Analysis on Blockchain Project, ICO and Tokenomics

Decentralized Storage - STORJ

1. Introduction

The utilization of blockchain technology by decentralized storage platforms such as Storj can potentially transform data outsourcing. By connecting cryptocurrency with local storage resources, these platforms ensure data integrity across distributed networks, leading to a revolutionized data storage approach **[1]**.

2. Evaluation of the business goal of this project

To meet the increasing need for data storage (Fig-2.1), STORJ's business goal provides an affordable, eco-friendly and decentralized storage system by utilizing unused space.

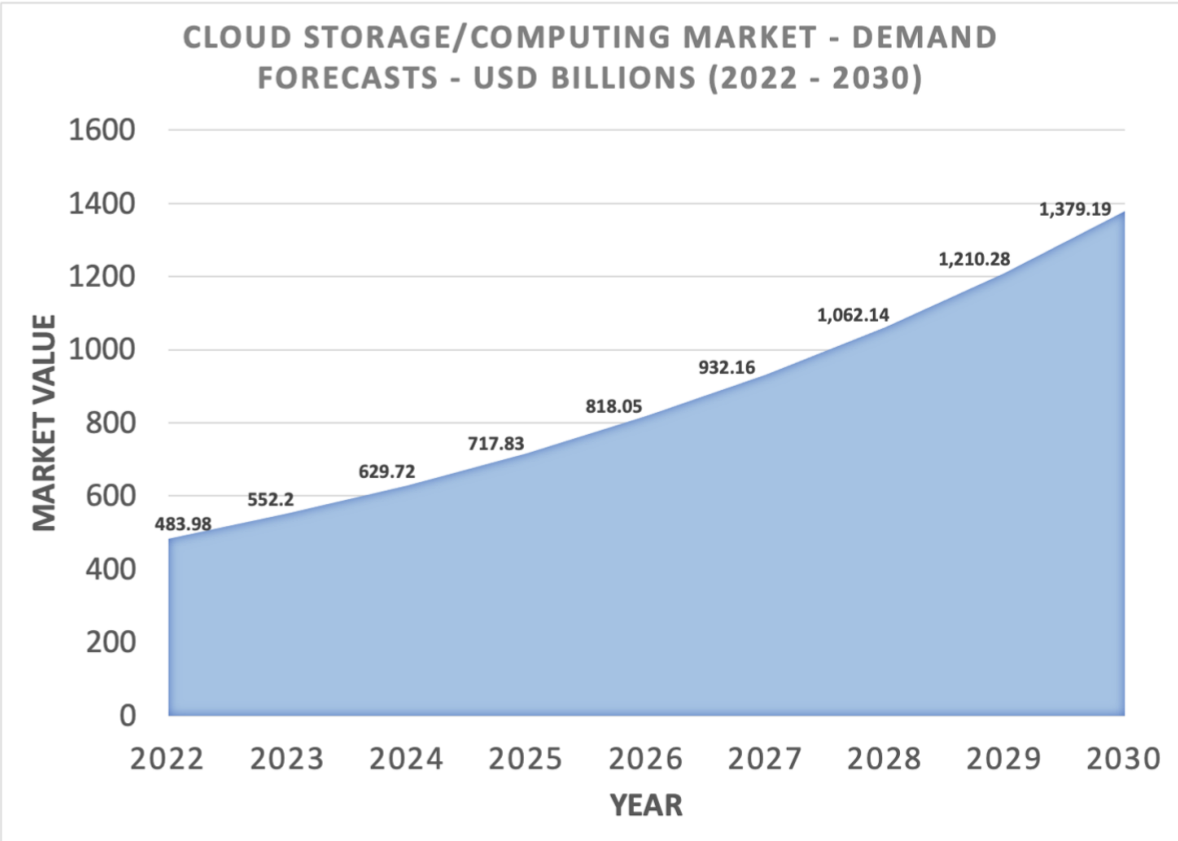


Fig 2.1 : Demand Forecast of Cloud Computing Market 2022 – 2030

Source : Created by Author

Data-Source : Grand-View Research

The project has a well-defined roadmap with specific objectives and goals, many of which have already been accomplished, such as the deployment of its V3 network and the new Storj DCS (Decentralized Cloud Storage) Interface.

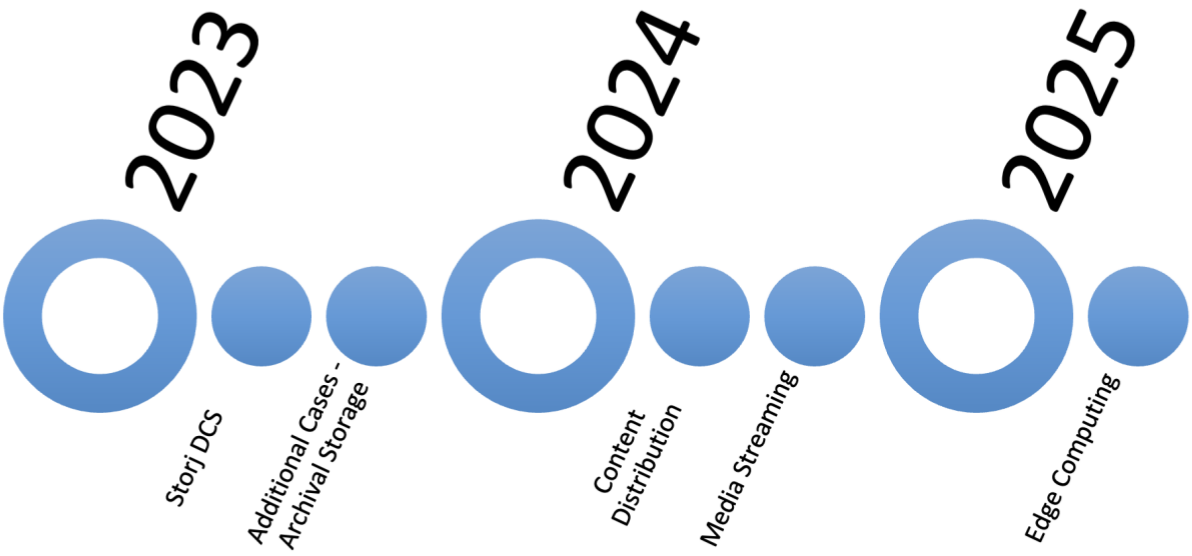


Fig 2.2 : STORJ Roadmap

Source : Created by Author

Data-Source : https://github.com/storj/roadmap

Sustainable, Censorship resistant and Durable Alternative to Cloud Storage: STORJ has a diverse range of applications, including Cloud Computing, Software Delivery, and Video Content Management. Users can pay to upload their files beyond a free tier of 150 GB and Storj doesn't use proof of work (Up-to 30 transactions/second) or proof of stake (Up-to 100,000 transactions/second) to reward its miners, as all the STORJ tokens have already been mined. Node owners are paid around 60% of revenue generated by the network based on the amount of data stored and bandwidth used per month, while the company retains the remaining 40% **[2]**. STORJ's S3 compatibility, a rare feature among platforms and a key sales point allows for easy migration to STORJ using AWS-S3 APIs.

According to estimates, data centres could contribute more than 3 percent to carbon discharge in the world by 2025. By 2040, this could rise to 14%, which is equivalent to the carbon output of USA. A recent survey of STORJ node operators found that over 69% of the capacity (Storj DCS) is located on unused/underutilized devices. Storj splits encrypted files into at least 80 pieces where each piece is stored on a unique drive at different area with distinct ISPs, and only 29 pieces are required to reconstruct the file. Carbon discharge of a platform using replication (EF of 9), reaches the same level of durability of STORJ (with EF of [80/29]= 2.7) only after storing 9-10 copies of a single file, which is more resource-intensive.

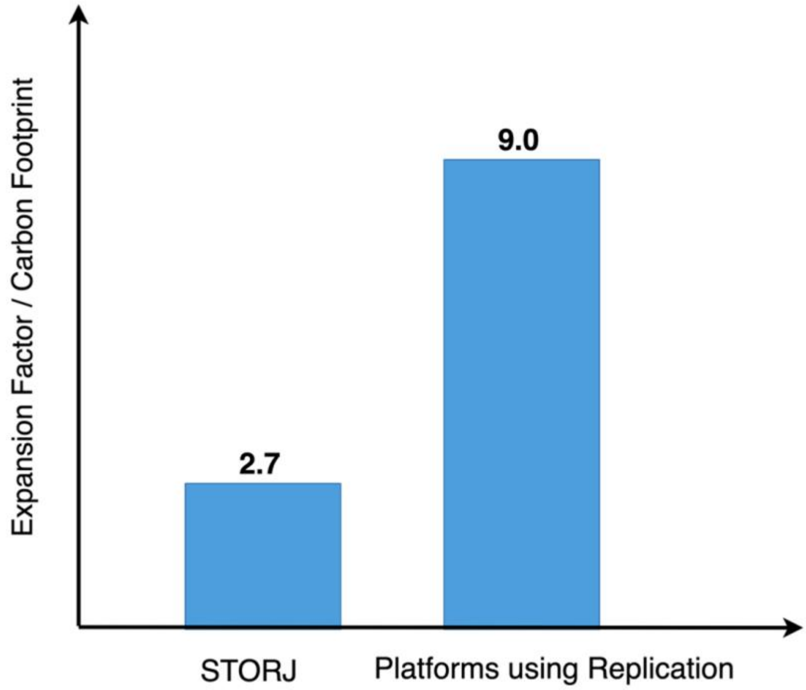


Fig 2.3 : *Expansion Factor - Cost - Carbon Foot-print* Comparison

Source : Created by Author

Data-Source : STORJ

Ethereum facilitates the management of the ERC-20 STORJ token, which serves as a means of exchange for storage services. Notably, the Storj network was one of the few Ethereum ICO projects to earn Vitalik Buterin's endorsement, and he was a contributor to the Storj whitepaper **[3]**. The Proof-of-Work system of Ethereum consumed 5.13-GW while the existing Proof-of-Stake system consumes 2.62-MW (2,000-times reduction as in Fig-2.4).

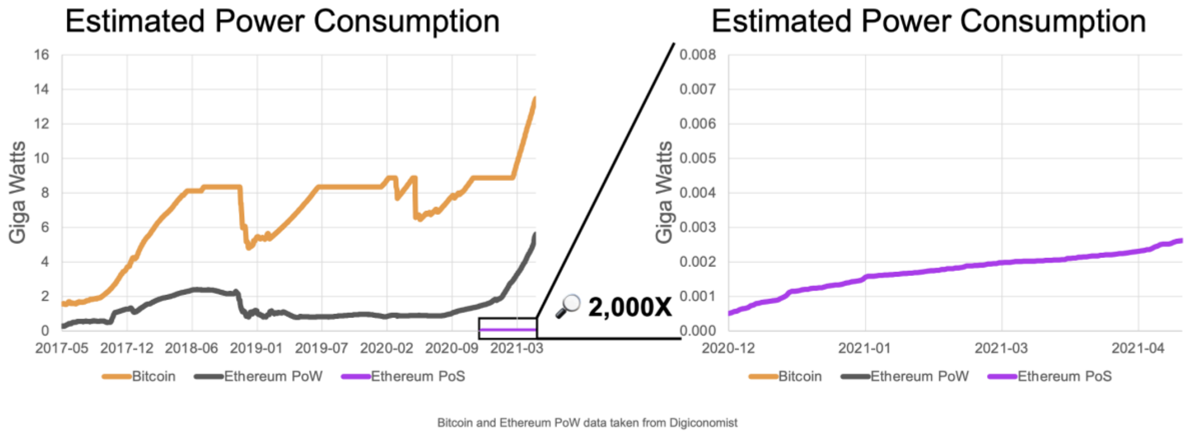


Fig 2.4 : Power Consumption PoW Vs PoS – Ethereum consumes 99% lesser at present

Source : Ethereum

STORJ plans to purchase carbon offsets for these 23,496 metric tons of carbon as shown in Fig-2.5.

	Storj Controlled Wallets	All STORJ transactions on Ethereum
Total Transactions as of 1/6/2022	222,545	1,702,952
Total Transaction Gas used	8,693,318,864	64,920,433,556
2020 Grams CO2e per Gas unit	0.18	0.18
2022 Adjustment factor	2.00	2.00
Average kg CO2e per transaction	14.14	13.80
Metric Tonnes of CO2e total	3,146	23,496

Fig 2.5 : Carbon Footprint of STORJ and its transactions on Ethereum

Source : STORJ

As a Not-for-Profit, the International Maize and Wheat Improvement Centre (CIMMYT) faces a challenge in sharing its research with the public, as some datasets can over 7 terabytes (TB). Many audience of CIMMYT reside in developing countries, where internet connections are often slow/unstable. Retrieving different pieces of each file from the closest Storage Node and downloading a single large file as multiple small files can better handle unstable connections. This distribution method helps them reach their audience more effectively **[4]**.

3. The role of Blockchain technology in pursuing this goal

Blockchain technology and cryptographic techniques play a crucial role in Storj. STORJ utilizes a Kademlia-distributed hash table (DHT) to store pointers to data on nodes across its network. File data is split into segments, encrypted, further split into stripes and then distributed across multiple nodes using erasure encoding (Fig-3.1). Erasure shares are used to reconstruct data in case of node failure, where each share represents a portion of the original data. Uplink API allows developers to access and interact with the network to store and retrieve data.

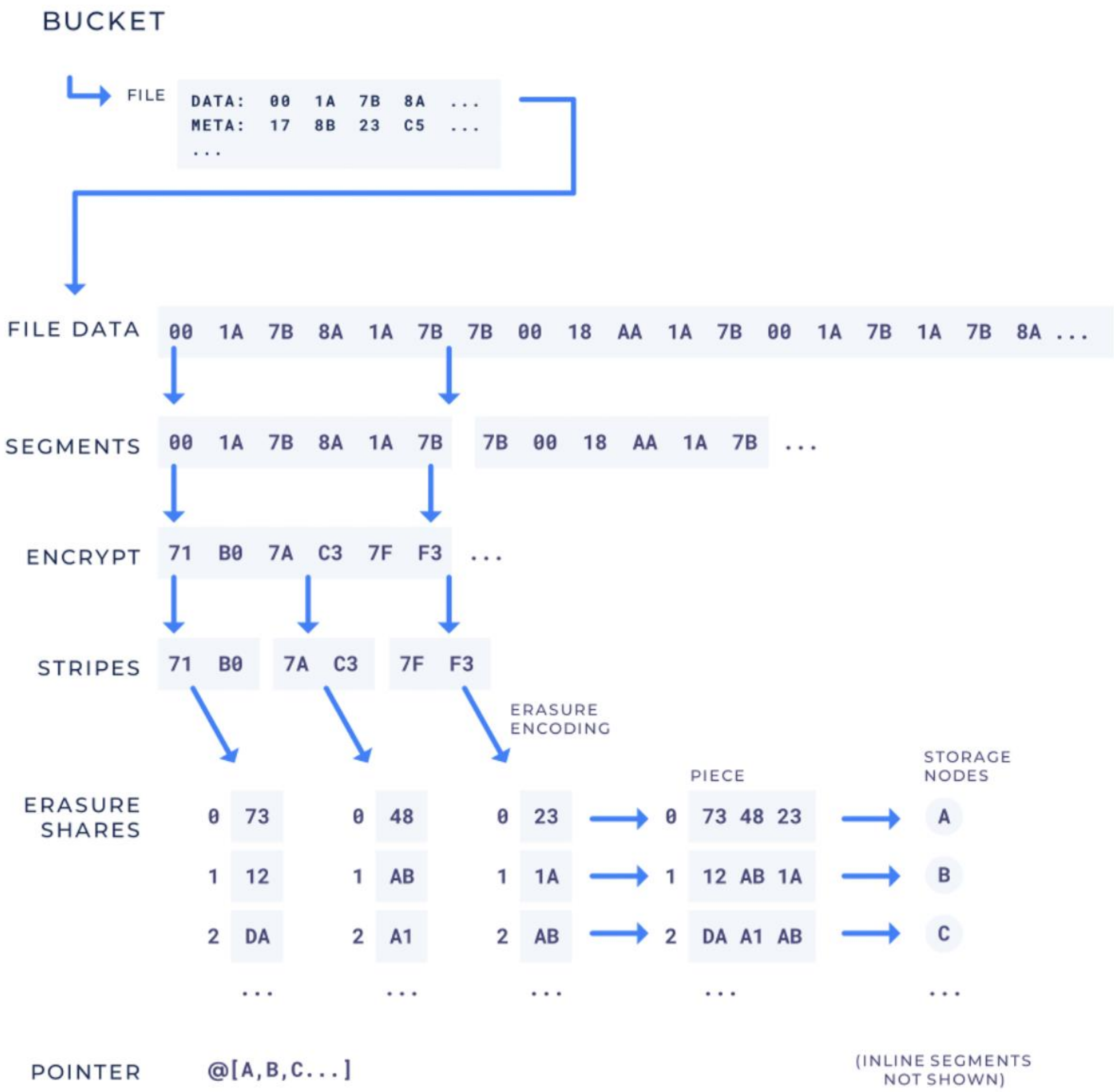


Fig 3.1 : Sharding and Retrieval

Source : STORJ v3 Whitepaper

Upon request, the network assembles the file from the collected shards and delivers it to the client. Storj uses an audit system to verify farmer compliance, which involves pre-generated audits (PRA) and the Merkle-Audit system (MA). Farmers must furnish the required shard to pass the PRA, while the MA compares the hash of each stored shard to the initial hash supplied by the client to verify adherence. Additionally, the use of end-to-end encryption and cryptographic hashes ensure that data is protected.

Dr.Antonin from University of Edinburgh conducted a case study "Storing high-performance computing data using Storj Decentralized Cloud Storage" **[5]**. According to the study, the most efficient method of transferring large files to and from Storj DCS is through Storj's native parallelism with multi-file parallel transfers.

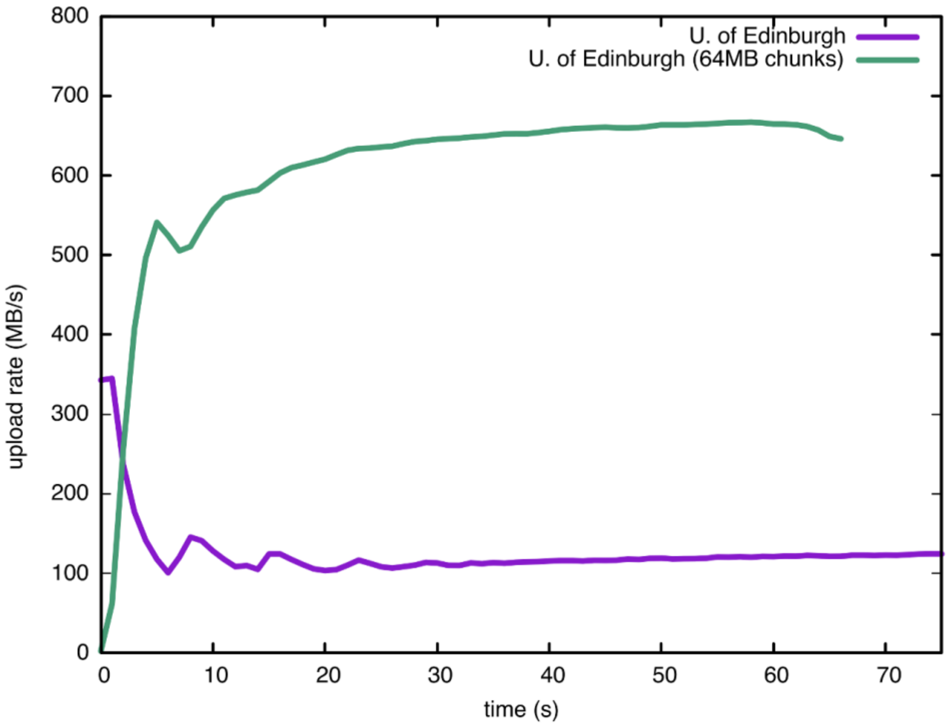


Fig 3.2 : Upload Rates from UoE DiRAC Supercomputer

Source : STORJ Case Study

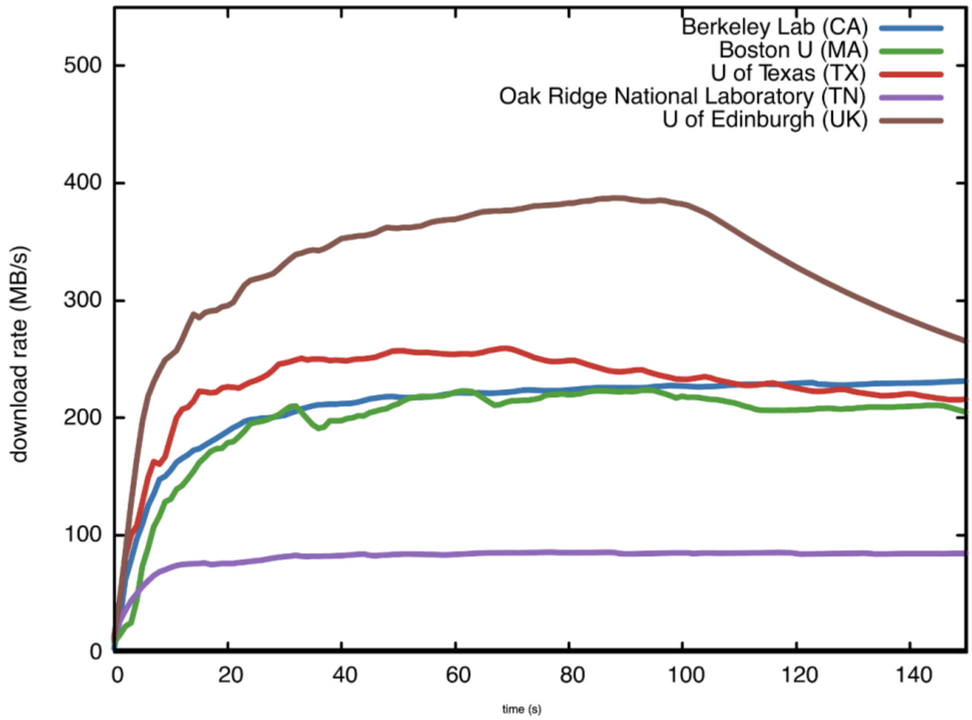


Fig 3.3 : Download Rates from Storj DCS for the original dataset

Source : STORJ Case Study

Parallelism and chunking approach utilize the user's full bandwidth, resulting in faster transfer speeds, reduced transfer times, and minimizes data loss risks. The Edinburgh-uploaded data (Fig-3.2) was retrieved from multiple supercomputing centres in the Edinburgh and USA (Fig-3.3 &3.4). Unsurprisingly, the study found that Edinburgh yielded the best results due to concentration of some data in Europe.

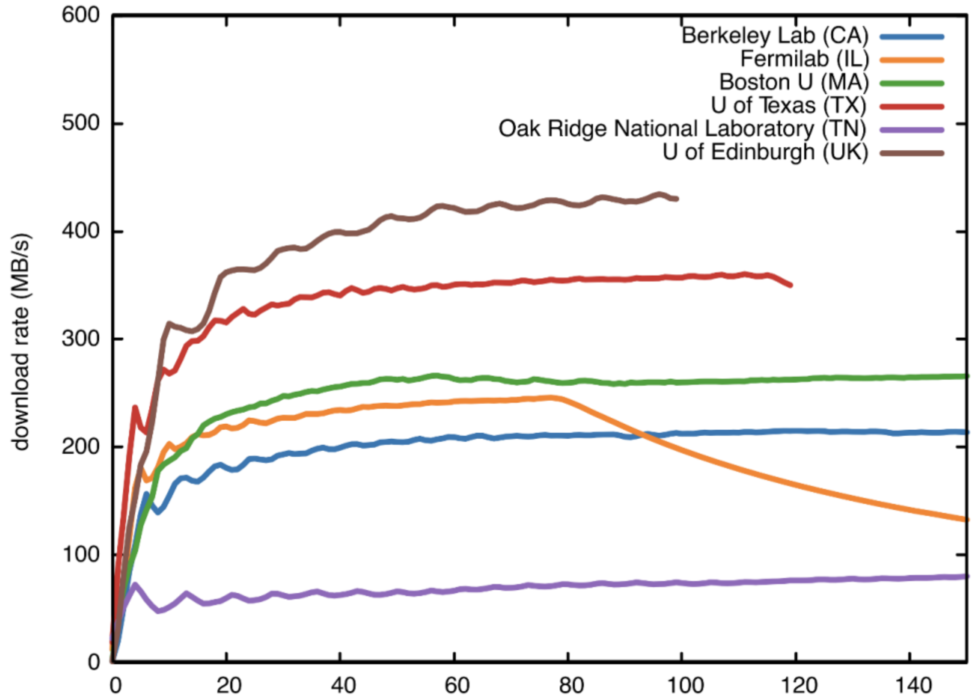


Fig 3.4 : Download Rates from Storj DCS for the chunked dataset

Source : STORJ Case Study

Chunking has a considerable effect on upload and improved by a six-fold increase (~650 MB/s Versus ~100 MB/s) except at Oak-Ridge. Additionally, chunking appears to enhance download rates between 10% and 30%, varying based on the centre's geographical position.

4. Evaluation of the feasibility of the mechanism

I) Cost and Storage Savings: A Storj Labs study found that using their network for storing data can be up to 90% cheaper than using Amazon-S3.

Fig 4.1 : Cost and Egress Comparison per TB

Source : STORJ

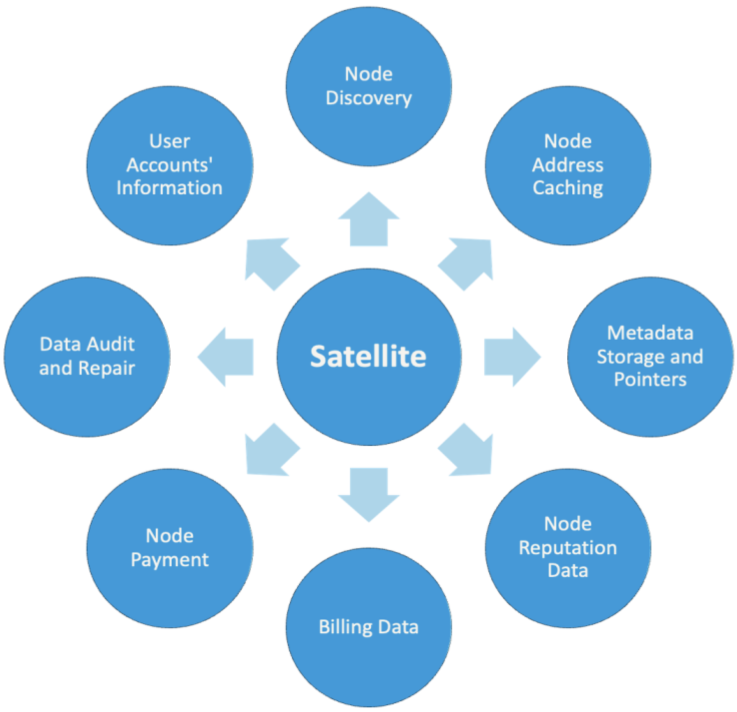
To achieve enterprise-grade 99.95% availability, Storj uses **Reed-Solomon** erasure coding that automatically splits each object file into 80 or more encrypted pieces across 13,500 geographically diverse nodes running on different ISPs and power grids across the globe **[6]**. The study demonstrated that erasure-encoded systems' mean time to failure (MTTF) is significantly higher than that of replicated systems with the same storage overhead and repair period. Therefore, using erasure coding and taking advantage of temporal and spatial locality can significantly enhance data availability to 8 nines (0.999999998) with only rate 1/2 erasure, while reducing storage and bandwidth consumption **[7]**.

# Storage Providers	# Losses To Tolerate	Replication Redundancy	Erasure Coding Redundancy	Storage Savings Over Replication
3	1	2	$1\frac{1}{2}$	$\frac{2-\frac{3}{2}}{2} \times 100$ or 25%
4	1	2	$1\frac{1}{3}$	$\frac{2-\frac{4}{3}}{2} \times 100$ or 33%
4	2	3	2	$\frac{3-2}{3} \times 100$ or 33%
5	1	2	$1\frac{1}{4}$	$\frac{2-\frac{5}{4}}{2} \times 100$ or 38%
5	2	3	$1\frac{2}{3}$	$\frac{3-\frac{5}{3}}{3} \times 100$ or 44%
5	3	4	$2\frac{1}{2}$	$\frac{4-\frac{5}{2}}{4} \times 100$ or 38%

Fig 4.2 : A representative example of a puppy image stored across different providers without encryption and Storage savings of Erasure Coding Vs Replication

Source : https://towardsdatascience.com/erasure-coding-for-the-masses-2c23c74bf87e

II) **Reliability:** STORJ ensures high levels of data reliability by using Satellites and distributed storage. The company conducted a benchmark test that resulted in achieving a 99.999999999% (12 nines) durability rate, which indicates that the chance of losing data stored on the network is extremely low.



Smart-Art 1 : Functions of a Satellite

Source : Created by author

Data-Source : STORJ V3 Whitepaper

Although any user can run their Satellite, most will likely choose to avoid the operational complexity and create an account on Trusted Satellite hosts like Storj Labs, which operates six Satellites.

III) **Proof of retrievability:** Proof of Retrievability (PoR) protocols typically operate by selecting a small, random subset of the file, called a challenge block, which the client must then return to the verifier. The verifier can then verify that the challenge block is indeed a part of the file and that the client possesses the rest of the file **[8]**. Storj platform operates in an untrusted environment where individual nodes are not necessarily trustworthy. To ensure reliability and fault tolerance in such an environment, Storj adopts the Byzantine, Altruistic, Rational (BAR) model and participants in the network are classified as follows: Byzantine nodes may actively try to sabotage the protocol. Altruistic nodes are good actors and participate in the protocol even if it is not in their best interest, while Rational nodes participate or deviate only when it is in their net best interest.

IV) **Security:** While AWS and Azure use erasure coding in their storage systems, the key difference is that the network utilizes erasure coding in conjunction with blockchain, distributed nodes, Kademlia-DHTs, BAR models and Merkle Audits. This combination provides an additional layer of censorship resistance that is not present in centralized cloud storage systems.

V) **Token Pricing and ICO:** Storj token pricing is influenced by a variety of factors, including supply/demand and network utilization. This is indirectly influenced by the platform's use of a reputation system, BAR model to incentivize/penalize good & bad behaviour and satellite billing system to create a transparent system.

- **Network utilization** of the people drives the demand for tokens and the price is likely to go up or down, accordingly.

- **Reputation and Incentivization system** tracks the performance of its storage providers. Storage providers with more uptime and consistent better reputations may be able to receive more tokens, driving up the price. This, in turn, improves the overall quality of the network, attracting more users and demand for tokens, which can drive up the price due to its limited supply.

	Information	Number of STORJ Tokens (Millions)	Percentage
1	Tokens allocated to projects and founders	427.95	85.59%
2	Tokens allocated to retail investors (30 million USD Raised - ICO)	72.05	14.41%
	Tokens Distribution amongst projects and founders (427.95 million - 100%)		
3	Storj Labs	107.92	25.22%
4	Reserves held by STORJ	245.03	57.25%
5	Burned Tokens	75	17.53%
	Total Supply after burning [2 + 3 + 4 = 425 million]		

Fig 4.3 : STORJ Tokens Distribution after Ethereum migration

Source : Created by Author

Data-Source : https://beincrypto.com/

Approximately 413 million tokens are in circulation after lost/burnt tokens in the past few years. Based on the supply breakdown and release schedule, the cryptocurrency could become deflationary - it is possible that the maximum total supply of tokens could be reached by 2027 or earlier.

5. Possible improvements and prediction of the projected future

I) Possible DOS Attacks in the Satellite: The satellite interface is vulnerable to DDoS attacks, which could render the entire system useless. During testing, it was found that closing connections in the 'pingSatelliteOnce' function was time-consuming, resulting in more connections being established in a shorter period. This could potentially overload the satellite's resources, leading to a DoS attack. The satellite spins up go routines for each connection, which can consume resources and ultimately lead to a crash if not correctly closed. Implementing attack detection techniques and modifying the dRPC library are possible solutions. Additionally, the server can filter requests by allowing only one connection within a specific time and blacklisting repeat offenders **[9]**.

II) Frameup and Optimized Attacks: Storj is vulnerable to frameup attacks where clients can upload malicious files to storage nodes (Fig-5.1). The design can be exploited by attackers to target farmers with the least latency and highest transfer rate to the renter, thus increasing the precision of their attack.

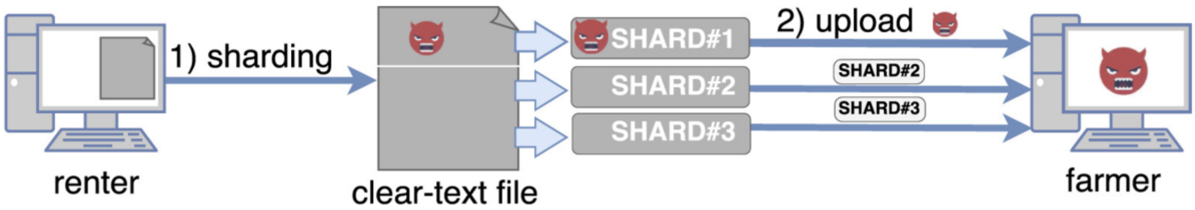


Fig 5.1 : Frame-Up Attack in STORJ

Source : https://doi.org/10.1016/j.diin.2019.02.003.

Storj could implement attack detection techniques such as entropy tests or signature-based approaches, as well as modifying the renter application to prevent users from disabling encryption. However, these approaches may result in rejected contracts and wasted network traffic, so balancing efficiency and complexity is necessary **[10]**.

III) Reserve Tokens: The platform holds a large percentage (57%) of reserve tokens, which could suggest a lack of demand for storage services or ineffective token utilization. Maintaining a balance between token reserves and usage is important for a healthy and sustainable ecosystem. Holding reserves of a centralized currency may not align with Decentralization Advocates' approach to technology.

IV) Data Regulations: HIPAA and GDPR are regulations that safeguard sensitive health and personal data in the US and EU, respectively. Both centralized and decentralized storage platforms must comply with these regulations. Decentralized platforms like STORJ, which operate globally, face challenges in adhering to data localization requirements. The 2021 Mastercard ban in India highlights the importance of compliance with data localization requirements for both centralized and decentralized storage platforms **[11]**. One potential solution could be for STORJ to partner with local data centre providers across different regions to ensure compliance, while maintaining its decentralized architecture at the same time.

V) Payment Mode: STORJ's "Payment-Agnostic" policy of accepting other popular cryptocurrencies as payment from renters could have some negative implications. It could also cause payment processing complexity, confusion for renters, and diversion of liquidity away from the STORJ token. The advantages of increased payment flexibility should be weighed against the disadvantages of increased complexity in managing payments and revenue fluctuations due to cryptocurrency volatility. The impact on STORJ's platform and token value will depend on the adoption of other cryptocurrencies and the ability to effectively manage payment processing complexity.

6. An Alternate plausible logic for token pricing of this project

- An alternative logic for STORJ's token pricing could be based on the computational power of nodes and utilization rate of storage.
- The platform could use a *Dynamic Pricing Model* that adjusts the token price based on supply and demand for storage and computational power, offering incentives to operators who provide high-quality and frequently utilized resources.
- Token burning could be implemented proportionally to the amount of storage and computational power utilized by renters, reducing the token supply and increasing its scarcity and value.

References

[1] Cai, C., Yuan, X. and Wang, C. (2017). *Towards trustworthy and private keyword search in encrypted decentralized storage*. [online] IEEE Xplore. doi: https://doi.org/10.1109/ICC.2017.7996810.

[2] Investopedia. (2022). *Storj (STORJ) Definition and Use Cases*. [online] Available at: https://www.investopedia.com/storj-cryptocurrency-definition-5224160 [Accessed 1 Mar. 2023].

[3] www.storj.io. (2022). *The Green Case for Storj*. [online] Available at: https://www.storj.io/blog/the-green-case-for-storj [Accessed 1 Mar. 2023].

[4] CIMMYT discovers sustainable object storage in Storj DCS. (n.d.). Available at: https://www.storj.io/documents/storj-case-study-CIMMYT.pdf [Accessed 1 Mar. 2023].

[5] Portelli, A. (2021). *Storing high-performance computing data using Storj Decentralised Cloud Storage Introduction and objectives*. [online] Available at: https://www.storj.io/documents/storj-univ-of-edinburgh-report.pdf [Accessed 1 Mar. 2023].

[6] GitHub. (2018). *Storj Whitepaper V3*. [online] Available at: https://github.com/storj/whitepaper [Accessed 28 Feb. 2023].

[7] Weatherspoon, H. and Kubiatowicz, J.D. (2002). Erasure Coding Vs. Replication: A Quantitative Comparison. *Peer-to-Peer Systems*, pp.328–337. doi: https://doi.org/10.1007/3-540-45748-8_31.

[8] Mahmood, W. and Wahab, A. (2018). *Survey of Consensus Protocols*. SSRN Electronic Journal. doi: https://doi.org/10.2139/ssrn.3556482.

[9] de Figueiredo, S., Madhusudan, A., Reniers, V., Nikova, S. and Preneel, B. (2021). Exploring the storj network. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*. [online] doi:https://doi.org/10.1145/3412841.3441908.

[10] Zhang, X., Grannis, J., Baggili, I. and Beebe, N.L. (2019). Frameup: An incriminatory attack on Storj: A peer to peer blockchain enabled distributed storage system. *Digital Investigation*, 29, pp.28–42. doi:https://doi.org/10.1016/j.diin.2019.02.003.

[11] Ahmad, P.A. (2021). Cyber Security Is More than Just a Question of Information Technology. *Journal of Image Processing and Intelligent Remote Sensing*, (12), pp.1–7. doi:https://doi.org/10.55529/jipirs12.1.7.