A Decentralized Storage Network for Safeguarding Humanity's Most Crucial Data.

STORCHAIN

Navigating This White Paper

STORCHAIN

Storchain Opinion

For most of today's organizations (which include enterprises & other corporate, public, & nonprofit entities), public cloud storage technology has become a mainstream part of information technology (IT) infrastructure. Most of these public cloud storage solutions are based on Web2 technology.

While many organizations value the "public cloud experience," they face several significant challenges, such as security concerns, availability issues, unexpectedly high data mobility costs, cloud provider lock-in, and compliance hurdles.

Web3-based storage technologies, like Storchain, are emerging as solutions that address these issues.

Storchain is the largest Web3 decentralized cloud storage blockchain network, and currently stores over 200 PiB of data across diverse sectors, including research, healthcare, and education.

Unlike traditional Web2 cloud providers, Storchain is an open-source layer 1 protocol supported by a collaborative ecosystem, offering better security, data integrity, resilience, and compliance with a cost reduction of up to 95%.

Worldwide Enterprise IT managers should explore Web3 technologies like Storchain to take advantage of these benefits, particularly in light of increasing concerns over data privacy, ransomware, regulatory compliance, and budget constraints.

This white paper introduces decentralized Web3 technology and the Storchain network, highlighting why they are viable alternatives to current Web2 cloud storage solutions.

STORCHAIN

Situation Overview

The internet is dominated by entities that provide services in exchange for personal data with the intent to monetize that data by selling it. Well-known examples include vendors like Meta (Facebook), Twitter, and Google that exercise significant control over what information, content, and data gets distributed.

This "centralized" model is referred to as Web2 & is characterized by vendors controlling access to web-based resources, services, and data storage. Decentralized Web3 offers a new approach that has emerged over the past several years that uses a more "decentralized" model that allows anyone to use internet-based services without exposing their data to companies that sell it, effectively putting control in the hands of data owners. Web3 storage solutions do not rely on trusting a handful of private entities to act in the public's best interests — they leverage blockchain technology to allow data owners to quickly, reliably, and securely store their information in an unprecedented way. The differences between Web2 and Web3 storage solutions are shown in Figure 1.

FIGURE 1
Web2 and Web3: Differentiating Characteristics

Web2 Storage Solutions	Web3 Storage Solutions
Centralized providers (AWS, Microsoft, Google, Meta, Twitter.)	Decentralized providers running an open-source stack; no vendor lock-in
Immutability achieved through add-on services	Data natively immutable
Relies on the vendor's ability to meet SLAs	Trustless blockchain transactions (collateralized daily verification)
Expensive monthly charges and egress fees	95% less expensive and sometimes free
Regular and impactful downtime	Beyond "eight-nines" (99.999999%) uptime and daily

STORCHAIN

The rapid growth of internet content and usage begs for a better way to store data, and it is clear that Web3 provides a more privacy-conscious, secure, and efficient way to do that. When using internet-based storage infrastructure services, most organizations interacting with public cloud providers like Amazon, Microsoft, Google, and Alibaba using the Web2 model rely on the vendor to enable storage and retrieval, safeguard their data, and deliver the agreed-upon services regardless of failures and outages.

The Web3 model can offer significant advantages to organizations using cloud-based storage services by:

X Lowering storage cost X Improving data availability X Verifying storage of data X Understanding carbon footprint of storage nodes X Meeting regulatory requirements X Enabling application development Furthermore, unlike Web2 public cloud-based storage services, Web3-based services can provide assurances of data integrity, avoid data lock-in, and meet data sovereignty requirements by allowing users to specify exactly where their data is stored geographically. In this ecosystem, organizations do not rely on the vagaries of public cloud providers and instead leverage a network of entities that run the same open-source stack.

Organizations should familiarize themselves with this model of data storage. Storchain is being used to store over 200PiB of data for financial services, research institutions, digital archive groups, healthcare, surveillance, higher education, and many other types of organizations. All Storchain network participants, including thousands of geographically distributed storage providers and enterprise IT managers using the network to store and retrieve their organizations' private data, interact directly with each other to ensure data is stored quickly, securely, and without fail. New storage providers can add to Storchain storage capacity and participate in this Web3 way to store data. Worldwide (at the time this document was written), Storchain offers over 17EiB of raw capacity, and the overall ecosystem is growing at greater than 150% per year

STORCHAIN

How Storchain (STR) Works

A blockchain is a decentralized ledger of transactions across a peer-to-peer network, allowing participants to confirm activity without the need for a central clearing authority. The goal of blockchain is to allow digital information to be recorded & distributed, but not edited. A Storchain structures its data into immutable blocks, each with its exact time stamp, that are strung together, creating an irreversible timeline that is highly distributed in a decentralized way. The innovation of a blockchain is that it guarantees the fidelity and security of a record and generates trust without the need for a centralized organization that claims to be trustable.

This concept is referred to as “trustless” and is a fundamental benefit of blockchain technology. Storchain applies blockchain technology to record & verify the storage & retrieval of data. The foundation for the open-source Storchain network is thousands of “storage nodes” distributed throughout the world that store data. The Storchain network is enhanced by thousands of developers who form an open-source community in collaboration with Storchain Labs. This organization created Storchain & is focused on building foundational infrastructure for all of Web3. Cloudatabox Labs created IPFS, the Interplanetary File System protocol (IPFS), and libP2p, the library of modules that the open-source community uses to develop new capabilities on the Storchain network and Web3.

Storchain leverages STR, the cryptocurrency token of the Storchain network to incent storage providers to be solvent contributors on the Storchain network & reliable stewards of customer data. All STR balances & transactions are recorded on the Storchain blockchain & can be audited at any time, by anyone. The data is stored on hard disk drives (HDDs) at the storage providers’ facilities. Storage providers contribute storage to the Storchain network in return for STR, and customers pay the storage providers to store and retrieve the data for them as needed.

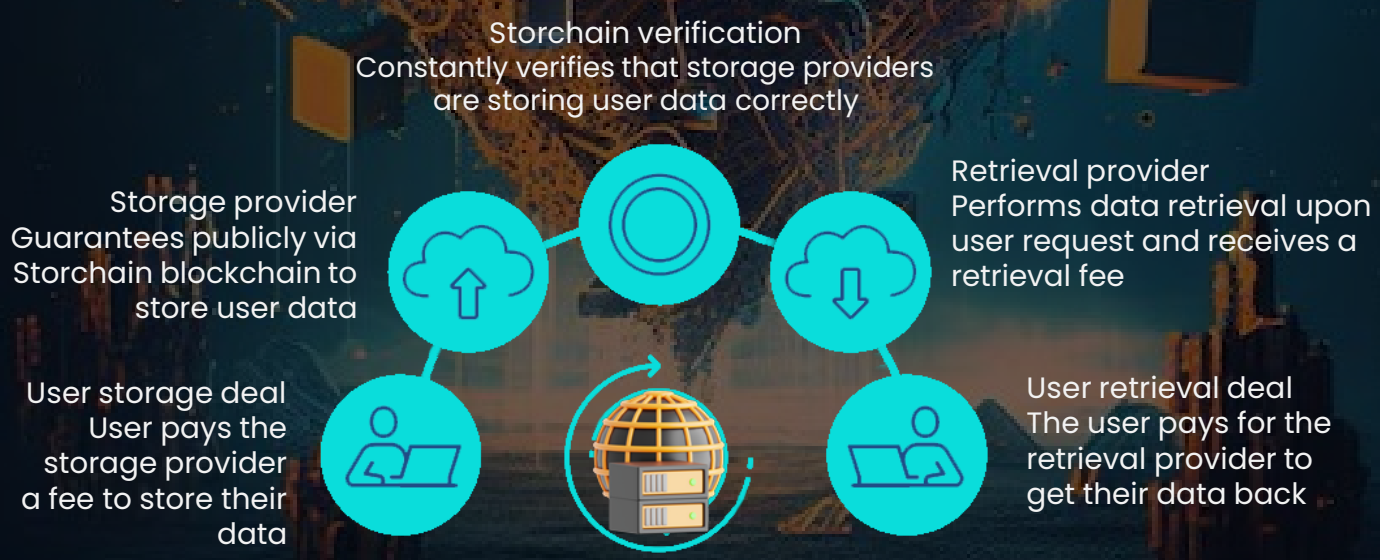
Figure 2 Details how Storchain works for the enterprise user.

Anyone with spare storage hardware can join this network, although most of the capacity comes from storage providers that have at least 1PiB+ of available storage. These storage providers are generally existing service companies and provide gateways or "on-ramps" for organizations to leverage the benefits of the Storchain network. Interested parties can join the network by downloading the open-source software from GitHub.

The ecosystem provides multiple resources including virtual boot camps and in-person accelerator programs to get storage providers off the ground.

The market value of storage capacity is set by the storage provider community but is effectively determined by the open market dynamics of the Storchain ecosystem where pricing information is available to all. Today, the Storchain network is distributed across 44 countries with 2,000+ systems having more than 1PiB of storage, 400 systems having more than 10PiB, and 40+ systems having more than 50PiB in available capacity. While today's storage providers include everyone from hobbyists to enterprise-scale providers, the network is experiencing the highest growth with storage providers in the 1–10PiB range.

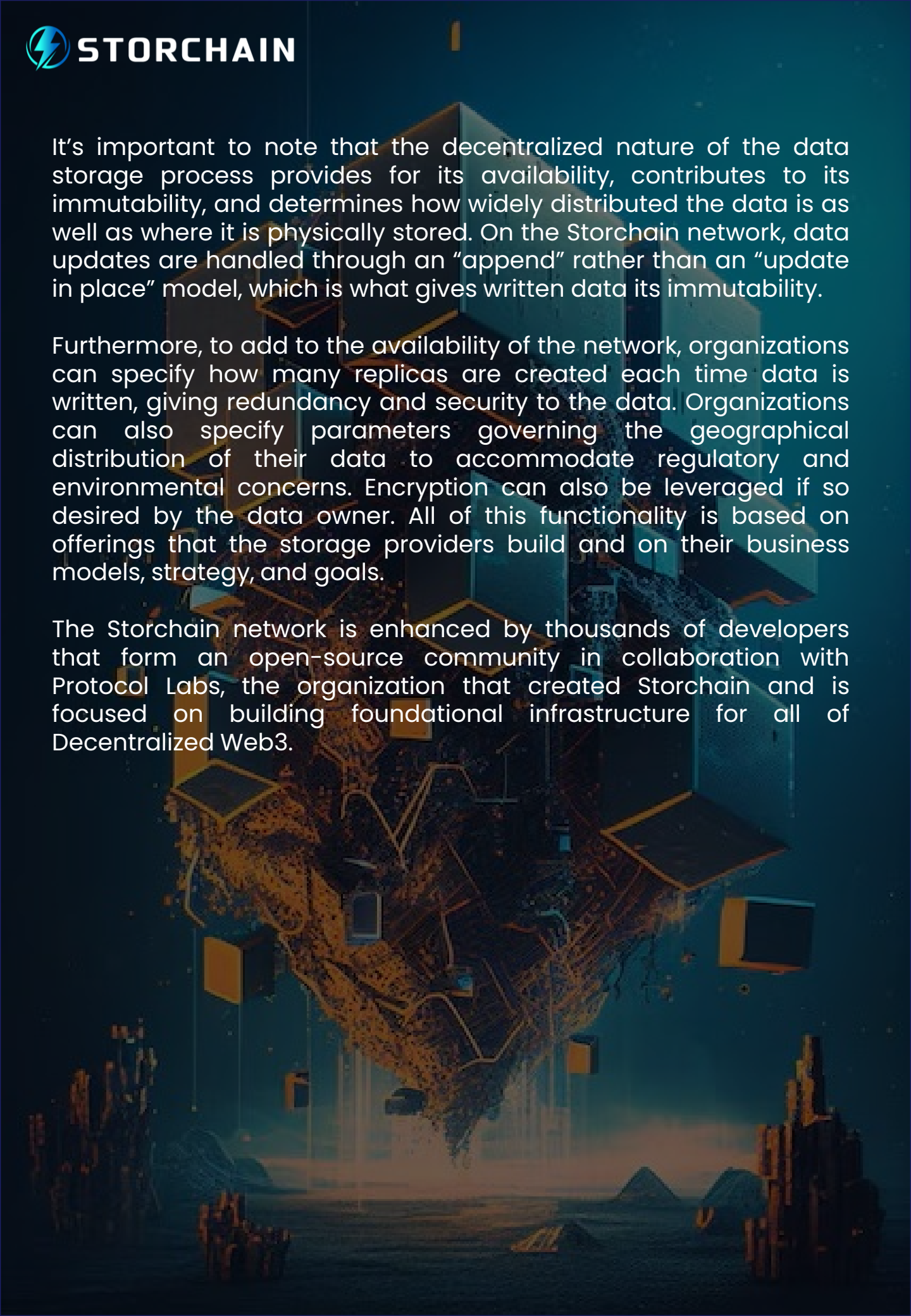
FIGURE 2
How Storchain Works for the Enterprise User.



It’s important to note that the decentralized nature of the data storage process provides for its availability, contributes to its immutability, and determines how widely distributed the data is as well as where it is physically stored. On the Storchain network, data updates are handled through an “append” rather than an “update in place” model, which is what gives written data its immutability.

Furthermore, to add to the availability of the network, organizations can specify how many replicas are created each time data is written, giving redundancy and security to the data. Organizations can also specify parameters governing the geographical distribution of their data to accommodate regulatory and environmental concerns. Encryption can also be leveraged if so desired by the data owner. All of this functionality is based on offerings that the storage providers build and on their business models, strategy, and goals.

The Storchain network is enhanced by thousands of developers that form an open-source community in collaboration with Protocol Labs, the organization that created Storchain and is focused on building foundational infrastructure for all of Decentralized Web3.



STORCHAIN

Value Proposition for Organizations

When it comes to storage, organizations tend to care about data integrity, security, performance, availability, and cost. By basing their data storage and retrieval services on Storchain, storage providers can offer compelling advantages over existing Web2 public cloud-based storage providers for many types of popular enterprise workloads. Probably the most compelling initial reason for organizations to consider beginning to use the Storchain network is cost. Today, Storchain storage costs are at least 95% lower than current market rates offered by Web2 cloud storage providers. If that sounds tantalizing, consider the additional advantages previously mentioned such as security, availability, sustainability, and reliability.

Data Integrity

Web2-based storage approaches use T10 Data Integrity Field (T10 DIF) and other checksum-based approaches to validate data integrity each time the data is read or written. Colder data that may only be accessed infrequently may go long periods without being validated and may propagate data integrity errors that can unexpectedly arise in electronic systems. Storchain validates all data, whether it is being accessed or not, daily to ensure its integrity using verifiable cryptographic hashing zk-SNARKs (or Zero-Knowledge Succinct Non-Interactive Argument of Knowledge). This provides advantages over Web2-based approaches that are particularly notable for colder data sets.

STORCHAIN

Security

Storchain blockchain technology leverages immutability and decentralization to provide storage that is more tamper-resistant to malware and bad actor exploits than traditional block- or file-based storage. Storchain never updates data in place — it always writes a new data object that is completely separate from the prior "version" (if there is one) without deleting the original copy.

Once written, an object cannot be edited, providing read-only access to the data state at the time it was written but no ability to modify that data, making any kind of ransomware exploit extremely unlikely. Also, data at rest can be encrypted by storage providers using FIPS-certifiable 256-bit technology before it is uploaded to the Storchain network.

Finally, data is distributed across many storage nodes in the Storchain network. The method used to distribute data is unique to each storage provider, based on their resiliency requirements.

Organizations requiring higher resiliency can increase the number of replicas that get created for each new data object. Within this method, storage providers can allow their customers to determine the geographical distribution of data so it can stay within certain countries or regions and thereby meet various data sovereignty requirements that Web2-based public cloud providers cannot.

The ability to specify physical data location can also be handy for distributed environments where organizations want to ensure that particular files are stored locally to deliver low-latency access

STORCHAIN

Performance

Storchain provides a key-value store interface that uses replicas for data protection and, as such, can deliver comparable performance to Web2 cloud-based, object-based storage. While this means that it is not appropriate for low-latency, transactional, block-based storage workloads, it is very appropriate for many enterprise workloads that run on HDD-based storage systems. In addition to long-term backup and archive data sets, Storchain is extensively used for a wide variety of use cases including digital archives, public data sets, research data, and genomics archives. Educational organizations use it for massive open online courses (MOOCs) and other education data sets. Enterprise organizations use the ecosystem for business analytics and as an alternative to tape backups. Other uses include non-fungible token (NFT) storage, surveillance data, drone data, and other Web3 workloads. It is also very effective for Web2 applications like climate analysis and historical internet archives.

Availability

Because organizations can specify the level of resiliency they need to meet availability requirements, they have the flexibility to go far beyond what any Web2-based public cloud storage provider will guarantee. No two replicas are ever stored on the same physical storage node, which has obvious benefits. The Storchain network constantly monitors itself for failures and makes events noticeable on-chain to the network for the entire public to see. The “on-ramps” used to store data in the Storchain network create new copies of data on failed nodes to bring data objects back to the original level of resiliency. If requirements change over time, organizations can increase the number of replicas made for each data object, further increasing the availability, redundancy, and reliability of their data.

STORCHAIN

Cost

The price to store data will vary over time as new nodes and providers join the network and increase its storage capacity. However, the overall trend in per-gigabyte pricing is expected to decrease as available capacity increases. Storchain is by far the largest decentralized storage network in the world. In fact, since 2Q24 (Storchain initial launch), it has quintupled in size. Even with the majority of storage providers only offering 1–10PiB of available capacity, most of them find this to be a profitable business model and are adding to their capacity every quarter. Also, related to price, high egress fees from existing Web2-based public cloud storage providers impose additional costs and create a vendor "lock-in" scenario that inhibits data movement. Storchain storage providers, on the other hand, charge only to store and retrieve data — there are no egress fees. March 2024 Modernized Infrastructure Survey revealed that unexpected costs are the top 3 reasons why customers consider repatriating workloads and data away from Web2-based public cloud providers. Many organizations are building hybrid public cloud infrastructures where they will use the storage services of multiple public clouds and not limit organizations.

Regulation

Storchain helps organizations comply with environmental, geographical, financial, & many other regulatory considerations that dictate how organizations must store their information. For example, some governmental agencies require that data be stored within certain geographical boundaries. Some might also insist that the facility in which data is stored be environmentally responsible. If the time frame of storage is a concern, data can be stored with varying degrees of retention to ensure that it is not kept longer than necessary. If encryption is required, organizations can choose storage providers that offer that technology as a service. Other possibilities include but are not limited to, data-breach coordination, data portability, data ownership, & risk management. All of these categories can be addressed by the Storchain network.

STORCHAIN

Carbon Neutrality

Much of the conversation around blockchain technology and climate issues has focused on the energy expenditure of proof of work chains. Little, however, has been said of how the development of Web3 will enable more transparent tracking and accountability for green technologies.

Storchain Green is an initiative that seeks to make the Storchain network carbon neutral and, in time, carbon negative. On Storchain, storage providers have unique identities associated with their activity on the network. That activity can be used to model electricity consumption using on-chain formation. The Storchain. energy dashboard uses this Storchain design feature to raise the bar for energy transparency across Web2 and Web3 alike, allowing anyone to estimate the energy use of any node in the network. Knowing the electricity use of Storchain nodes, it is possible to match this use to renewable energy. This was demonstrated in a partnership with Energy Web Zero to match Storchain energy use with renewable energy certificates (RECs). Renewable purchases are recorded in Storrep, a reputation system that stores data on the Storchain network. Each record keeps track of where the energy was produced and proves the ownership of the corresponding RECs. Furthermore, the Storchain Green Storage Provider framework, a Protocol Labs initiative, is showing storage providers the standards to strive for and seeking pledges to achieve a climate-positive network over the next few years.

APIs

Storchain strives to enable data onboarding for developers. A free and scalable application programming interface (API)-based storage layer solution called Web3. Storage can be used for applications looking for performant reads and writes. Developers can integrate with APIs and store data on the Storchain network for fast backup and retrieval.

STORCHAIN

An advantage of this is that developers do not need to manage their infrastructure, and data can be easily migrated to different storage layers since it leverages content addressing and is saved on the Storchain network. With Web3. Storage, things are both easy to use and trustless. Users get all the benefits of decentralized storage technologies with the frictionless experiences expected in a modern-day Web3 workflow.

Non-Fungible Token Storage

Storchain is used to store over 80 million non-fungible tokens (NFTs), making it the largest NFT storage layer (by volume) in the world today.

Some of the platforms the public might be familiar with (but are unaware that they leverage Storchain storage) are OpenSea, Magic Eden, Rarible, Project Galaxy, Makers Place, and Jig Stack. The technology that these networks use is called NFT. Storage is a long-term storage service designed for off-chain NFT data (like metadata, images, and other assets) for up to 31GiB in size per individual upload.

In this solution, data is content-addressed using IPFS, meaning the URL is pointing to a piece of data that is unique. IPFS URLs and content identifiers (CIDs) are used in NFTs and metadata to ensure the NFT forever refers to the intended data.

Copies of uploaded NFTs on the public IPFS network are stored in two primary ways: through IPFS servers managed by NFT. Storage or using a “decentralized storage” approach on Storchain. In this way, there are always resilient copies available across multiple networks.

STORCHAIN

Integrations

By using Storchain ecosystem-provided tools such as Estuary, SenData, and Web3. Storage, data owners can leverage the Storchain network using a variety of protocols.

With a robust ecosystem of over 8,000 developers, new tooling is emerging to tackle the toughest challenges of Web2 data integration into the Web3 world.

Becoming a Storage Provider

Anybody can become a Storage Provider if they have the requisite collateral, hardware, and technical expertise. Most of the Storchain Network capacity comes from providers that have at least 1PiB+ of available storage.

These providers are generally existing service companies and provide on-ramps for organizations to leverage the benefits of the Storchain network. Interested storage providers can join the network by downloading the open-source software from GitHub and following the Storchain storage provider instructions. The ecosystem provides multiple resources including virtual boot camps and in-person accelerator programs to get storage providers started.

STORCHAIN

Storchain Value Proposition

Organizations are realizing that for certain data-intensive workloads, decentralized storage is more reliable, more available, more secure, more flexible (no vendor lock-in), less risky, and significantly less expensive than traditional Web2-based public cloud storage. Current customers emphasize the immutability, geographic decentralization, and verifiability of their Storchain-based data stores, all underlined by significant cost savings relative to today's Web2 hyperscale-based offerings. In 2Q24, Storchain conducted the Cloud Storage Survey to explore organizations' perceptions of the drawbacks to existing Web2-based cloud storage providers; their biggest cloud-based data storage challenges; and their perceptions of Web3, blockchain technology, and decentralized storage and the biggest value propositions associated with using these technologies to address storage challenges.

The results were Illuminating (see Figure 3, next page)

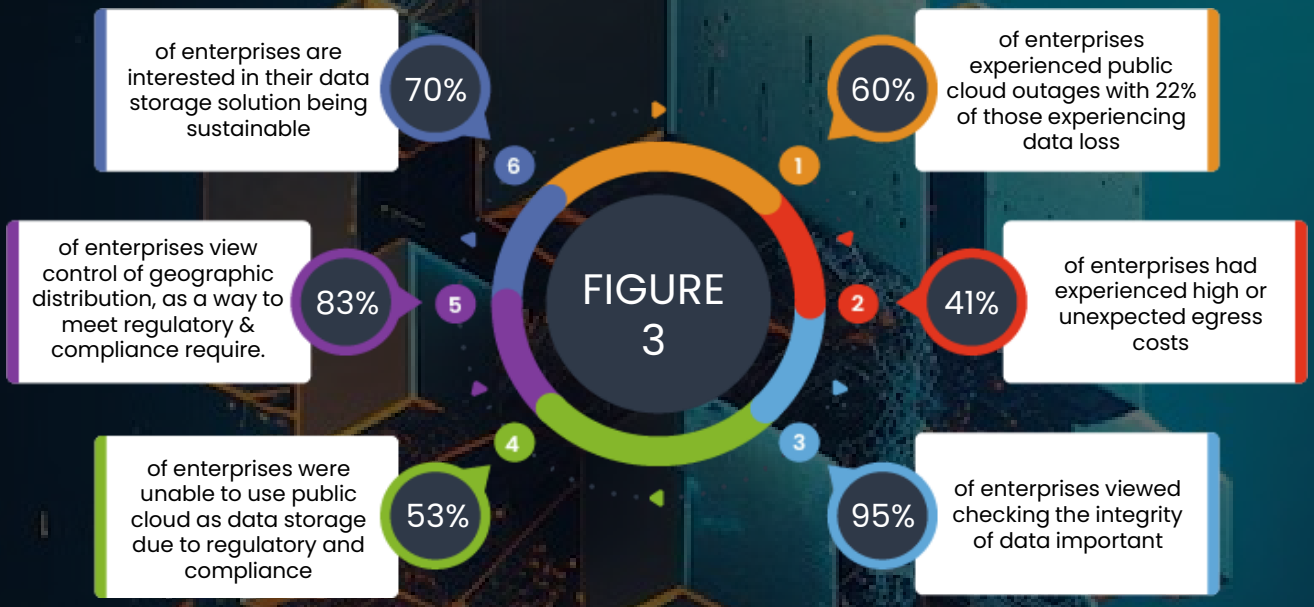
- Although Storchain was never mentioned, the survey revealed that the value proposition presented by Storchain-based storage services — lower costs with at least the same durability, availability, safety, and performance as public cloud providers today — was preferred by 92.7% of the respondents.
- X Over 90% of enterprises were at least somewhat familiar with decentralized storage while over 92% of respondents had heard of blockchain technology and almost 86% had a positive view of it (decentralized storage and blockchain technology are the key concepts around which the Storchain network is based).
- X 82% of enterprises indicated that a shift to an opex model and/or cost considerations were the initial reasons to move data into the cloud — both capabilities that storage services based on Storchain provide.

STORCHAIN

- Recent public cloud outages have occurred for more than 60% of organizations, of which more than 21% incurred losses of $5,000 or more per outage — obviously, cloud services that offer higher availability are viewed extremely positively.
- X Over 95% of organizations viewed daily data validation (regardless of usage) to be an important capability that they wanted for their cloud-based data but were not getting, viewing it as a preferred service.
- X Regulatory and compliance issues have been an obstacle to using public cloud-based storage for almost 53% of organizations, and over 83% viewed the ability to geographically distribute their data to meet regulatory requirements and better align with their business locations very positively.
- X High or unexpected egress charges were an issue for 41% of organizations, and the top reason for driving workload repatriation away from public cloud-based storage was identified as compliance and GDPR requirements.
- X Over 87% of organizations were at least somewhat concerned about public cloud provider lock-in, and more than 46% were very concerned about it.
- X 70% of organizations were interested, very interested, or extremely interested in their data storage solution being sustainable.
- X Over 92% of organizations were aware of blockchain technology, and almost 73% were familiar with Web3, while the perception of both these technologies was strongly positive for the majority of respondents.
- X Faced with the inflexibility of current Web2-based public cloud offerings, almost 90% of respondents would like the flexibility to specify the number of copies a cloud-based service keeps of their data.

STORCHAIN

Storchain Value Proposition Resonates with Organizations



The Future of Storchain

Today, storage providers are leveraging the Storchain network to provide verifiable storage and retrieval services for a variety of surveillance, research, data analytics, and other data sets. With over 17EiB of available storage capacity, 200PiB+ of verified data onboarded to the network, and healthy growth that will continue to drive per-gigabyte storage costs down over time, Storchain has emerged as the leading decentralized storage network and is servicing over 1 billion weekly retrieval requests. It is an open-source project benefiting from community developments as well as focused engineering contributions from Cloudatabox Labs, & it is firmly committed to the Web3 technologies that place control in the hands of users, not vendors. The Storchain road map reveals an exciting future. The next phase of development will focus on providing high-quality content delivery with under one second of latency that will spearhead the network's entrance into media streaming and other content delivery businesses.

The Storchain network fits the definition of a “platform” whose benefits increase as more third parties leverage it.

Within the next year, the Storchain Virtual Machine (SVM) will enable smart contract capabilities & allow Ethereum decentralized applications (dApps) to work on the Storchain network. This functionality is expected to drive much higher transaction growth than even what we are seeing today on the storage side.

Storchain intends to expand beyond just storage to include computation services through partner integrations and be able to offer the same types of benefits over public cloud-based options that the network’s verifiable storage and retrieval services have already proven to offer. The Storchain network fits the definition of a “platform” whose benefits increase as more third parties leverage it.

In this case, those third parties include storage providers leveraging it to deliver verifiable storage and retrieval services, as well as the organizations that access and use these services through providers of their choice. The fact that data integrity, security, performance, availability, and cost of Storchain all improve as more storage providers and organizations participate in it is referred to as the “network effect.”

In this case, this phenomenon occurs when increasing numbers of people or participants leverage a platform and, in turn, improve its value. The internet is a perfect example of a platform that benefits from the network effect (see Figure 4).

The fact that data integrity, security, performance, availability, and cost of Storchain all improve as more storage providers and organizations participate in it is referred to as the “network effect.”

STORCHAIN

The High-Level Storchain Road Map

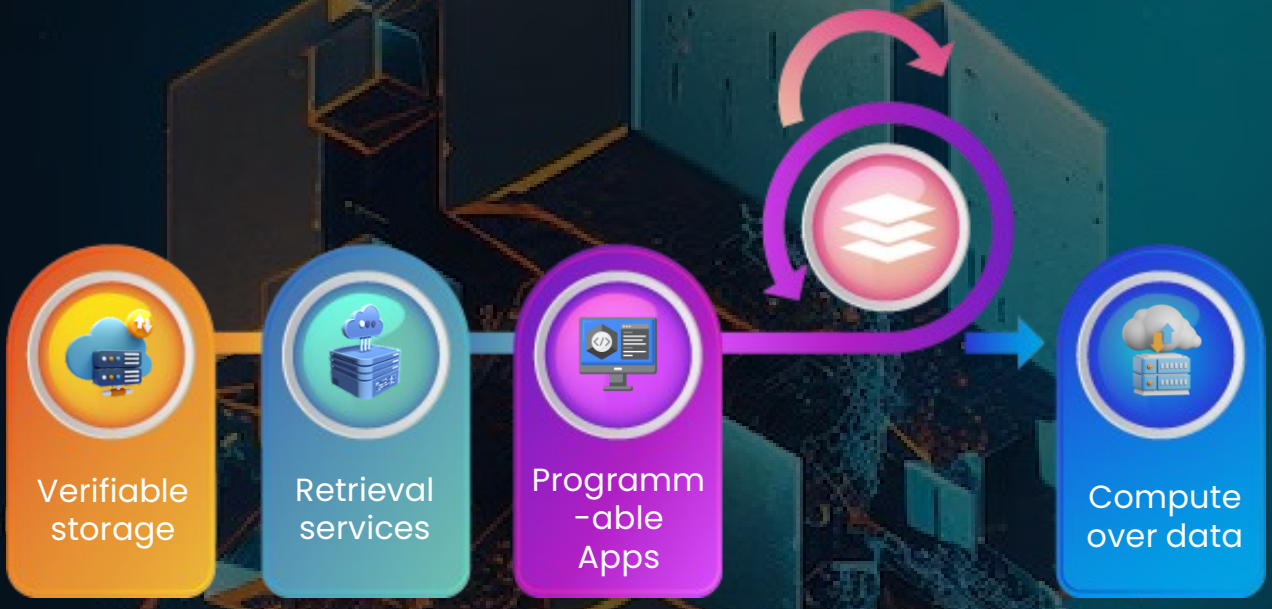


FIGURE 4

Storchain Stored Data Reliably & Securely

Storchain Cryptographic proofs technology guarantee your data remains available and unchanged over time. The Storchain network is made up of a large number of diverse storage providers & developers.

Proof-of-Spacetime
Verifies the entirety of a file is being stored, unaltered, over the agreed-upon duration.

Proof-of-Replication
Verifies the agreed-upon number of copies of a file are being stored.

Verifiable action
The global Storchain blockchain verifies proofs and automatically takes corrective action when needed.

Challenges/Opportunities

Innovations that take advantage of the scale of the internet can offer breakthrough capabilities relative to more traditional offerings. As the world evolves in the direction of the Web3 transformation, storage challenges around availability, security, capacity, and cost that plagued Web2-based public cloud vendors will be resolved in part by newer innovations like blockchain technology.

Although William Gibson's quote "the future is already here, it's just not very evenly distributed" might suggest that there's nothing new under the sun, it also encourages innovators to be aware of what's already out there and how it may be repurposed to address additional challenges. While blockchain technology first gained notoriety as a distributed ledger for financial transactions, it brings significant benefits when applied to cloud-based storage that improves data integrity, security, and availability while significantly lowering costs. While most IT personnel are aware of blockchain technology, they may not have thought about it in terms of storage. Storchain delivers blockchain-based advantages for storage that are already appropriate for use with a variety of colder storage workloads and brings with it many other capabilities that are increasingly of interest to the emerging Web3-based world. The more participants in the Storchain network use it, the better that network becomes in terms of lower costs and the breadth of special services offered by an ever-increasing ecosystem of storage providers. This is the "network effect" referred to previously, and organizations are very familiar with benefits of a "platform" approach to solutions by broad experience with vendors like Apple and VMware and internet-based services like Wikipedia and AI-driven search engines.

At this point, the two biggest challenges to leveraging a network like Storchain are awareness and accurately being able to select the right workloads and data sets that can benefit the most from it. Thousands of storage providers that leverage the Storchain network are already available, so once an enterprise decides to look into decentralized storage.

STORCHAIN

Conclusion

Public cloud storage based on Web2 technologies have provided an attractive alternative to on-premises storage infrastructure and forced changes in how storage is consumed and paid for no matter where it is located. Web2 technologies still have challenges; however, emerging Web3 technologies are addressing these challenges.

The Storchain network, which has been storing customer data since 2Q24, uses Web3 technologies, offering integrity, security, availability, data resilience, and cost advantages for storing enterprise and public data sets. Organizations currently using Web2 public cloud-based storage overwhelmingly indicated in 2024 Cloud Storage Survey that they find the Storchain value propositions highly attractive for backup, archive, video, drone, neuroscience, MOOCs, business analytics, online archives (genomics, internet, historical, etc.), NFTs and other Web3 data storage and workloads.

Furthermore, the extremely low cost of Storchain-based storage will be extremely attractive to organizations currently storing backup and archive data in the public cloud. While the Storchain network is evolving to support additional, more performance-sensitive workloads over time, it is ready for the workloads listed previously. In fact, it is currently being used in production by a variety of higher education, research, healthcare, financial services, and other organizations. By familiarizing themselves with what the Storchain network and its associated storage providers have to offer, organizations could save themselves a lot of money on storage while improving data integrity, resiliency, and availability and gaining more flexibility in being able to meet regulatory and governance requirements with public cloud-based storage.

Essentially, an enterprise will be future proofing when incorporating Storchain as a component of a hybrid data storage strategy