordination points called "Satellites" to manage storage and retrieval operations (Figure 3). Developed by Storj Labs [30], Storj targets businesses and developers seeking privacy-preserving, cost-efficient alternatives to traditional cloud storage.

Satellites handle metadata indexing, user authentication, and coordination but do not store data themselves. This creates a quasi-centralized architecture, as all routing must pass through operator-controlled Satellites. While this improves performance and reliability, it places Storj toward the centralized end of the routing spectrum (Figure 1).

For data persistence, Storj applies erasure coding and encryption to uploaded files. It splits the files into shards and distributes them across the network nodes. Like Filecoin and Swarm, it rewards providers using the STORJ token on the Ethereum blockchain. However, only Satellites interact with the blockchain. As a result, participants face lower resource requirements (Figure 1), which improves accessibility but reduces decentralization in the design.

III. Data collection and Network analysis

Our first aim is to understand the adoption patterns and user characteristics of each decentralized storage platform.

A. Collecting network snapshots

We collected data from March 26th of 2024 to May 15th of 2024, a period of 50 days, and obtained the following information: 1) total number of peers in the network, 2) each peer's IP address, and 3) the peer's neighbor information. We collect this data every 12 hours so in total we have 101 snapshots of the entire network for each platform. We also collect additional data from March 26th of 2025 to April 16th of 2025, total of 50 days, to observe how each system changes after about one year.

a) IPFS & Filecoin: Both IPFS and Filecoin use an overlay DHT network. We get a snapshot of the network by performing a Breadth First Search. We start with the root node, query its neighbors, and continue querying each neighbor's neighbors until no new peers are discovered. To this end we utilize Nebula, a crawler designed to navigate through overlay DHT networks [31] and which has also been used in previous IPFS studies [7].

b) Swarm: While Swarm also uses Kademlia DHT for its overlay network, the DHT works differently compared to IPFS and Filecoin. We relied on swarmscan [32], a monitoring tool for the Swarm network to retrieve all peer information.

c) Storj does not use a DHT architecture and there are no tools available to query the network. We designed a custom crawler, taking advantage of Storj's satellites upload API which returns a list of available peers. We use synthetic upload requests to the satellite nodes and record all returned peer data. The crawler continues sending requests until it receives no new peer information for ten consecutive rounds.

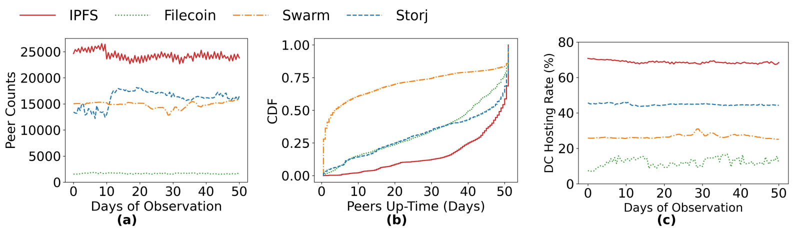

Fig. 4: *Network characteristics for all decentralized storage platforms over 50 days.* ***(a)*** *Daily network size counts.* ***(b)*** *CDF of peer up-time.* ***(c)*** *Daily data-center hosting rate.*

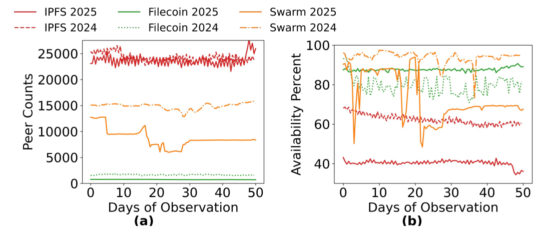

Fig. 5: *Network characteristics comparison (2024 vs 2025) over 15 days for Filecoin, IPFS, and Swarm.* ***(a)*** *Daily network size counts.* ***(b)*** *Daily network availability.*

B. Network characteristics

Network size directly reveals each platform's adoption level and is heavily influenced by design choices related to resource requirements and participation barriers. Figure 4(a) shows the day-to-day variation in network size for each platform.

IPFS maintains the largest network with an average of 23,000 peers per day, consistent with previous studies [7], [33], [34] and official reports (Table I). This high adoption is thanks to IPFS's medium resource requirements and free participation model. However, the same design choice leads to significant network size fluctuations over time, which we discuss later in this section.

At the opposite end, Filecoin has the smallest network with only 1,700 peers, despite being built on IPFS technology. This limited adoption directly results from Filecoin's blockchain-based design requiring substantial computational resources. Each node must verify blockchain transactions and maintain chain integrity, with blockchain data reaching 2.4 TB [35] and growing by up to 50 GB daily [36].

Storj and Swarm show moderate adoption levels with 18,000 and 15,000 peers respectively. Storj's network size reflects its balanced approach with low resource requirements. However, though our observations suggest that its satellite-based architecture introduces information asymmetry. When upgrading from a free to paid subscription, we observed an increase in discoverable peers, indicating that satellites may limit peer information based on account status. This design choice creates a situation where a single entity controls network visibility, potentially affecting reported network statistics and explaining the difference between our count (18K) and official figures (23K). For the purpose of this work, we perform our analysis on the 18K peers that we obtained through our crawl.

Swarm maintains consistent peer numbers around 15,000, lower than IPFS and Storj despite its comparable resource requirements. This difference in network size is attributable to Swarm's architecture, which requires more resources for running full nodes that participate in the storage network.

Network Availability. Incentivization mechanisms play a crucial role in determining peer availability and network stability. Uptime is defined as the duration a peer remains reachable across our snapshots. For instance, if a peer is present and reachable in two of our snapshots but does not appear or is unreachable in subsequent ones, we consider its uptime to be one day, given that we conduct our snapshots every 12 hours.

Figure 4(b) shows peer up-times across different decentralized storage platforms. IPFS, which lacks direct economic incentives, shows the poorest availability with 50% of peers online for less than 4 days. In contrast, the other three platforms employing cryptocurrency-based incentives demonstrate significantly higher peer stability. Swarm shows the highest uptime with a median of 48.5 days, followed by Filecoin and Storj with median uptimes of 45 and 40 days respectively. By rewarding peers for maintaining availability, these platforms create motivation for consistent participation, resulting in more stable networks.

Peer Geo-location. Geographic diversity can impact user experience through latency effects and also reflects how successfully each implementation has attracted a worldwide user base. Figures 6-9 show the geographical distribution of peers for IPFS, Filecoin, Swarm, and Storj.

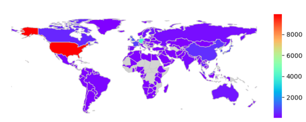

Fig. 6: *IPFS Peer's IP geo-location heat map with presence in 185 countries.*

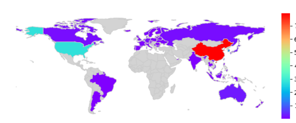

Fig. 7: *Filecoin Peer's IP geo-location heat map with presence in 49 countries.*

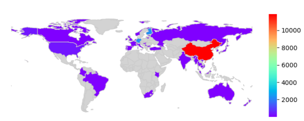

Fig. 8: *Swarm Peer's IP geo-location heat map with presence in 41 countries*

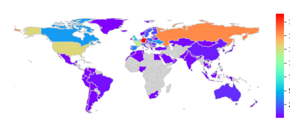

Fig. 9: *StroJ Peer's IP geo-location heat map with presence in 104 countries*

There is a significant difference in geographical reach across platforms. IPFS shows the most global presence with nodes in 185 countries. This widespread adoption is due to its minimal barriers to participation. In contrast, Filecoin's peer distribution shows a more concentrated presence primarily in the U.S. and China. This concentration results from its high computational requirements, which effectively limit participation to regions with advanced server infrastructure.

Swarm exhibits a similar geographic pattern to Filecoin, with significant node clusters in China and Europe, and presence in only 41 countries—the lowest among all platforms studied. This limited distribution may impact content availability and retrieval performance for users in underrepresented regions. Storj presents a more balanced distribution across Europe, North America, and Asia (with a concentration in Europe), spanning 104 countries.

Hosting Infrastructure. Datacenter hosting refers to the use of commercial data centers to serve as peers in decentralized storage networks. While professionally managed infrastructure offers advantages in reliability, up-time, and connectivity, they also re-introduce the issues associated with centralized services.

To quantify hosting patterns across platforms, we rely on reverse lookups using Udger [37], an IP database that identifies data center hosts. This methodology has been validated in related work [7], [33]. To ensure accuracy, we cross-verified results using two additional services (ipinfo[38] and ip-api[39]), achieving over 99% labeling agreement.

Figure 4(c) shows the datacenter utilization across platforms. IPFS shows the highest rate of data center hosting, consistently approaching 80%. This finding aligns with previous studies [33]. IPFS's free-to-use model and moderate resource requirements make it particularly suitable for data center deployment, where operators can easily set up and maintain nodes without significant investment or specialized hardware. Further analysis of the ASNs reveals that 69.41% of these datacenter nodes are concentrated among just four providers: Constant [40], Contabo [41], AmazonAWS [42], and HostPapa [43]. This significant concentration raises concerns about whether IPFS is achieving its intended goals of true decentralization and censorship resistance, as these dominant entities could become centralized points of failure or control within the network.

In contrast, Filecoin demonstrates the lowest datacenter hosting rate at approximately 20%. This is due to the requirement of high-performance computing for blockchain verification. This forces participants to deploy custom hardware rather than rely on standard cloud infrastructure with recurring costs. Swarm and Storj fall between these extremes with datacenter hosting rates of 40% and 50% respectively. Their moderate resource requirements allow for some datacenter hosting, but their cryptocurrency incentive models also encourage a significant portion of participants to maintain their own infrastructure for greater control and potentially higher rewards.

C. Network characteristics in 2024 vs 2025

We compared network characteristics between data collection in 2024 and a subsequent collection in early 2025 which revealed important insights into platform sustainability.

Figure 5(a) shows the network size of IPFS, Filecoin, and Swarm in 2025 relative to 2024. We omitted Storj as its network remained stable. IPFS maintained consistent popularity with approximately 23,000 peers, yet Figure 5(b) reveals a concerning decline in peer availability from 60% to 40%. This illustrates the sustainability challenge of systems lacking economic incentives. Without financial motivation, fewer peers maintain continuous participation despite stable adoption numbers.

Filecoin experienced a decline in network size, suggesting that its cryptocurrency-based incentives no longer sufficiently offset the high operational costs of running nodes. This aligns with our findings in §V about the relationship between cryptocurrency value fluctuations and participation. However, despite having fewer total peers, the remaining Filecoin nodes show increased availability and consistency, indicating a consolidation where committed participants maintain higher reliability. Swarm demonstrates challenges to long-term sustainability with decreases in both peer count and availability. This dual decline suggests the platform is struggling to attract new participants while not improving the reliability of existing nodes. Our data comparison demonstrates how incentivization mechanisms influence long-term network health.

Takeaway: Our network analysis reveals resource requirement and platform incentivization directly impact adoption and sustainability. IPFS achieves the largest network through its free model but suffers from poor peer reliability without economic incentives. Platforms with cryptocurrency rewards (Swarm, Filecoin, Storj) demonstrate significantly higher peer stability but face different challenges: Filecoin's high resource requirements limit participation size, and all platforms show varying degrees of concentration within datacenter environments. Although using datacenters does not inherently undermine decentralization goals, reliance on a few major cloud providers introduces potential points of central control. These trade-offs between network size, reliability, and infrastructure independence highlight fundamental tensions in decentralized storage design.

IV. Performance Analysis

While decentralization offers benefits in terms of resilience and censorship resistance, users ultimately need storage solutions that deliver acceptable performance for their use cases. In this section, we study the performance of the four decentralized storage systems and compare the performance against Google Drive as a representative centralized solution.

A. Experimental Methodology

We measured upload and download capabilities under different file sizes and network environments as discussed below.

1) Data and Environment Setup

- **Test Files**: We selected three representative file sizes: small (5 MB), medium (50 MB), and large (512 MB). These sizes reflect typical user scenarios from simple documents to multimedia content. For statistical validity, we tested three files of each size across all platforms, resulting in 9 uploads per platform. All files are uploaded sequentially with an interval of 1s in each upload.
- **Upload Environment**: All uploads originated from controlled US and Germany nodes with a 1 Gbps network connection. We used each platform's latest client: IPFS (Kubo v.0.28.0), Swarm (Bee v.2.0.0 and swarm-cli v.2.9.0), Filecoin (boost version 2.1.1+mainnet), and Storj (Uplink v1.96.6).
- **Download Environment**: To evaluate geographic performance variations, we conducted download tests from five locations: United States (US), Australia (AU), Japan (JP), Germany (DE), and Chile (CL). We deployed identical client configurations at each location using Vultr [44] cloud computing services.
- **Centralized Baseline**: We included Google Drive as a representative centralized storage service (which has over 2 billion monthly active users [45]).

2) Performance Measurement Approach

Due to architectural differences between platforms, we adapted our measurement methodology to accurately capture comparable metrics:

Upload Time Measurement:

- **IPFS**: Time from upload initiation until the file's Content Identifier (CID) appears in the DHT network and becomes retrievable via provider record.

- **Filecoin**: Time from deal submission to blockchain confirmation. This reflects when the file becomes officially retrievable within the Filecoin network.
- **Storj**: Duration of the upload command execution, as files become immediately accessible after command completion.
- **Swarm**: Time from upload command initiation until file information propagates through the DHT. We configured Swarm to upload directly to the network rather than using its default batching mechanism.
- **Google Drive**: Time from initiation to completion of the upload process measured using the official Google Drive API. We implemented a Python script that authenticates with the API, initiates file uploads, and records timestamps at the beginning and successful completion of each upload operation.

Download Performance Measurement:

- We measured both download time and Time-to-First-Byte (TTFB) for all files across all platforms and locations. For Google Drive, we used the same Google Drive API and recorded the download time and TTFB.
- Download throughput was calculated using total file size divided by download time.

The detailed platform-specific upload and download commands are provided in Appendices §C and §D respectively.

B. Upload Performance

Table IV and Table V present our measured upload times across all platforms compared to Google Drive as a centralized baseline from both the US and Germany.

TABLE IV: *Upload time for each platform with various file sizes from the US. Filecoin does not allow uploads of small-sized files and has a minimum upload size of 512MB.*

File Size	IPFS	Filecoin	Swarm	StorJ	Google Drive
5 MB	3.5 (s)	N/A	11.76 (s)	1.18 (s)	0.94 (s)
50 MB	7.72 (s)	N/A	99.98 (s)	2.20 (s)	1.51 (s)
512 MB	19.54 (s)	24.6 (Hours)	1002.06 (s)	17.88 (s)	7.42 (s)

TABLE V: *Upload time for each platform with various file sizes from Germany.*

File Size	IPFS	Filecoin	Swarm	StorJ	Google Drive
5 MB	12.56 (s)	N/A	8.98 (s)	1.44 (s)	1.23 (s)
50 MB	13.30 (s)	N/A	86.17 (s)	2.65 (s)	2.02 (s)
512 MB	23.50 (s)	10.68 (Hours)	879.79 (s)	19.97 (s)	13.42 (s)

For small files (5MB), the fastest upload is using Google Drive in both US and Germany (0.94 and 1.23 seconds average), which is not surprising as centralized services are more efficient in handling direct uploads. In contrast, decentralized platforms show varied results. Storj has the closest performance to centralized storage, with an average upload time of 1.18 and 1.44 seconds for 5 MB file. Recall that Storj uses a hybrid data routing, as shown in Figure 3. When uploading a file, Storj contacts the satellites first to retrieve a list of storage peers for storing the file. This hard requirement of getting all routing information by satellites makes the architecture essentially centralized where all information comes from one source. However, this architecture also improves upload performance since the file information does not have to be propagated across the decentralized network.

IPFS and Swarm show significantly longer upload times even for small files from both locations. While both use similar Kademlia DHT-based routing (Table II), IPFS outperforms Swarm by approximately 70% for small files (3.5s vs. 11.76s) from the US upload. This substantial difference stems from their divergent approaches to data persistence. As shown in Figure 1, Swarm prioritizes guaranteed data persistence through duplication and erasure coding, while IPFS offers no persistence guarantees. This design choice manifests in Swarm's smaller chunk size (4KB compared to IPFS's 256KB), creating more overhead as more chunks need to be processed and uploaded. We also notice IPFS performs significantly poorer from Germany, which is due to fewer nodes being available in the network as we observed in §III-C.

As file sizes increase, the performance gap between centralized and decentralized solutions widens. Storj's upload time for 512MB files (17.88s and 19.97s) is more than double Google Drive's (7.42s and 13.42s), reflecting the additional processing required for erasure coding as file sizes increase. Similarly, IPFS and Swarm show increased upload times for larger files, with Swarm's performance degrading most dramatically due to its small chunk size. However, we notice that Swarm performs better when uploading from Germany due to its large number of nodes present in Europe, as we observed in §III-B.

Filecoin has the largest file upload time among the four platforms. For 512MB file size, it takes more than 24 hours from the US and over 10 hours from Germany to upload the file. This extreme delay directly results from Filecoin's position on the design spectrum (Figure 1), where it combines high resource requirements with cryptocurrency-based incentivization. Filecoin's protocol requires storage providers to "seal" a sector before pushing the transaction to the blockchain. To optimize operations, providers typically aggregate multiple deals into a single sector, maximizing space utilization before committing transactions. During our upload experiment, we observed that our data transferred to the storage provider within 8 minutes on average. However, the file is not available until the transaction is committed to the blockchain, which accounts for the extended upload time. The shorter 10-hour upload time from Germany likely resulted from fortunate timing, where our data was bundled together with other clients' deals, triggering earlier sector sealing and blockchain commitment.

C. Retrieval Performance

Retrieval performance is important for user-facing applications. Similar to upload performance, the data routing and persistence strategies chosen by each platform significantly impact retrieval capabilities. In fact, download performance suffers even more acutely from decentralization design choices, as retrieval requires not just locating content across distributed nodes but also efficiently assembling fragmented data into usable files. We evaluate retrieval performance across three key metrics: download throughput, download time, and time-to-first-byte (TTFB).

Figure 10 and Figure 11, along with Table VI, reveal complex performance dynamics across platforms. While Google Drive predictably provides the highest overall throughput, it doesn't consistently deliver the fastest experience, particularly for small files. This counterintuitive result stems from Google Drive's surprisingly high TTFB (Figure 11), especially for clients distant from its US-based infrastructure. Google Drive's centralized architecture creates initial latency that becomes less significant only as file size increases and its superior throughput compensates for the slow start.

TABLE VI: *Average download throughput (Mbps) for different storage platforms across different geo-locations.*

Platform/Geo-location	US	AU	JP	DE	CL
StorJ	147.55	60.47	60.47	185.57	70.81
Swarm	3.03	1.7	1.84	3.59	2.35
Filecoin	13.49	13.51	19.05	8.81	25.11
IPFS	60.74	21.76	24.02	52.86	38.18
Google Drive	247.15	67.22	82.94	130.99	197.96

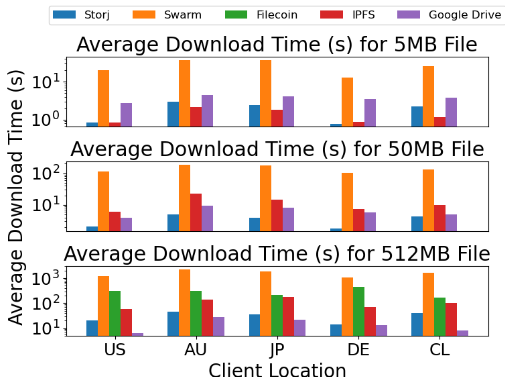

Fig. 10: *Average download time for 5, 50, 512 MB for all storage platforms across all geo-locations. Filecoin only allows storing and retrieving files of size 512MB or greater.*

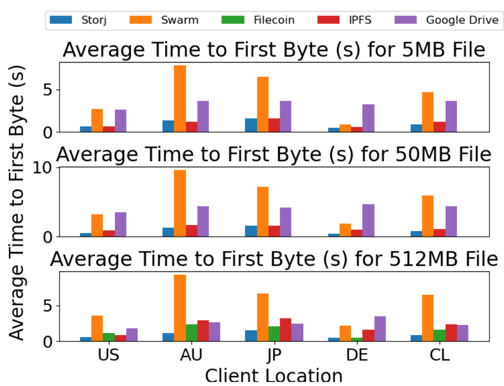

Fig. 11: *Average Time to First Byte (TTFB) for different storage platforms across different geo-locations*

Among decentralized platforms, Storj demonstrates the most consistent performance across all metrics due to its satellite-based architecture (Table II). Its quasi-centralized approach to routing delivers both efficient throughput (Table VI) and consistently low TTFB (Figures 11). When retrieving files, clients simply query the satellites for file locations rather than searching a distributed network. Additionally, Storj's erasure coding implementation allows clients to reconstruct files even before all chunks are downloaded, creating a double advantage in both responsiveness and overall download speed. However, this advantage diminishes with larger files as the computational overhead of processing numerous chunks increases, as evident in Figures 10. Storj's performance also exhibits geographical sensitivity, with download speeds in Asia and South America reaching only half those observed in Europe and North America, reflecting the geographical distribution of its nodes shown in Figure 9.

IPFS and Filecoin present an interesting comparative case. Despite sharing technological foundations, they deliver different retrieval experiences. IPFS consistently achieves approximately 300% higher throughput than Filecoin across all locations. For TTFB, IPFS demonstrates variable responsiveness dependent on file size. Its Merkle DAG (Directed Acyclic Graph) structure allows quick responses for small files but creates increasing latency for larger ones as the tree structure becomes more complex to traverse. A Merkle DAG is a tree-like data structure where each node contains cryptographic hashes of its children, enabling content verification and deduplication while maintaining hierarchical relationships between data blocks. Filecoin often delivers faster initial responses than IPFS due to its well-provisioned storage nodes, but this TTFB advantage is quickly overwhelmed by significantly slower overall download times. The performance gap between these platforms comes from Filecoin's blockchain-based guarantees, which require complex validation processes.

Swarm demonstrates the lowest performance across all metrics and locations, a direct consequence of its design choices favoring true decentralization (Figure 1). The poor performance is shown in both throughput (Table VI) and download times (Figures 10). The swarm's small 4KB chunk size creates numerous fragments distributed across different peers. Unlike IPFS, which stores related chunks on the same provider, Swarm distributes each chunk to different providers based on proximity to the chunk's address. This fragmentation creates a double penalty: high TTFB as clients first download metadata files containing chunk hashes, and then slow overall downloads as each chunk must be individually retrieved from different sources.

Takeaway: Data routing and persistence designs significantly impact both upload and download performance. A clear trade-off exists between decentralization and performance: platforms with more centralized components achieve higher throughput, while those prioritizing decentralization face notable performance penalties. This fundamental tension suggests that current decentralized storage systems cannot fully optimize for both goals at once. However, emerging hybrid approaches such as Filecoin Saturn [46], which incentivize edge networks, offer promising ways to improve performance without sacrificing decentralization.

V. Incentives and Pricing

One of the key reasons for the growth in decentralized storage solutions is the presence of financial incentives for user participation in their storage network. This is a shift from traditional centralized storage systems that are still within the domain of large, well-established, companies. In this section, we examine how these incentives influence network participation, sustainability, and adoption from both provider and client perspectives.

A. Methodology

To analyze incentives across the platforms, we collected data from multiple sources to understand both the provider-side economics and client-side costs:

1) **Cryptocurrency price data**: We obtained historical and current price data for Filecoin (FIL), Storj (STORJ), and Swarm (BZZ) tokens from CoinGecko [47].
2) **Network participation metrics**: For provider participation, we collected total storage capacity statistics from official platform sources including Filecoin's Starboard dashboard [20]

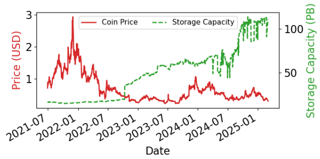

Fig. 12: *Storj's cryptocurrency (STORJ) price and storage capacity over time.*

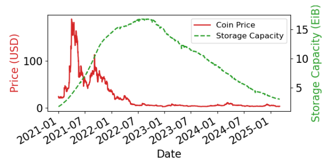

Fig. 13: *Filecoin's cryptocurrency (FIL) price and storage capacity over time.*

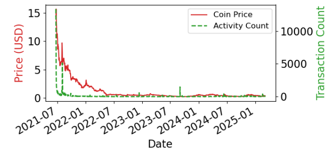

Fig. 14: *Swarm's cryptocurrency (BZZ) price and activity transaction count over time.*

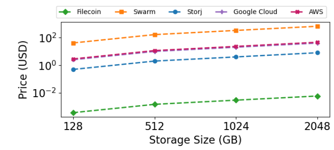

Fig. 15: *Monthly cost for hosting different capacity on Filecoin, Swarm, Storj, GCP, and AWS.*

and Storj's network statistics portal [22]. For Swarm, which doesn't publish storage capacity metrics, we used blockchain transaction volume as a proxy for network participation.

3) **Token distribution data**: For Ethereum-based tokens (Swarm and Storj), we used Etherscan's API [48] to extract transaction data. We analyzed token holder information by replaying the complete transaction history on each platform's blockchain and calculating final token balances for each address. We cross-validated our calculated distributions against the published token distribution statistics provided by Etherscan. For Filecoin, we utilized Filscan [49] directly, as the full blockchain is significantly large to replay locally.

We omitted IPFS because it does not use financial incentives.

B. Provider Incentives

1) Cryptocurrency Price vs. Storage Commitment

Figure 12 illustrates Storj's inverse relationship between token price and network capacity. Despite STORJ token value declining from approximately \$3 to \$0.5 between 2022 and 2024, the network's storage capacity grew substantially from 20 PB to over 100 PB. This counterintuitive pattern suggests that as token value decreases, storage providers must contribute more capacity to maintain similar earnings, effectively driving network growth even during market downturns.

In stark contrast, Filecoin (Figure 13) demonstrates a direct correlation between token price and network capacity. As FIL price dropped precipitously from over \$200 to approximately \$6, the network's committed storage capacity also decreased significantly. This correlation drives from Filecoin's high hardware and resource requirements for participation. The substantial upfront investment needed to become a Filecoin storage provider means that when token values fall, the economic equation becomes unfavorable, and participants exit the market.

For Swarm (Figure 14), the data reveals an initial surge of activity during the platform's launch phase, followed by declining engagement as token prices fell. This pattern indicates that Swarm faces sustainability challenges when token prices decrease, limiting its ability to maintain network participation. We also observed this trend during our network size analysis in §III-B.

The relationship between cryptocurrency price and provider participation is not straightforward and depends on several factors. While we observed direct correlations between token value and network engagement for Filecoin and Swarm, Storj demonstrates that other factors (such as resource requirements) could also impact provider behavior regardless of the pricing.

2) Token Distribution and Ownership Patterns

Since anyone can purchase storage tokens as investments, we examine whether tokens primarily circulate among actual storage participants, or are concentrated in the hands of external investors. Note here that storage participants are those who provide decentralized storage service in exchange for cryptocurrency. Figure 16 shows the top five holders for each platform from the token distribution data (full details shown in Table VIII at Appendix).

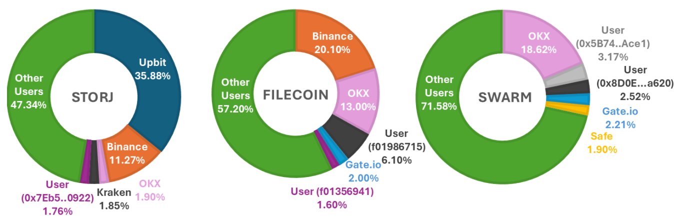

Fig. 16: *Token distribution showing the top 5 holders and other users for Storj (STORJ), Filecoin (FIL), and Swarm (BZZ).*

Storj exhibits the most concentrated ownership structure, with exchanges controlling approximately 50.9% of all tokens (including 35.9% held by Upbit alone). This concentration among financial entities rather than actual storage providers signals a disconnect between token economics and network operation. Interestingly, this exchange dominance may partially explain the network growth phenomenon we observed in the previous section, where provider participation increased despite declining token prices. This pattern suggests participants have more confidence in the long-term stability of rewards, possibly due to the presence of these major exchanges.

Filecoin maintains a more moderate distribution pattern, with 42.8% of tokens held by the top five addresses (predominantly exchanges), while 57.2% circulates among a broader user base. With over 3.4 million accounts compared to Storj's 99,707 and Swarm's 23,787, Filecoin demonstrates wider market participation due to its higher peak valuation attracting broader interest.

Swarm shows the most distributed ownership model among the three platforms, with 71.6% of tokens held outside the top five addresses. This distribution aligns with Swarm's philosophical commitment to decentralization in all aspects of design decisions. The distributed ownership of the token demonstrated closer alignment between token holders and actual network participants.

All platforms experience some level of exchange dominance in their token distribution, with varying degrees of concentration. This dominance creates uncertainty about long-term platform sustainability as market trading directly influences incentive outcomes rather than participant involvement.

C. Client Economics: Storage Pricing and Adoption Incentives

Next, we look at the pricing for each storage system. Figure 15 shows the monthly cost of hosting files for different sizes. We include two centralized storage systems—Google Cloud and AWS for comparison. Pricing data for Storj, Swarm, and centralized storage providers was collected from their publicly listed rates, while Filecoin prices were calculated by converting provider-requested FIL amounts to USD at current exchange rates.

Filecoin offers the most economical storage solution at less than $1 per TB per month—dramatically undercutting centralized providers like Google Cloud and AWS that charge approximately $30 for equivalent capacity. This aggressive pricing stems from Filecoin's massive committed storage capacity (measured in exbibytes) combined with currently low token values, creating a buyer's market for storage clients.

Storj presents a middle ground at approximately $4 per TB per month, still significantly more affordable than centralized alternatives while maintaining reasonable performance characteristics as demonstrated in our earlier analysis. This balance between cost and performance positions Storj as a practical alternative for cost-conscious users seeking reliable decentralized storage.

In contrast, Swarm's storage costs exceed both decentralized and centralized options at $335.54 per TB per month. This prohibitive pricing directly correlates with the platform's declining participation and transaction activity observed in our cryptocurrency analysis. The high costs create a negative feedback loop: lower participation leads to higher prices, which further discourages adoption.

These pricing differences directly impact adoption potential. Filecoin's extraordinarily low costs make it attractive for large-scale storage needs despite its performance limitations, while Storj's balance of affordability and performance appeals to a broader use case range. Swarm's current pricing structure presents a significant barrier to adoption regardless of its technical merits.

Takeaway: Economic incentive models directly influence the adoption and sustainability of decentralized storage platforms. They shape both provider participation and client onboarding, creating a difficult balance to maintain. Filecoin offers attractive pricing for clients but struggles with provider retention during cryptocurrency downturns due to high hardware demands. Storj sustains growth with lower entry barriers, but compromises on decentralization. Swarm upholds decentralization ideals but suffers from high storage costs that hinder adoption. These trade-offs reflect the challenge of simultaneously attracting storage providers and clients; optimizing for one often undermines the other. Interestingly, while Filecoin shows a direct relationship between token value and network capacity, Storj exhibits an inverse trend. This contrast highlights how incentive mechanisms and system architecture interact in complex ways that can produce counterintuitive economic outcomes.

VI. Related Work

Previous studies on decentralized storage platforms have typically focused on individual platforms and specific aspects of the platforms. IPFS is the most researched platform with studies on network size [7], [50], churn rate [34], transportation protocol [51], I/O performance [52], and security aspects [53]. Trautwein et al. [7] perform a large-scale study on the number of users, global reach, and performance of IPFS. This study also reports that the IPFS network is highly decentralized in terms of individual hosting. A follow-up study by Balduf et al. [33] use network and traffic analysis to show that many IPFS nodes are hosted in data centers. Other works [54] also show improvement in IPFS content discovery when using a centralized infrastructure. Recent study also investigate content moderation in IPFS [55].

Filecoin and Storj have also been analyzed in related work, but not to the extent of IPFS. Studies have examined Filecoin's blockchain storage mechanism [56] and scalability issues [57]. Studies have also shown that Filecoin's reward system can cause centralization on the blockchain [13]. However, these studies rarely connect technical choices to economic incentives. In terms of Storj, studies have evaluated Storj's performance with AWS [14], studied potential security issues with Storj satellites [58], [59], and have forecast Storj's token prices [60].

Few studies provide a comparison of different decentralized storage platforms. A survey by Daniel et al. [61] compares different decentralized storage platforms, but only in terms of system design. Other studies have compared the overall cost of different decentralized storage platforms [62]. The study by Huang et al [63] conducts an overview of the rationale and layered structure of IPFS and Swarm. However, there are no comprehensive comparisons analyzing how design decisions such as routing architecture, data persistence, and incentive structures directly affect adoption, performance, and long-term sustainability of decentralized storage solutions.

VII. Concluding Remarks

Our analysis demonstrates that decentralized storage platforms embody fundamental trade-offs across network characteristics, performance, and incentive mechanisms. No single platform excels in all dimensions. IPFS offers the largest network but suffers from poor reliability. Storj delivers performance approaching centralized solutions but compromises on true decentralization. Filecoin provides affordable client pricing but struggles with provider retention during market downturns. Swarm maintains stronger decentralization principles but faces significant performance and cost penalties.

While we did not explore in detail in this study, data privacy represents another critical dimension worth acknowledging. Despite their censorship resistance, these systems handle data privacy differently: IPFS's open data model makes content publicly accessible by default, Storj and Swarm implement client-side encryption, and Filecoin has the same open data model as IPFS by default, but it allows private deals. This data privacy-decentralization balance constitutes another important trade-off in this ecosystem.

The architectural trade-offs among routing, persistence, incentives, resources, and privacy directly influence each platform's capabilities. Our systematic comparison provides practical guidance for users selecting platforms based on specific needs while offering valuable design insights for researchers. Future systems must innovate in creating mechanisms that ensure provider sustainability, client affordability, true decentralization, and robust privacy guarantees. This remains a challenging but essential goal for creating viable alternatives to centralized storage infrastructure.

VIII. Acknowledgements

We thank all the reviewers and our shepherd for their constructive feedback. This work was supported in part by NSF grants CNS-1909356 and CNS-1941617.

References

[1] A. Ratiu, "Enhanced trust and confidence in the digital economy," May 2021. [Online]. Available: https://www.atlanticcouncil.org/content-series/geotech-commission/chapter-3/

[2] "Privacy and data protection: Increasingly precious asset in digital era says un expert," Oct 2022. [Online]. Available: https://www.ohchr.org/en/press-releases/2022/10/privacy-and-data-protection-increasingly-precious-asset-digital-era-says-un

[3] Sandvine, "2023 global internet phenomena report," *Sandvine*, 2023. [Online]. Available: https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2023/reports/Sandvine%20GIPR%202023.pdf

[4] M. Hicks, "2023 internet outage trends and the new outage landscape," 2024. [Online]. Available: https://www.thousandeyes.com/blog/internet-report-pulse-update-2023-internet-outage-trends#Cloud-Service-Provider-Outages-Trending-Up

[5] "Turkey extends the censorship of youtube," 2010. [Online]. Available: https://edri.org/our-work/edrigramnumber8-12turkey-extends-blocking-youtube/

[6] S. Alder, "Itrc: Data compromises reach all-time high in 2023," 2023. [Online]. Available: https://www.hipaajournal.com/itrc-data-compromises-record-2023/

[7] D. Trautwein, A. Raman, G. Tyson, I. Castro, W. Scott, M. Schubotz, B. Gipp, and Y. Psaras, "Design and evaluation of ipfs: a storage layer for the decentralized web," in *Proceedings of the ACM SIGCOMM 2022 Conference*, ser. SIGCOMM '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 739–752. [Online]. Available: https://doi.org/10.1145/3544216.3544232

[8] D. Trautwein, "Network-measurements/results/rfm21-hydras-performance-contribution.md at master · protocol/network-measurements," Jan 2023. [Online]. Available: https://github.com/protocol/network-measurements/blob/master/results/rfm21-hydras-performance-contribution.md

[9] Filecoin, "A decentralized storage network for the world's information," 2024. [Online]. Available: https://filecoin.io/

[10] "Featured clients," 2024. [Online]. Available: https://filecoin-explorer.com/expanded-view

[11] W. Hanamura, "What information should we be preserving in filecoin? — internet archive blogs," Oct. 2020. [Online]. Available: https://blog.archive.org/2020/10/22/what-information-should-we-be-preserving-in-filecoin/

[12] "The seti institute works with filecoin to simulate extraterrestrial contact." 2023. [Online]. Available: https://destor.com/seti

[13] B. Guidi, A. Michienzi, and L. Ricci, "Evaluating the decentralisation of filecoin," in *Proceedings of the 3rd International Workshop on Distributed Infrastructure for the Common Good*, ser. DICG '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 13–18. [Online]. Available: https://doi.org/10.1145/3565383.3566108

[14] H. Li, X. Mi, Y. Dou, and S. Guo, "An empirical study of storj dcs: Ecosystem, performance, and security," in *2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS)*, 2023, pp. 1–10.

[15] IPFS, 2024. [Online]. Available: https://ipfs.tech

[16] S. DCS, 2024. [Online]. Available: https://www.storj.io/

[17] EthSwarm, 2024. [Online]. Available: https://www.ethswarm.org/

[18] P. Lab, 2024. [Online]. Available: https://probelab.io/ipfs/amino/

[19] filecoin, 2024. [Online]. Available: https://storage.filecoin.io/

[20] starboard, 2024. [Online]. Available: https://dashboard.starboard.ventures/dashboard

[21] S. DCS, 2024. [Online]. Available: https://stats.storjshare.io/

[22] storjstats, 2024. [Online]. Available: https://storjstats.info/d/storj/storj-network-statistics?orgId=1

[23] siascan, 2024. [Online]. Available: https://siascan.com/

[24] viewblock, 2024. [Online]. Available: https://viewblock.io/arweave

[25] [Online]. Available: https://protocol.ai/

[26] J. Benet, "Ipfs-content addressed, versioned, p2p file system," *arXiv preprint arXiv:1407.3561*, 2014.

[27] Messari, "State of filecoin q3 2024," 2024, accessed: 2025-05-11. [Online]. Available: https://messari.io/report/state-of-filecoin-q3-2024

[28] F. Foundation, "Introducing proof of data possession (pdp): Verifiable hot storage on filecoin," 2024, accessed: 2025-05-11. [Online]. Available: https://fil.org/blog/introducing-proof-of-data-possession-pdp-verifiable-hot-storage-on-filecoin

[29] S. Foundation, "The origins of swarm," 2024, accessed: 2025-05-11. [Online]. Available: https://blog.ethswarm.org/hive/2024/the-origins-of-swarm/

[30] S. Labs, "Why storj is the smarter cloud for data storage," 2024, accessed: 2025-05-11. [Online]. Available: https://www.storj.io/why-storj

[31] D. Trautwein, A. Raman, G. Tyson, I. Castro, W. Scott, M. Schubotz, B. Gipp, and Y. Psaras, "Design and evaluation of ipfs: A storage layer for the decentralized web," Amsterdam, NL, Aug. 2022. [Online]. Available: https://github.com/dennis-tra/nebula

[32] J. Guljaš, 2024. [Online]. Available: https://swarmscan.io/

[33] L. Balduf, M. Korczyński, O. Ascigil, N. V. Keizer, G. Pavlou, B. Scheuermann, and M. Król, "The cloud strikes back: Investigating the decentralization of ipfs," in *Proceedings of the 2023 ACM on Internet Measurement Conference*, 2023, pp. 391–405.

[34] E. Daniel and F. Tschorsch, "Passively measuring ipfs churn and network size," in *2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2022, pp. 60–65.

[35] "Lily's dump," 2023. [Online]. Available: https://lilium.sh/data/dumps/

[36] "Software components — filecoin docs," Mar. 2024. [Online]. Available: https://docs.filecoin.io/storage-providers/architecture/lotus-components

[37] "User agents analysis," 2024. [Online]. Available: https://udger.com/

[38] "The trusted source for ip address data, leading ip data provider," 2024. [Online]. Available: https://ipinfo.io

[39] "ipapi - ip address lookup and geolocation api," 2024. [Online]. Available: https://ipapi.com/

[40] Constant.com, "Constant cloud infrastructure services," 2024, accessed: 2025-05-13. [Online]. Available: https://constant.com

[41] C. GmbH, "Contabo - vps, dedicated servers and cloud hosting," 2024, accessed: 2025-05-13. [Online]. Available: https://contabo.com

[42] A. W. Services, "Amazon web services (aws)," 2024, accessed: 2025-05-13. [Online]. Available: https://aws.amazon.com

[43] H. Inc., "Hostpapa web hosting services," 2024, accessed: 2025-05-13. [Online]. Available: https://www.hostpapa.com

[44] [Online]. Available: https://www.vultr.com/

[45] F. Duarte, Sep. 2023. [Online]. Available: https://explodingtopics.com/blog/google-workspace-stats

[46] [Online]. Available: https://saturn.tech/

[47] CoinGecko, "Cryptocurrency prices, charts, and crypto market cap — coingecko," 2024. [Online]. Available: https://www.coingecko.com/

[48] etherscan.io, "Ethereum (eth) blockchain explorer," 2024. [Online]. Available: https://etherscan.io/

[49] [Online]. Available: https://filscan.io/en/

[50] S. Henningsen, M. Florian, S. Rust, and B. Scheuermann, "Mapping the interplanetary filesystem," in *2020 IFIP Networking Conference (Networking)*, 2020, pp. 289–297.

[51] A. De la Rocha, D. Dias, and Y. Psaras, "Accelerating content routing with bitswap: A multi-path file transfer protocol in ipfs and filecoin," 2021.

[52] J. Shen, Y. Li, Y. Zhou, and X. Wang, "Understanding i/o performance of ipfs storage: a client's perspective," in *Proceedings of the International Symposium on Quality of Service*, 2019, pp. 1–10.

[53] B. Prünster, A. Marsalek, and T. Zefferer, "Total eclipse of the heart–disrupting the {InterPlanetary} file system," in *31st USENIX Security Symposium (USENIX Security 22)*, 2022, pp. 3735–3752.

[54] Y. Wei, D. Trautwein, Y. Psaras, I. Castro, W. Scott, A. Raman, and G. Tyson, "The eternal tussle: Exploring the role of centralization in IPFS," in *21st USENIX Symposium on Networked Systems Design and Implementation (NSDI 24)*. Santa Clara, CA: USENIX Association, Apr. 2024, pp. 441–454. [Online]. Available: https://www.usenix.org/conference/nsdi24/presentation/wei

[55] S. Sokoto, L. Balduf, D. Trautwein, Y. Wei, G. Tyson, I. Castro, O. Ascigil, G. Pavlou, M. Korczynski, B. Scheuermann, and M. Król, "Guardians of the galaxy: Content moderation in the InterPlanetary file system," in *33rd USENIX Security Symposium (USENIX Security 24)*. Philadelphia, PA: USENIX Association, Aug. 2024, pp. 1507–1524. [Online]. Available: https://www.usenix.org/conference/usenixsecurity24/presentation/sokoto

[56] I. Vakilinia, W. Wang, and J. Xin, "An incentive-compatible mechanism for decentralized storage network," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 4, pp. 2294–2306, 2023.

[57] B. Fisch, J. Bonneau, N. Greco, and J. Benet, "Scaling proof-of-replication for filecoin mining," *Report. Protocol Labs Research*, 2018.

[58] S. de Figueiredo, A. Madhusudan, V. Reniers, S. Nikova, and B. Preneel, "Exploring the storj network: a security analysis," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, ser. SAC '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 257–264. [Online]. Available: https://doi.org/10.1145/3412841.3441908

[59] X. Zhang, J. Grannis, I. Baggili, and N. L. Beebe, "Frameup: An incriminatory attack on storj: A peer to peer blockchain enabled distributed storage system," *Digital Investigation*, vol. 29, pp. 28–42, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1742287618303438

[60] R. Bush and S. Choi, "Forecasting ethereum storj token prices: Comparative analyses of applied bitcoin models," in *2019 International Conference on Data Mining Workshops (ICDMW)*, 2019, pp. 216–223.

[61] E. Daniel and F. Tschorsch, "Ipfs and friends: A qualitative comparison of next generation peer-to-peer data networks," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 31–52, 2022.

[62] A. Ismail, M. Toohey, Y. C. Lee, Z. Dong, and A. Y. Zomaya, "Cost and performance analysis on decentralized file systems for blockchain-based applications: State-of-the-art report," in *2022 IEEE International Conference on Blockchain (Blockchain)*, 2022, pp. 230–237.

[63] H. Huang, J. Lin, B. Zheng, Z. Zheng, and J. Bian, "When blockchain meets distributed file systems: An overview, challenges, and open issues," *IEEE Access*, vol. 8, pp. 50 574–50 586, 2020.

[64] Sia, 2024. [Online]. Available: https://sia.tech/

[65] Arweave, 2024. [Online]. Available: https://www.arweave.org/

[66] I. Drive, 2024. [Online]. Available: https://internxt.com

[67] Holepunch, 2024. [Online]. Available: https://holepunch.to/

[68] S. Network, 2024. [Online]. Available: https://safenetwork.tech/

[69] [Online]. Available: https://holochain.org/

[70] Zus, 2024. [Online]. Available: https://zus.network/

[71] O. Storage, 2022. [Online]. Available: https://opacity.io

[72] B. Lee, "Introducing lassie - a retrieval client for filecoin and ipfs," 2023. [Online]. Available: https://filecoin.io/blog/posts/introducing-lassie---a-retrieval-client-for-filecoin-and-ipfs/

Appendix

A. Ethics

This work does not raise any ethical issues.

B. Existing Decentralized Storage Platforms

TABLE VII: *Existing decentralized storage platforms available in the market (information obtained from the official website for each platform)*

Platforms	In Production	Network Information Availability	Storage Size Availability	Last Updated
IPFS [15]	✓	✓	N/A	2025-02-14
Filecoin [9]	✓	✓	✓	2025-02-11
Swarm [17]	✓	✓	✗	2025-03-11
Sia [64]	✓	✓	✓	2025-02-25
StroJ [16]	✓	✓	✓	2025-02-18
Arweave [65]	✓	✓	✓	2025-02-11
Internext [66]	✓	✗	✗	2025-03-04
Hypercore [67]	✓	✗	✗	2025-03-11
SAFE [68]	✓	✗	✗	2025-02-25
Holochain [69]	✓	✗	✗	2025-01-29
0Chain [70]	✗	✗	✗	2024-11-13
Opacity [71]	✗	✗	✗	2022-07-25

C. Upload Commands

- **IPFS:** To measure the upload time for IPFS, we calculate the duration when the CID of the file appeared in the DHT network to the time of initial upload. To find the provider record for any given CID we used the IPFS command *ipfs routing findprovs ⟨CID⟩*
- **Filecoin:** We measure the time when the Filecoin deal status being *Proving* and calculate the duration when we submitted the *deal*. To keep track of our deal status we used Filecoin command *boost deal-status –provider=⟨PID⟩ –deal-uuid=⟨UID⟩*.
- **Storj:** We measure the execution time using the command *uplink cp ⟨local file path⟩ ⟨storj storage path⟩*
- **Swarm:** We measure the uploadtime as the execution time of Swarm's upload API */bzz* with options *swarm-deferred-upload: false* in the POST header to the API call.

D. Download and Time-to-First-Byte Commands

- **IPFS:** To measure the download time, we measured the execution time of the IPFS download command *ipfs get ⟨CID⟩*. For TTFB, we use *CURL* to access IPFS gateway link *http://localhost:8080/ipfs/⟨CID⟩*
- **Filecoin:** We used a retrieval tool designed for Filecoin called *lassie* [72]. We measure the execution time for lassie's download command *lassie fetch ⟨CID⟩*. For TTFB, we use *CURL* to fetch Lassie's HTTP API *http://localhost:port/ipfs/⟨CID⟩*
- **Swarm:** We use *CURL* to measure both download time and TTFB. The download time is the total duration of *CURL* fetch through Swarm's API *http://localhost:1633/bzz/⟨swarm hash⟩*.
- **Storj:** we measure the execution time for the download command *uplink cp ⟨storj file path⟩ ⟨save path⟩*. For TTFB, we first share the file through *uplink share* and use *CURL* to access the shared file via URL *https://link.storjshare.io/raw/key/filepath*.

E. Coin Distribution Data

Table VIII show the detail coin distribution by platform.

Storj (STORJ) Total Supply: 424,999,998		Filecoin (FIL) Total Supply: 643,115,202		Swarm (BZZ) Total Supply: 63,149,437	
Address	Asset Amount	Address	Asset Amount	Address	Asset Amount
0xCB1C98A7Fbf4fDE5f27dB695434932A580b8D5b3	152500813	f086971	129415544	0x6cC5F688a315f3dC28A7781717a9A798a59fDA7b	11757776
0xF977814e90dA44bFA03b6295A0616a897441aceC	47911761	f01986715	83726721	0x5B7402902dc1f3a6Ab01c69F1bad99a8372bAce1	1999772
0x94dBF04E273d87e6D9Bed68c616F43Bf86560C74	8054261	f03385436	39213227	0x8D0EA62f0365F8D206EFe701e94581659Ee3a620	1593292
0x6cC5F688a315f3dC28A7781717a9A798a59fDA7b	7854044	f0224439	12810482	0x0D0707963952f2fBA59dD06f2b425ace40b492Fe	1395211
0x7Eb52D0c25717e6fC6704d8EC2CBe54Ae7750922	7471446	f01356941	10063684	0xA6871939e46E654A3d2642085916276921898c67	1200000

TABLE VIII: *Top 5 holder address and asset amount for Storj, Filecoin, and Swarm*