COLDSTACK: Decentralized Cloud Storage Aggregator

Lite Paper v 1.4.7   updated Feb 2 2021

Abstract

Nowadays the internet is still in the middle of a revolution: centralized proprietary services are being replaced with decentralized solutions and monolithic services replaced with peer-to-peer algorithmic markets. That's exactly what **Decentralized Cloud Storage Platforms** such as *Filecoin*, *SIA*, *Arewave* and *Storj* are doing to cloud storage market, providing their users with more security and privacy, lower costs and much higher redundancy.

However, there are still a lot of challenges Decentralized Cloud Storage is facing on the way to the mass adoption:

- Integration can take a lot of time and effort;
- No unified API or billing mechanism exists for Decentralized Clouds;
- Developers cannot reuse existing tools or SDK clients;
- It is no easy way to combine multiple Decentralized Clouds in one dApp;
- Migration from one Cloud Storage to another is a very hard task.

ColdStack is an **Uber for Decentralized Clouds**: the world's first Aggregator, which provides users with unified API and unified billing. We buy the cheap storage space, optimize the final costs with our AI-based pipeline and offer a **single entry point** to any Cloud.

That's how we supply our users with the world's most affordable storage space from any Cloud, just like Uber supplies its clients with cost-effective rides and deliveries from a lot of drivers and taxi services. It represents a fundamental shift in the efficiency and economics of Data Clouds that enables us to provide the **world's best price** for data storage: $2.4 per Terabyte per month.

Contents

1. Introduction

1.1 Project Overview

ColdStack is the world's first **Decentralized Cloud Storage Aggregator**. We supply our users with a unified API that allows them to upload and retrieve data to any Cloud Storage.

Thus, they were spared from having to deal with different APIs of different Decentralized Clouds. When the file is uploaded, we provide users with a unique ID and use our blockchain to store all the upload (and, later on, billing and retrieve) transactions. That allows users to combine multiple Decentralized Clouds in one dApp and makes it unnecessary to perform any data migration while switching from one Decentralized Cloud to another. It is designed to deal with a bulk data such as logs or archives for projects like DeFi analytics dApps, Web3 infrastructure services or DAO Platforms.

ColdStack is a decentralized platform so we issue **5,000,000 $CLS tokens (ERC20)** for our ecosystem. Its circulation volume is strictly **pegged** to the volume of the currently purchased storage space.

The space is 100% **auditable** and **verifiable** because all the storage/retrieve transactions will be logged in the blockchain.

ColdStack has its own **NFT** (non-fungible tokens: ERC-721/ERC-1115) engine. It allows users to create NFT token for any of their files stored into the System. Token format is 100% compatible with major **NFT marketplaces** like Super Rare or Open Sea. This enables users to seamlessly tokenize their files, turning them into digital assets that can be immediately offered for sale on those marketplaces.

This feature makes ColdStack the most affordable **Universal Digital Asset Storage**.

We have **two blockchains** in ColdStack: **Ethereum** mainnet and **Polkadot** blockchain for billing and storage/retrieve transactions. We have chosen to use Polkadot infrastructure because of speed and affordability of transaction fees. We supply users with a **cross-chain bridge** that will allow users to deposit / withdraw ETH $CLS tokens.

It's important to note about our **Yield Farming**: ColdStack offers **reward tokens** for staking $CLS tokens (see "Yield Farming" in Token Economy section, page 18).

1.2 What are Decentralized Clouds

Decentralized Clouds, or Decentralized Storage Networks (**DSNs**), are blockchain-based **Cloud Storage Platforms**. They store clients' files in shards on the storage space provided by the crowd of independent suppliers. Like an Internet, they have no central authority and use advanced decentralized technologies to deal with the data.

Current total capitalization of top DSNs like Filecoin, SIA, Arweave, Lambda and Stroj is over **$2bn**. Each DSN has its own crypto currency to carry out transactions between the clients and other ecosystem participants.

Main features of DSNs:

1. **Low cost**   cutting all the overhead costs like banking and management costs make DSNs really cheap;

2. **High latency**   due to its blockchain-based technology, DSNs data access speed is relatively slow;

3. **Reliability**   DSNs are decentralized: i.e., truly distributed and high secured;

4. **Privacy**   all the data in DSNs is fully encrypted and all the clients are 100% anonymous.

1.3 What are Cold Data Storages

Cold Data Storage is used to store the inactive data that is rarely used or accessed. However, that "cold" data but must be stored for a long time or even forever for business or compliance purposes. That data is like the books on the very top or bottom shelf of a bookcase in your office (information a business is required to keep for regulatory compliance, video, photographs, and data that is saved for backup, archival or disaster recovery purposes.) For example, no less than 80% of Facebook data according to the report of K. Patiejunas (Facebook Storage Architect) can be called cold data.
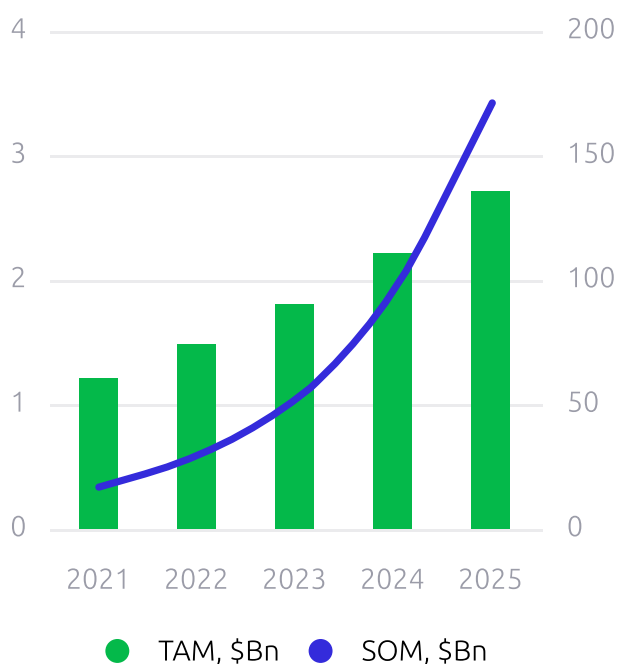
Advantages of Cold Data Storages:

- Prevent primary storage from becoming overloaded with inactive data;
- Reduce overall storage costs;
- Simplify storage and data management.

2. Industry overview

Total Available Market (TAM) for Cloud Storages is $50.1Bn in 2020 and will be $137.3Bn in 2025, at a Compound Annual Growth Rate (CAGR) of 22.3% during the forecast period. The Cloud Storage market saw a rise mainly due to growing data volumes across enterprises, rising need for providing the remote workforce with ubiquitous access to data and files, and cost-saving and low Total Cost of Ownership (TCO) benefits of Cloud Storage solutions.

Analysts from IDC noted the amount of data created in 2020 exceeded 44 Zettabytes and continue to accelerate exponentially from there. Up to 60% of that data isn't active or frequently accessed and may be classified as a cold data.
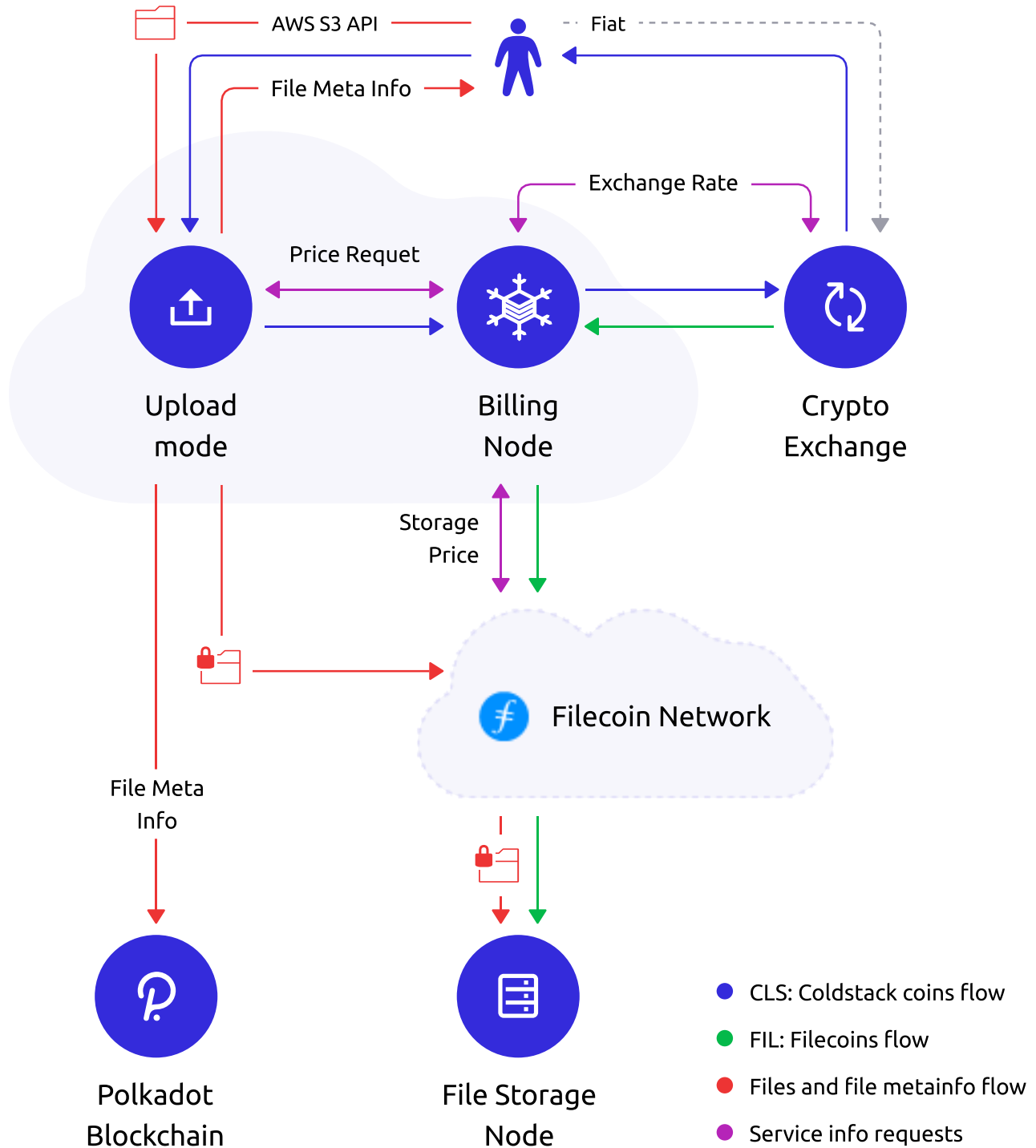

TAM, $Bn • **SOM, $Bn**

ColdStack is an aggregator of Decentralized Data Clouds. Their distinctive features are low cost, high latency and quite slow access speed. And the same is true for the market niche of **Cold Data Storages** like *Amazon Glacier*, *Google Coldline Storage* and the same services of *Microsoft*, *IBM* and *Oracle*: services, which were specially designed to store and retrieve the cold data.

Serviceable Available Market (SAM) for Cold Data Storages is $12.5Bn in 2020 (20% of the whole Cloud Storages market), at a current CAGR of 60%.

We have most competitive prices but there is also a whole bunch of very strong competitors like Amazon and Google, so we can estimate our **Serviceable Obtainable Market** (SOM) in 2021 as 2% of the available market with the prospective growth to 10% in 2025: that will be approximately **$300M** in 2021 and up to **$3.5Bn** in 2025.

3. Product Description

3.1 High-level system description

AWS S3 API

Fiat

File Meta Info

Exchange Rate

Price Requet

Upload mode

Billing Node

Crypto Exchange

Storage Price

Filecoin Network

File Meta Info

Polkadot Blockchain

File Storage Node

- CLS: Coldstack coins flow
- FIL: Filecoins flow
- Files and file metainfo flow
- Service info requests

To start using the Service, user must have his own ETH wallet: all the account information is linked to the wallet and **no KYC** or other identity verification required. User can specify his e-mail address for notification purposes, but it is optional. All the AWS-style security tokens assigned to each account will also be linked to the ETH wallet.

To start the data storage / retrieval, user must transfer necessary amount of $CLS tokens to his wallet (i.e., personal account) on our Polkadot blockchain using our cross-chain bridge. All the transactional billing will be performed there to save a lot of gas and accelerate transaction processing.

This part of the System is developed based on the **Substrate framework** and can run independently on the blockchain. It will provide us with multi-level permissions, encrypted transactions and state, limiting to asynchronous calls optionality and governance tools and methods such as stakeholder referendums, approval voting, and qualified abstention biasing.

3.1.1 Upload

The **Upload Node** is responsible for all the user's interactions during the file upload.

Easiest way to upload files to a Cloud is **Amazon Simple Cloud Storage (S3) API**, which becomes the de-facto standard for the Cloud Storage industry. We have implemented the AWS S3 API to allow the users to move and manage data with no or minimal coding and makes it extremely easy to migrate data from *Amazon* to ColdStack. As it was mentioned before, all the AWS security tokens are linked to $CLS ETH wallet.

ColdStack stores data under a **Universally Unique Identifier** (UUID) key. User stores his data by paying to the System in $CLS tokens. Minimum charge time of the data storage is 30 days. Incoming traffic is free of charge.

User's file will be eventually stored in some of the Decentralized Clouds, or Decentralized Storage Networks (DSNs): SIA, Arweave, Filecoin, Lambda, Storj, 0Chain etc. Filecoin is presented on a flowchart above as an example of DSN and most popular of them.

Before storing the data, Upload Node performs a price query. It is necessary to check the user's balance before charging and calculate the exact sum, that will be charged as a file storage price.

The query will be submitted to the **Billing Node**, which will perform the following operations:

1. **Request the token's exchange rate**

   The rate is requested from Crypto Exchange, where DSN's token is listed. Usually, it will be the cross pair (like FIL/ETH and ETH/CLS). To save time and money, requests will be cached (1 request per 10 minutes).

2. **Clarify the storage cost**

   Storage cost for the exact file will be calculated in DSN tokens. Billing node need to request the DSN (in the example displayed on the flowchart above, it is Filecoin) for the exact price.

3. **Charge the Upload Node**

   Billing Node will charge the Upload Node for the file storage in $CLS tokens. In it's turn, Billing Node will arrange the storage payment for the Upload Node in DSNs tokens.

That will make able for the Upload Node to store user's file in the DSN, and DSN in turn will take care about storing user's file on some of its Nodes.

3.1.2 NFT issuance

Immediately after file uploading, upon a user request, we can issue a **Nonfungible Token (NFT)** linked to the uploaded file. Non-fungible tokens are used to create verifiable digital scarcity, as well as digital ownership, and the possibility of asset interoperability across multiple platforms.

We can issue ERC-721 Standard or ERC-1155 Standard token at the user's choice. Token issuance cost will be charged from owner's $CLS ETH wallet. NFT will be transferred to that wallet.

Our Nonfungible Tokens are 100% compatible with major NFT marketplaces such as *SuperRare* or *OpenSea*. That feature allows the user to seamlessly tokenize his files, making them Digital Assets, which can be immediately offered for sale on those marketplaces.

Users therefore can consider ColdStack as a handy and very affordable Universal Digital Asset Storage.

3.1.3 Storage

ColdStack is a DSN Aggregator that allows users to store their files to any DSN with unified API and billing. The task of the Upload Node is to get the file from the user and cache it on its side: DSN, therefore, cares about storing file on its Nodes.

It is worth clarifying that Billing Node will charge Upload Node in $CLS and DSN will charge Upload Node in its tokens. Those tokens will be accumulated to the reasonable-sized exchange pools, which will be established and maintained on each Billing Node. Both Billing Node and the Upload Node will charge their commission fees.

As an Aggregator, we delegate the responsibility for the redundancy and security of users' data to DSNs and fully focus on **storage cost minimization** in the interests of our users. According to our studies, all the DSNs, which we use (or consider to use) as file storages for our users, guarantee the highest level of security and reliability. That means anytime we can select any of them as a file storage provider, by aiming at a better cost.

Accordingly, in order to provide the selection of most suitable DSN for each file from users request flow, we developed an AI-based pipeline. It allows to save more than 50% of the storage cost for the end user.

The savings would be achieved through:

* As a Cold Storage, we focused on cheapest storage space available (however, it still must have high redundancy!): geo regions with low demand, high latency Nodes etc.;
* AI-based purchasing system that will make smart bids on DSN internal auctions, saving 10-35% from average DSNs' selling prices;
* Stack of advanced lossless AI-based compression algorithms, saving 7-18% storage space and thus storage price;
* We constantly search for new opportunities to buy reliable and affordable storage space beyond existing sources like Decentralized Storages.

We also consider the feature that will allow users to manually select one of the DSNs, which we aggregate and use our engine to purchase space and our unified API to store the files. We will charge in $CLS tokens at DSN prices + traffic costs + 1% service fee.

3.1.4 Retrieval

File Retrieval flow is mostly similar to the Storage flow. The **Retrieval Node** is responsible for all the user's interactions, when he is retrieving the file.

User can retrieve the previously stored file via S3 REST API (by specifying UUID) or via URL (Amazon virtual buckets engine). During the retrieval transaction he will be charged in $CLS tokens. In the Retrieval Request he can also specify the time of file caching (if necessary). In this case all the requests to the data will be cached during the specified time. Retrieval Node will perform the caching services upon a user request.

Retrieval time may vary from 2 to 200 seconds (if the file is not yet cached).
Retrieval price is avg $1.2 per Terabyte (**53x cheaper than Amazon**).

3.2 Architecture overview
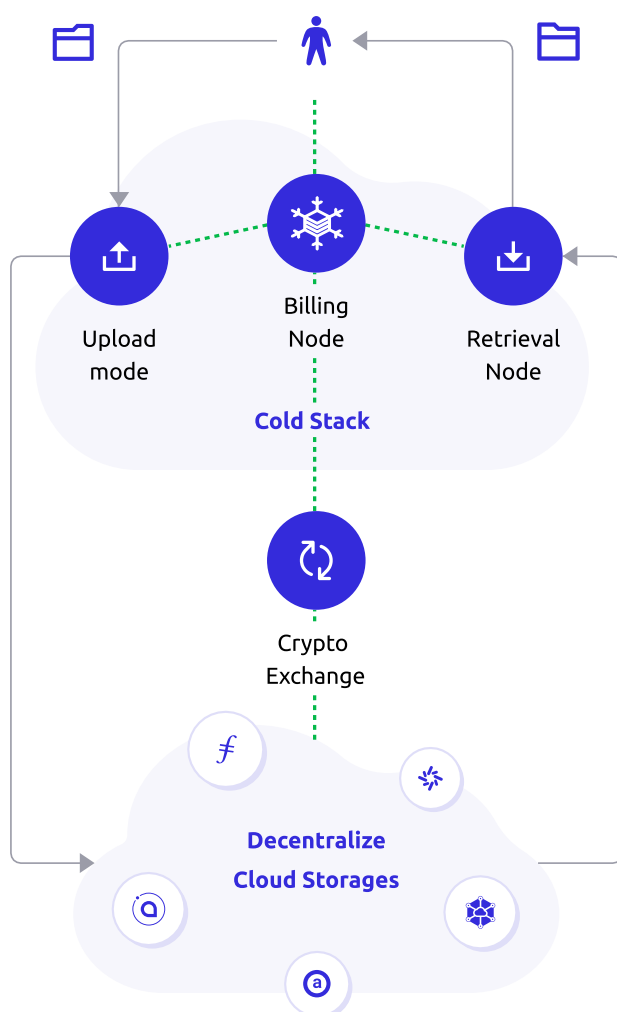
As it was mentioned above, there are three types of Nodes in ColdStack:

1.   Upload Node
This Node handles all the user interactions while uploading a file and hosts AI-based pipeline that allows to keep the extra-low storage costs.

2.   Billing Node
This Node performs all the billing, cross-chain and cross-token operations to calculate the payments and commissions and maintain the exchange pools with DSNs tokens.

3.   Retrieval Node
This Node handles all the user interactions while retrieving a file and cache the files upon user request.

Crypto Exchange, shown in figure above, plays a crucial role in the System: it allows to supply us with necessary tokens exchange rates replenish the exchange pools on the Billing Nodes (see more in Token Economy area). As it was mentioned before, ColdStack has two blockchains: ETH mainnet and Polkadot (storing the billing and data transactions info). We have a cross-chain bridge to perform the tokens transfer operations between these two blockchains.

All the Nodes in the System currently are trusted. Cross chain bridge is implemented now as simple P2P bridge (as in Bounce Finance dApp). We are going to implement consensus mechanism and fully decentralized bridge during later stages of development.

3.3 System Use Cases

Here are some ColdStack use cases, which can illustrate the possible adoption of ColdStack. We are going to publish a separate article on Medium that will be specifically dedicated to this topic. To fit the Lite Paper format, we will just give a few examples below:

1. **Web3 infrastructure project**
   ColdStack can provide the client's dApp with Web3 based "all-in-one" experience for unified data management to deal with heterogeneous data of any size.

2. **DeFi and DEX Service**
   ColdStack can store and retrieve large hybrid datasets for these services: like payout schedules and token bonding curves: most of these data isn't frequently accessed but must be stored for a very long time or even forever.

3. **Blockchain-based Game**
   ColdStack can help game developer with the cheap and reliable storage for tons of data concerning the players activities, which must be archived and stored for a very long time for different purposes (to make the game process, payments and rewards most transparent and auditable).
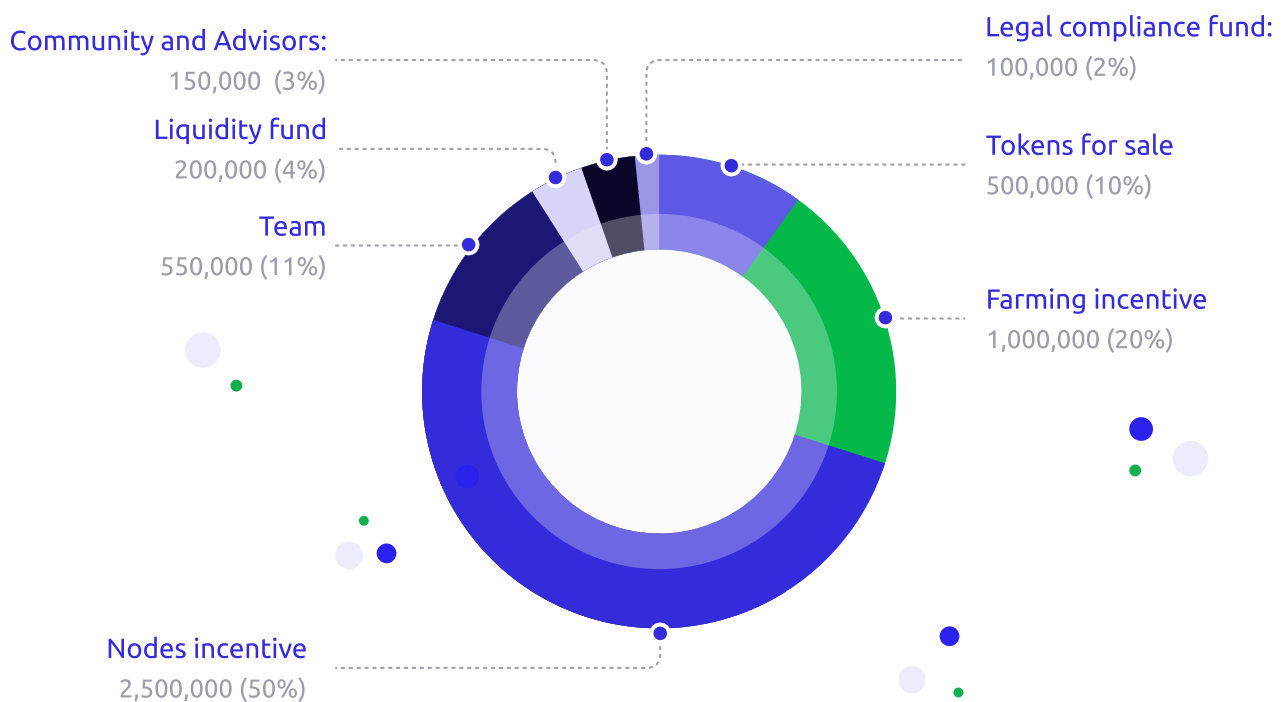
4. **NFT Marketplaces**
   Users can consider ColdStack as a very affordable Universal Digital Asset Storage for any NFT marketplace. Large files make it even more important (please note that high latency of the file storage is of little importance from the standpoint of NFT purchase transactions).

4. Token Economy

4.1 Token Metrics

ColdStack is a decentralized platform with a ERC20 token $CLS to receive payments for its services. Those tokens may be burned, but cannot be minted.

1. Total token supply: **5,000,000**
2. Initial token exchange rate: **1 USD**
3. Tokens for sale: **500,000** (10%) no locks
4. Farming incentive: **1,000,000** (20%) monthly unlock, see the details below
5. Nodes incentive: **2,500,000** (50%) demand-based unfreeze, see the details below
6. Team: **550,000** (11%) 12 months vesting
7. Advisors: **75,000** (1.5%) 12 months vesting
8. Community grants: **75,000** (1.5%) 6 months vesting
9. Liquidity fund: **200,000** (4%)
10. Legal compliance fund: **100,000** (2%)

Community and Advisors:
150,000 (3%)

Liquidity fund
200,000 (4%)

Team
550,000 (11%)

Legal compliance fund:
100,000 (2%)

Tokens for sale
500,000 (10%)

Farming incentive
1,000,000 (20%)

Nodes incentive
2,500,000 (50%)

4.2 Tokenomics

4.2.1 Storage Space and Token Supply

General idea of ColdStack tokenomics is to make $CLS **pegged** to the Data Storage space actually purchased by the clients. That requires to keep $CLS circulation volume pegged to the Data Storage volume. Thus, $CLS can only be unfrozen when the appropriate data storage volume is purchased (see details below).

The circulation volume will be pegged to the volume of the storage space, which is currently purchased. The space is 100% **auditable** and **verifiable** because all the storage transactions will be logged in the blockchain (see the diagram below).

Initially all the tokens (excluding token sale, liquidity and compliance tokens) will be frozen.

We have vesting periods for tokens intended for the team, advisors, community etc.

10% of the tokens will be sold during the token sale.

50% of total supply will be unfrozen by the smart contract, when it will be necessary for the increased storage space (see below).

As it was mentioned before, ColdStack has two blockchains:

- ETH mainnet to deal with crypto exchanges, DEX, DeFi services etc.
- Polkadot to store the billing and data transactions information

When the user starts to use ColdStack to store his data, he will need to transfer $CLS from ETH to Polkadot via cross-chain bridge to deposit it on his personal account or vice versa to withdraw.

The example chart below indicates that there are two types of Data Storage Updates records in the Polkadot transactional blockchain:

- **Data Point Creation**: new file with the specified size and ID is added
- **Data Point Completion**: old file with the specified size and ID is removed
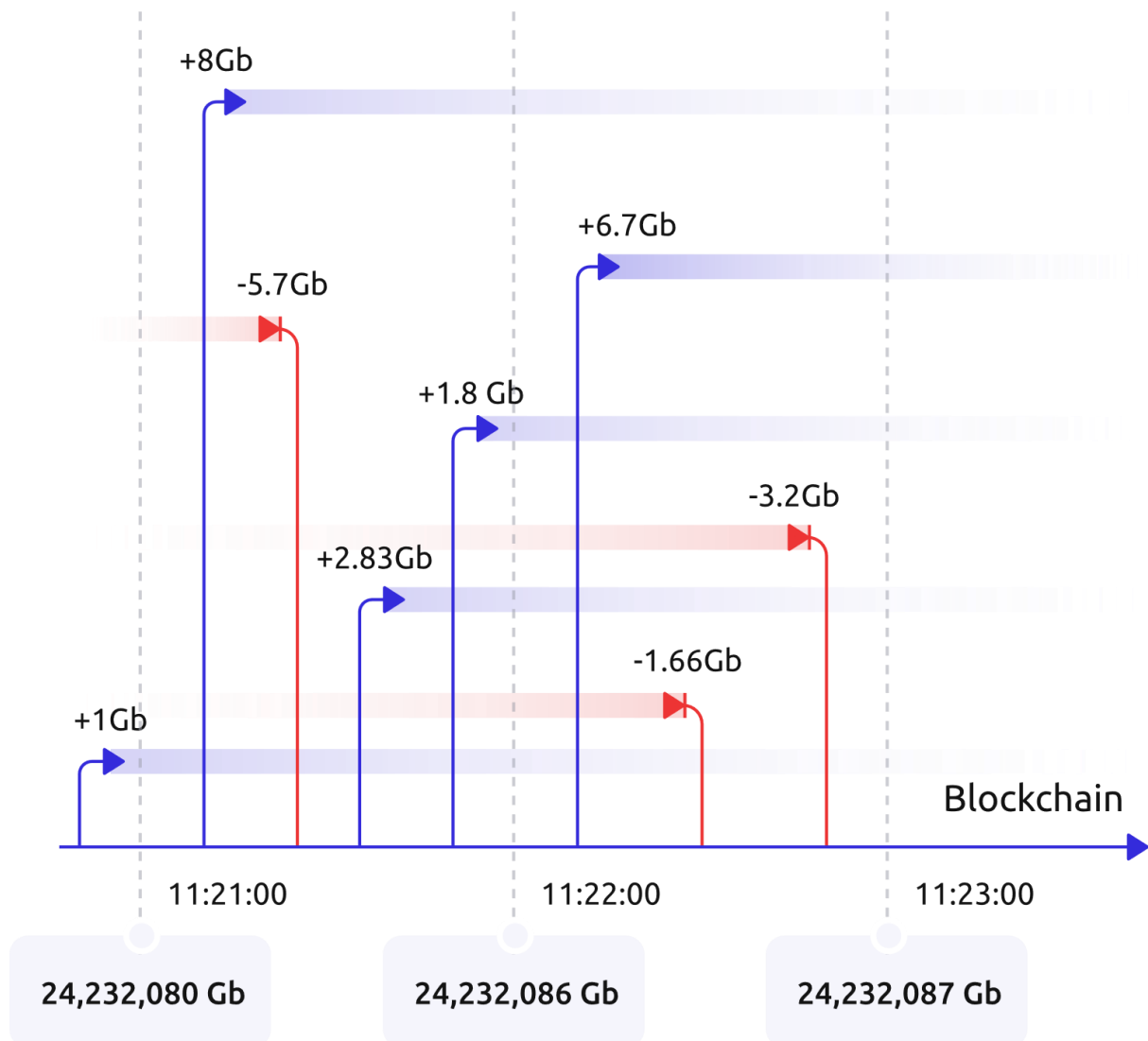
According the chart, during a minute since 11:21:00 there were 4 new blockchain records: new file 8Gb added, old file 5.7 Gb removed (30 days period is over and file's user did not pay for next 30 days so it will be purged from the storage), new files 2.83 and 1.8Gb added.

Total storage volume increased by 9.93Gb (=8-5.7+5.83+1.8).

It makes really easy to calculate the total storage volume at any moment of time for the System smart contract as well as for any third-party audits. We will periodically publish verifiable Project Status Reports with aggregated info.

4.2.2 Demand-based token unfreeze mechanism

The Minimum period to store data in ColdStack is 30 days. Therefore, we can predict the token circulation volume, which will be necessary 30 days later. Prediction will help us to calculate the exact number of tokens we will need to unfreeze.

Calculations are based on:

1. Growth rate of the Data Point Creation records for the last 30 days
2. (Data Point Creation) / (Data Point Completion) ratio for the last 30 days
3. The Download traffic / Storage volume ratio for the last 30 days
4. The 4-month moving average of the $CLS exchange rate
5. The current volume of the circulating $CLS tokens

To ensure a sustainable appreciation of the token, we aim to consistently supply the lower number of tokens compared to the required circulation forecast. This is controlled by adding the **Deficit Rate** parameter, an additional multiplier, which brings the next month's circulation volume by **15%** below the demand forecast.

The circulation volume Ct+1 to be supplied for the next month is thus calculated as:

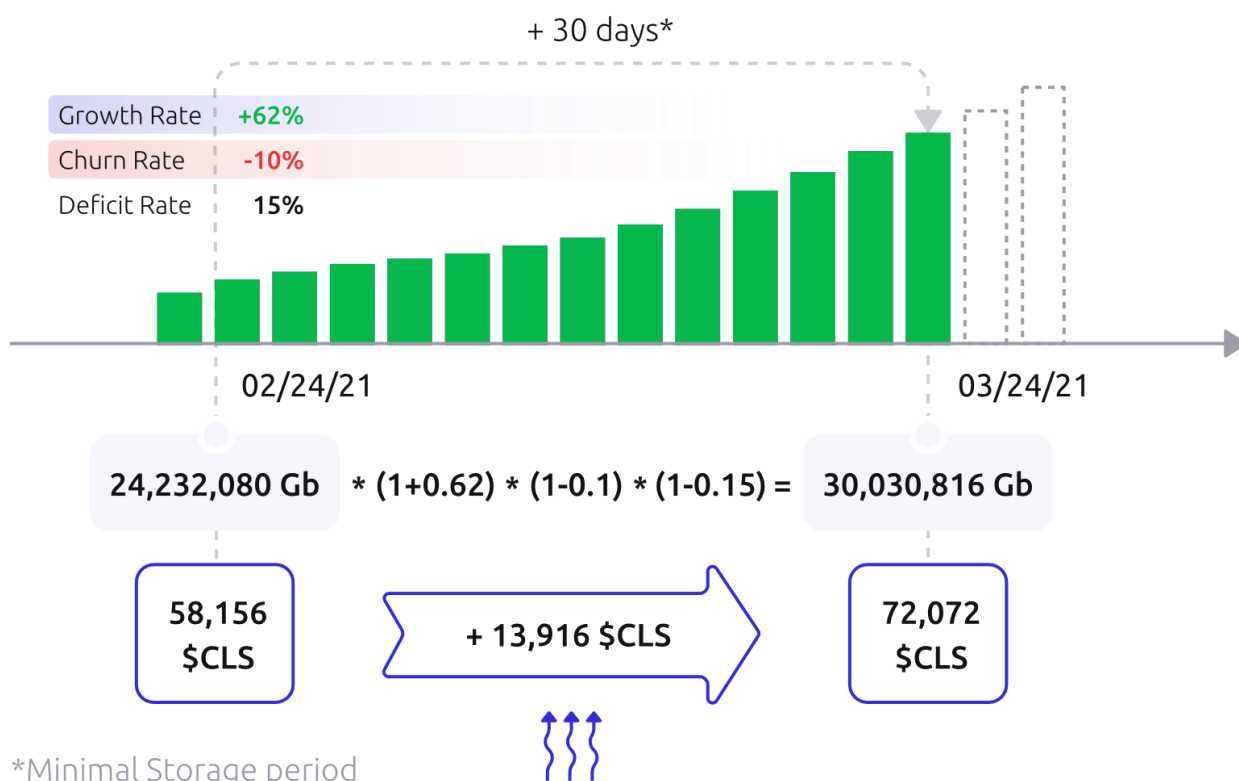$$C_{t+1} = \frac{T_t}{R} * \left[(1 + g_t) * (1 - h_t) * P_s + (1 + l_t) * P_l \right] * (1 - d) \, ,$$

Where Tt is the current purchased storage volume in Tb,

gt - growth rate, ht - churn rate, lt - download traffic rate,
Ps - storage price, Pl - download price, denominated and fixed in USD,
d - deficit rate,
R - the 4-month moving average of the exchange rate.

The actual token volume to be unfrozen is, in its turn, identified as the positive difference between the next month's circulating volume, Ct+1, and the current circulation Ct. The final unfrozen volume, Ut+1 is defined as

$$U_{t+1} = max\{0, C_{t+1} - C_t - F_t\}.$$

This takes the monthly due farming rewards, Ft, into account, as detailed in the next section.

+ 30 days*

Growth Rate	+62%
Churn Rate	-10%
Deficit Rate	15%

02/24/21   03/24/21

24,232,080 Gb * (1+0.62) * (1-0.1) * (1-0.15) = **30,030,816 Gb**

58,156 $CLS   **+ 13,916 $CLS**   **72,072 $CLS**

*Minimal Storage period

The example chart above indicates how unfreezing mechanism works: for instance, the total size of the files in all Data Point Creation records for the previous period was 62% of the total Data Storage space at the beginning of the period. Total size of **Data Point Completion** records was 10% total Data Storage space. Let's assume that current token price and Download traffic / Storage volume ratio have not changed considerable and Deficit Rate is 15%.

4.2.3 System growth

Initial / current demand of $CLS tokens is 50K per month (initial token exchange rate $1), according to the current commitments with the existing clients (companies, which have paid initial deposit and are going to start using ColdStack as soon as it's ready).

Hard cap of the initial public sale is **500K**.

According to our Financial Plan, that will be sufficient to accomplish all the necessary development and launch a System. We are going to make it fully operational within four to six weeks after the date of Public Sale (including all the initial Nodes deployment and uploading initial set of clients' data).

Storage   Smart CDN   Retrieve
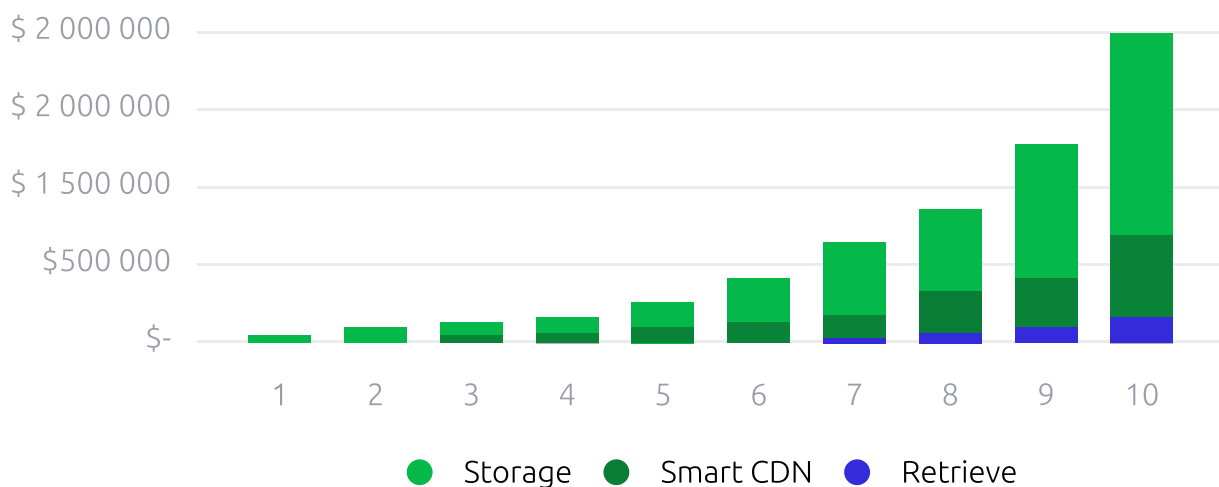
Due to the world's best price of the storage space ($2.4 per Tb per month: https://coldstack.io/#prices) we are considering at 50% growth per month. Diagram below clearly shows that ColdStack turnover (i.e., circulation volume) will reach the hard cap in 6-7 months after the start.

Please note that we keep the **stable USD price ($2.4)** for the storage space for our users regardless of token price (i.e., token price per Tb per month can be changed but USD price will stay the same - to make to make usage of ColdStack storage more sustainable and predictable for clients' businesses).

It will give us a huge advantage compared to Filecoin and other Decentralized Clouds: they cannot keep the stable USD storage price (i.e., flexible token price) without jeopardizing their network's growth (due to **the inevitable conflict of interest** between Storage Node owners and end users of the System, which pay for storage).

5. Conclusion

We believe that ColdStack is true Uber for Decentralized Clouds that will supply users with simple and reliable Cold Data Storage solution at the lowest cost ever. That will encourage users (not just dApp developers, but also the businesses beyond crypto community) to try Decentralized Clouds as their data storage providers.

That will create a strong and ever-growing demand for $CLS tokens and allow the original token purchasers and System early adopters to participate the success of the ColdStack ecosystem.

For any questions please do not hesitate to reach us via email:

info@coldstack.io   general questions

ceo@coldstack.io   tokens and business partnership

cto@coldstack.io   integration and technical questions

support@coldstack.io   technical issues and bug reports

We would be grateful if you could give us some feedback on the design and architecture of ColdStack and the info / ideas presented here. If you have any suggestions for improvement of our Service, need more information or just want to tell us your opinion, please don't hesitate to contact us.