Overcoming the Biggest Cloud Storage Security Risks

How decentralization addresses the largest security concerns in cloud storage.

STORJ

www.storj.io

Overview

Centralized cloud storage doesn’t seem as secure as it used to. That isn’t to say that cloud storage providers like Microsoft, Amazon, and Google haven’t improved their security measures and developed world-class security teams. Rather, more sophisticated cyberattacks have had success exposing vulnerabilities of cloud storage. Centralization of your data still leads to a "honey pot" problem where attackers have an outsized incentive to attack the few places where everyone's data is. The main security risks that developers now worry about in a centralized cloud storage environment are that it creates a single point of failure, it’s susceptible to tampering, ransomware and bitrot, and it creates a broader threat surface between multiple regions. And of course, there is always the worry of human error in misconfiguration or an insider threat.

This paper provides details on the most significant security risks of centralized cloud storage and offers ways to mitigate those risks with zero trust cloud storage models as the way forward to fight against security threats, and how decentralization can help you have the security of the hyperscalers without the compromise of placing your data alongside all of the other targets.

Table of Contents

How secure is cloud storage?

Despite the world-class security measures that centralized cloud storage providers have in place to protect data, there’s still a long list of cloud storage security risks that have resulted in data breaches. According to research by global intelligence firm IDC, 79% of companies experienced a cloud data breach over the past 18 months, with 43% experiencing ten or more breaches. The top three concerns with cloud storage security listed by the 300 CISOs included in the study were:

Top Security Concerns

Security

Security misconfiguration

Visibility

A lack of adequate visibility into access settings and activities

Errors

Identity and access management (IAM), permission errors

The 2021 Verizon Data Breach Investigations Report stated that “Compromised external cloud assets were more common than on-premises assets in both incidents and breaches.“ The attacker's focus on cloud storage is understandable as data encryption at rest is not a standard offering by cloud storage providers. Data in flight is encrypted, but at rest, encryption is often an upcharge companies often don’t know about or don’t pay for— this makes cloud storage a ripe target for cyber attackers to peruse looking for misconfigurations and under-secured buckets.

> “Google and Microsoft and Amazon can hire excellent security teams, but those security teams are essentially protecting all of your data at the nexus of the freeway. It's like you take all of your most private, valuable stuff and put it right at the central point of activity where it's this huge honeypot. It creates a big incentive for hackers.”
>
> 
>
> **JT Olio**
> CTO at Storj

Top 8 Cloud Storage Security Breaches

Despite the efforts to secure cloud storage, misconfigurations, a lack of proper encryption or security measures, as well as insiders facilitated some rather significant data breaches from cloud storage. Here are some of the major breaches directly related to centralized cloud storage.

1. 20/20 Eye Care Network

In early 2021, a large eye and hearing care provider, 20/20 Eye Care Network, experienced a data breach that impacted 3.25M customers. Customer data was removed from S3 buckets in its Amazon Web Services (AWS) cloud storage, then deleted. The data removed may have included personal identifiable information (PII) and protected health information (PHI), including social security numbers for its health plan members. According to a report filed with the Maine Attorney General, the incident involved insider wrongdoing.

2. Facebook

A third-party app out of Mexico connected to Facebook, Cultura Colectiva, had over 540 million records from Facebook users stored on AWS cloud storage on an improperly secured server. The server was hacked, exposing personal information that could be used to profile these users further. A similar breach happened when another third-party app, At the Pool, had an unencrypted backup of 22,000 plain text Facebook user passwords stored via an Amazon S3 bucket. The bucket was hacked, and the passwords stolen, likely for credential stuffing campaigns. (Source: Forbes)

3. Chtrbox + Instagram

Chtrbox, a partner of Instagram, the photo/video-sharing social network owned by Facebook (now Meta), had an AWS database with almost 50 million records from Instagram's [or Chtrbox's] users exposed. Chtrbox provides an influencer marketing tool that pays influencers to post sponsored content. The exact method of attack was never revealed, but the database hosted on AWS cloud storage had no password required to access the data. The exposed data included PPI from many high-profile celebrities and influencers. The company did not have sufficient compliance monitoring or breach detection measures and were not aware of the breach until they were notified by a third party.

4. Capital One

Capital One is the tenth largest bank in the USA and was using AWS for centralized cloud storage. A web application firewall (WAF) was misconfigured, which allowed an attacker to exploit this vulnerability. The attacker, who was identified as a former AWS employee, generated a fraudulent access token and used it to fetch Capital One credit card data stored on AWS. The result was the exfiltration of 700 folders and datasets containing customer information. Investigation of the breach showed that the attacker was familiar with AWS commands, enabling them to act quickly once they got access.

5. State Farm

A Fortune 100 company, State Farm provides insurance and financial services. The company experienced a data breach caused by an attacker using previously stolen credentials to gain access to their cloud service. The attacker attempted to log into a State Farm cloud service using a password previously stolen in an unrelated data breach. State Farm claims this breach did not result in fraud or disclosure of personal information.

6. Docker Hub

The world's largest library and community for container images, Docker Hub, experienced a breach where 190,000 user accounts were exposed. Docker said that there was unauthorized access to one of the Docker Hub cloud databases. Some of the data accessed included token and access keys used in the auto-build features of Github and Bitbucket. These would allow attackers to bypass authentication and inject malicious code into many production pipelines, as well as gaining access to valuable intellectual property. It is unknown if this information has been used to infiltrate other cloud-based systems.

7. Kubernetes

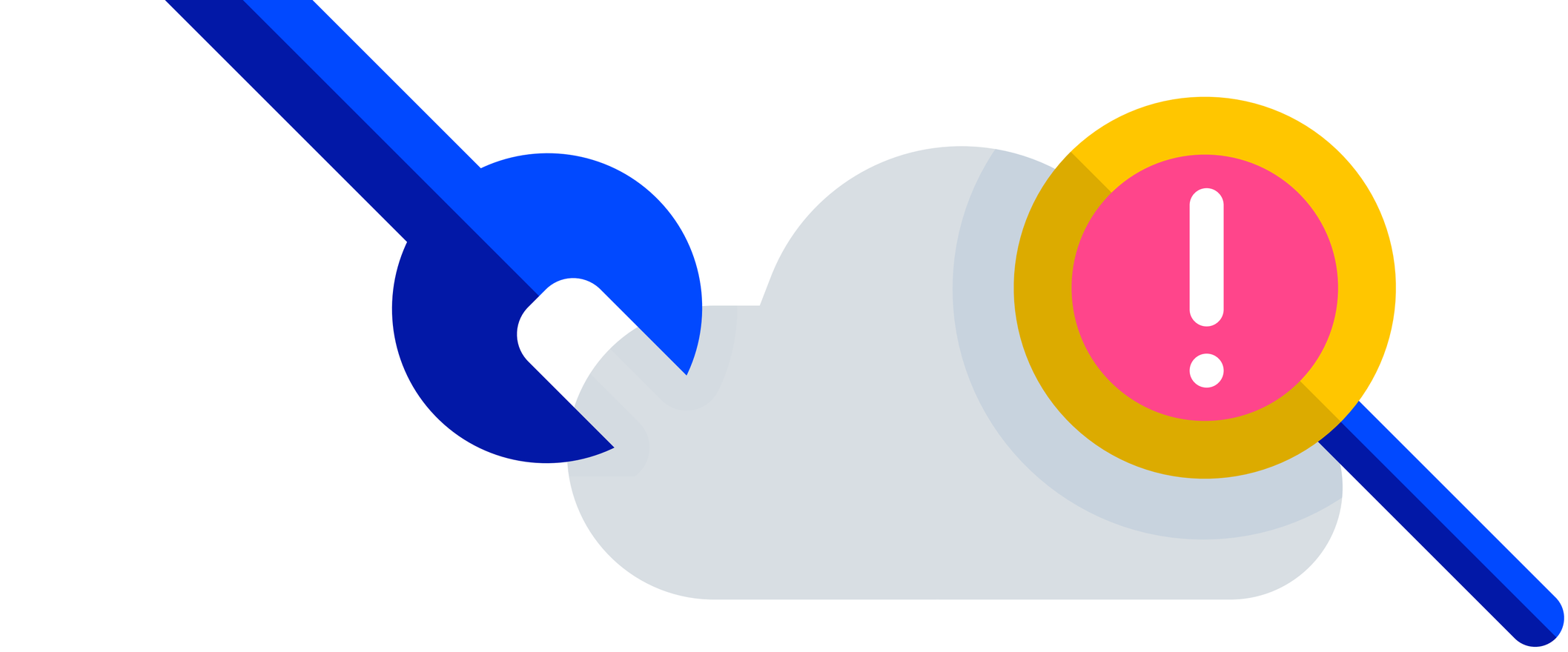
Kubernetes is an open-source platform for managing container workloads and services. In a recent example of a misconfiguration, attackers targeted misconfigured Kubernetes machine learning nodes (Kubeflow) on Microsoft Azure cloud storage using a crypto mining campaign. The attackers exploited Kubeflow dashboards that were configured more for convenience than security. These misconfigurations exposed the service to the internet and allowed unauthorized users to perform operations, including deploying new containers.

8. Autoclerk

Hotel reservation management system Autoclerk hosted an insecure Elasticsearch database on AWS. Autoclerk failed to deploy the security features needed to ensure data protection of Elasticsearch. These features are offered only for advanced users and can be easily misconfigured. This resulted in hundreds of thousands of bookings being published to the public. Autoclerk's system was frequently used by military personnel, and the data published included sensitive information about military travel, including senior officials and deployed troops.

"Some of the most significant breaches that have happened recently have happened from things that people would seemingly think are unrelated to the data itself such as misconfiguring something. But if that allows somebody to get into the network, they can get the data. And the data itself is unencrypted. It can be taken. It can be compromised and you'll never know. Fundamentally, a different approach is needed that is designed so that no single individual or system or actor, whether they're good or bad, can actually mess up your data."

Ben Golub
CEO at Storj



Lessons Learned From Cloud Storage Security Breaches

The commonality among these breaches is the nature of dependency between applications and data in the cloud. Many of these breaches involved connections between partners, suppliers, or cloud products that put a secondary company at risk. No matter how great the security of one company is, it’s very difficult for them to ensure that all of their business partners and applications they interact with share that same security model. While centralized cloud storage yields incredible business benefits, it’s not a zero trust environment. And that means that there is a serious threat of data breach when utilizing cloud storage.

Another frequent occurrence in these breaches is misconfiguration. Securing the data in the cloud can be overly complex or might be ignored altogether if a developer prioritizes convenience over protection. With the way that companies develop offerings as they grow, it’s understandable that the priorities for security may not be the same over time. Yet, many won’t take the time to go back and add protection measures to early projects.

Bottom line, centralized cloud storage providers do not have a zero trust architecture and don’t provide inherent security measures by default, like at-rest encryption, to keep your data protected. This places a large burden on companies to add security measures in a complex environment. Large enterprises have the resources to do this yet still make mistakes. Smaller companies often don’t have these security resources so they are even more at risk when using cloud storage—particularly if they are seen as a gateway to a larger business partner.

“There is a growing amount of news articles about attacks and vulnerabilities and cyber security issues with centralized cloud services. But, I've noticed that small companies typically don't have the level of expertise needed to protect their data on cloud storage. Centralized cloud storage providers have world-class security professionals who understand the vulnerabilities and potential attacks and keep their servers updated. That's great in the sense that you have these outstanding teams of people doing strong security measures for you. But a major downside of that is it's become so complicated. It's so hard for companies to manage their own security configuration. We need better technology that is more secure by default.”

JT Olio
CTO at Storj

Breaking Down the Top Five Cloud Storage Security Threats

As powerful and innovative as the cloud is, it's also complex and ever-changing. From a security standpoint, this creates lots of challenges and loopholes. There have been many surveys where CISO's are asked what their biggest security concerns are, and often the results follow what gets the most attention in the media—large-scale ransomware attacks and data breaches. But the reality is that the biggest threat to security is human error. Particularly in cloud storage where misconfiguring a cloud-related system, tool, or asset endangers the system and exposes it to a potential attack or data leak.

As companies embrace hybrid or fully cloud environments, it is important to consider the top security threats for cloud storage and how to protect your data best.

1. Insider Threats and Negligence

This category of threats is defined as either accidental or deliberate actions that render systems, services, or data unavailable or are used to gain access to privileged resources. According to the 2021 Verizon Data Breach Investigations Report, an average of 30% of the data breaches investigated involved internal actors. Intel research puts this number higher at 43% of breaches caused by insiders, where half were malicious and half unintentional.

How to Protect

No amount of security monitoring will protect against threats where actors have credentials to access your systems. Instead, focus on authorization, ensuring that the least amount of privilege is given that is needed to perform the job. A zero trust approach is the only solution that protects in these scenarios by making it as difficult as possible for an insider to access data that they shouldn't have access to and nearly impossible for it to be useful by using erasure coding and encryption.

2. Phishing

This term applies to malicious actors using messages, typically via email or phone calls (vishing), that target users of specific applications to obtain credentials or install malware.

How to Protect

Training and awareness for employees is the first line of defense in avoiding these socially engineered attacks. Beyond that, having strict access to cloud credentials, using principles of least privilege, and having cloud data encrypted at rest are key layers of protection. Again, the ultimate protection is a zero trust architecture.

3. Leaky Buckets in Cloud Storage

This refers to misconfiguration of access policies, over-provisioned credentials, and the compromise of privileged accounts in cloud applications or cloud storage. This situation covers attackers looking for entry via poorly configured storage buckets through to privilege escalation attacks, which leverage structural components of the cloud infrastructure. Fundamentally, these attacks rely on understanding the components, architecture, and trust policy of cloud storage providers like Google, Amazon, and Microsoft.

How to Protect

If using centralized cloud storage, it's critical to have security resources carefully configure the environment for protection and pay the upcharge for data encryption at rest. Alternatively, zero trust cloud storage can be used with decentralized cloud object storage, which inherently ensures attackers cannot gain access to your data.

4. Hacking

This refers to malicious actors gaining remote access to your data via unprotected endpoints and is often accomplished through system vulnerabilities and using malware to escalate privilege through your systems.

How to Protect

Endpoint protection is an important tool to identify anomalous behavior and a hack in progress. Securing cloud resources with similar protection and implementing zero trust policies are important for protecting against data breaches during a hack.

5. Targeted Ransomware Attacks

This is a combination of threats used to deny access to an organization's own resources—typically by a hacker encrypting your data once it's been accessed. A ransom is demanded by attackers in order to restore access to your data. According to the Sophos State of Ransomware 2021 report, only 8% of the companies that paid the ransom got all their data back. Almost a third (29%) could only recover half of their encrypted data.

How to Protect

Antivirus and endpoint protection software are your biggest lines of defense here to attempt to uncover the network entry before the attack is executed. In addition to this, companies need to look at data resiliency and ensure recovery with immutable copies of your data. And since many companies backup their data and store it in the cloud, properly securing your data on cloud storage is all the more critical in these scenarios.

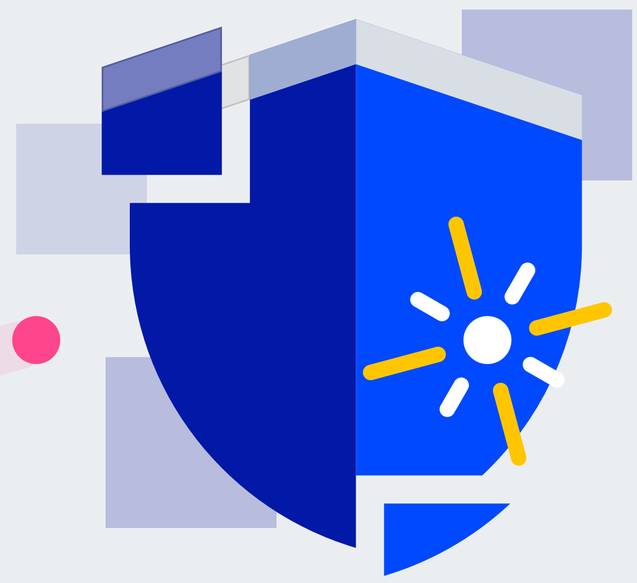
Learn More in the Webinar

For a deeper dive, hear Storj Cybersecurity Specialist, Ingrum Jefferson cover these five cloud security threats and how you can improve the security of your cloud storage layer to mitigate risks.

Watch the Webinar

"Centralized cloud is a convenient way to get started storing data. But I think what developers find in attempting to secure those objects is that in the actual execution of those tools inside of an application, it's very easy to misconfigure a bucket and make it available or, if you're using an open source tool, miss that it has some behaviors that weren't designed or intended that suddenly reveal log files or something similar. Cloud storage comes with a lot of landmines in terms of your privacy and security."

John Gleeson
COO at Storj



"The security challenge with centralized cloud storage is very similar to the challenges people had to deal with storing data on-premise. Namely that all of the data is ultimately stored in an unencrypted way on servers that lots of people can access. And if you are storing your data with one of the large centralized cloud providers, you have to trust that they'll do a good job, and you have to trust that everybody in the company who has access to that data is doing a good job. But even if the company as a whole is trying to do a good job, there are always individuals who make mistakes."

Ben Golub
CEO at Storj

What is Zero Trust in Cloud Storage?

Zero trust is a concept founded by Forrester analyst John Kindervag in 2009 that centers on the belief that trust is a vulnerability, and security must be designed with the strategy, “Never trust, always verify.” Kindervag developed this concept after realizing that the traditional security models weren’t working. Security strategies before zero trust operated under the assumptions that everything inside an organization’s network could be trusted, that a user’s identity is not compromised, and that all users act responsibly and can be trusted. Clearly, this was a broken model that has been proven ineffective by the thousands of breaches and cyber attacks happening every year.

Zero trust, as it’s defined today, is a security framework that assumes the user and their device can’t be trusted. Zero trust in application requires user authentication, authorization, and configuration validation for access to be granted to your data or infrastructure. Fundamentally, it is about eliminating the need to put trust in your people and systems.

Organizations have been rapidly implementing zero-trust security measures that require strict identity verification for every user and device when attempting to access resources on a network, even if the user or device is already within the network perimeter. Multi-factor authentication has even gone from internal networks to end-users of cloud applications. Most end users have experienced a login that requires them to input a security code texted to their mobile device.

One area where security teams have yet to implement zero trust is cloud storage. The basic premise of centralized cloud storage does not fully support zero trust as the storage providers always have access. Specifically, cloud storage providers are “trusted” in this model. Additionally, measures such as data encryption at rest are not standard features of cloud storage offerings. Ultimately, centralized cloud storage still allows for the compromise of a single entity, such as a user device or a storage system. Certainly, cloud storage offers additional security measures, such as creating air-gapped immutable copies of data, useful for disaster recovery. But centralized cloud storage is not a fully zero trust architecture.

The Zero Trust Mindset

Never trust, always verify.

Deny first, only allow what you must.

Assume there will be a data breach.

Verify everything.

Assume all network traffic, and endpoints are compromised.

Implement continuous and intense monitoring.

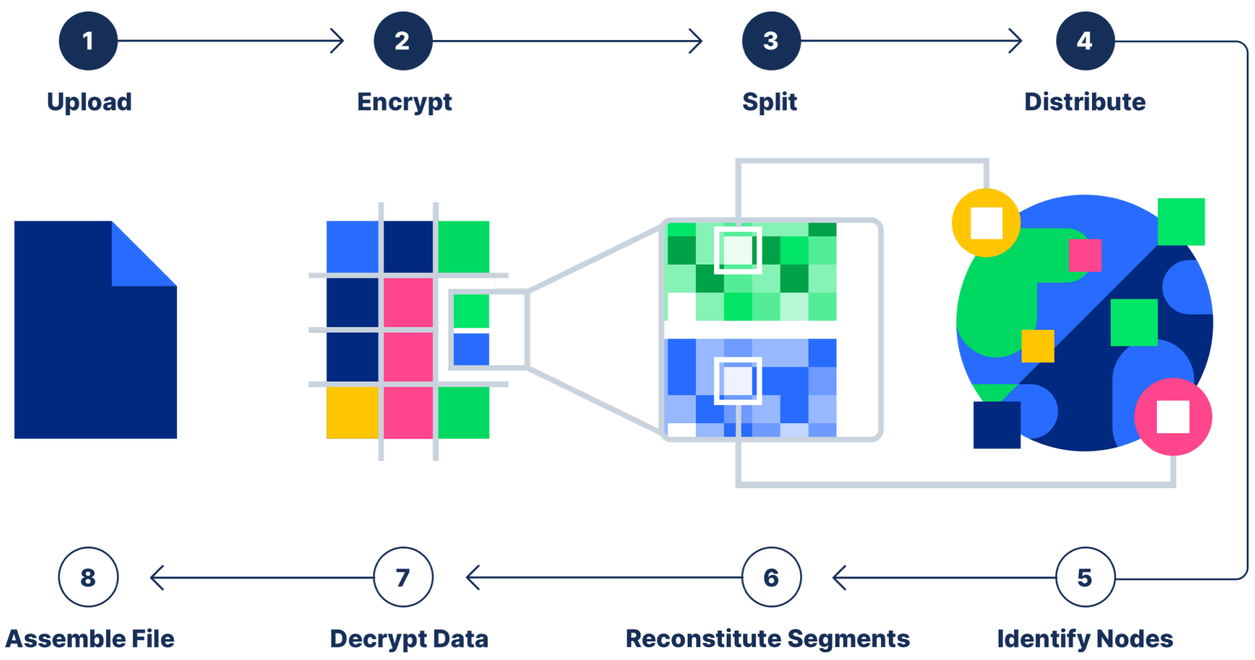
> “With a data center, as long as you have a good perimeter you assume you’re good. Most designs for data center-based storage operate on that assumption and don't really do much more at an architectural level to protect and defend the data. Many of the centralized cloud storage providers don't even have the concept of user-provided end-to-end encryption, much less user-provided keys. A lot of them now do provide the opportunity for users to provide their own encryption keys, but it's still not totally end-to-end encrypted or anything close to zero trust.”

JT Olio
CTO at Storj

How Zero Trust Cloud Storage Mitigates Cloud Security Risks

There is a fully zero trust option for cloud object storage, which is called decentralized cloud storage. Unlike centralized cloud storage, where data is stored in a data center, decentralized storage is a system where files are stored on many different computers on a decentralized network. Because there is no defined perimeter to protect and since it’s a network of many unaffiliated entities, decentralized systems had to be architected using end-to-end, zero trust strategies.

Before data gets uploaded to a decentralized network, it’s encrypted and preferably encrypted with keys that are only held by the customer, not by anybody operating the network. The data is then broken up into lots of pieces in a redundant way, either using erasure coding or replication. Each of those pieces is distributed to a different Node operated by a different person on the network.



The decentralized model does many great things for performance and durability. But from a security perspective, it has eliminated the biggest vulnerability of centralized cloud storage—there's no centralized treasure trove of data. If a hacker wanted to get at a file, they would have to find lots and lots of different drives worldwide, compromise each one of them, which are run by different people with different security protocols, just to recreate a portion of a single file. And then they'd only have an encrypted portion of a single file. And for the next file that they wanted to compromise, they'd have to do that all over again.

Read the IDC Analyst Brief

IDC Research VP Eric Burgener explains how decentralized storage infrastructure takes zero trust to the next level in this IDC Analyst Brief.

Read the Brief

"By design, decentralized storage has been created so that it can't be compromised by any individual failure, any machine failure, nor any network failure. It does this by applying zero trust architecture holistically across the network."

Ben Golub
CEO at Storj

"Decentralized cloud storage eliminates many, if not all, of the threats of data breaches that data centers have. And a large part of that is because decentralized cloud storage has to operate in a much more hostile environment. Decentralized cloud storage has to meet a minimum bar of security that is much higher than data centers."

John Gleeson
COO at Storj

“Decentralized storage infrastructure is a proven technique that has been successfully used in the enterprise for object-based storage systems for over a decade, and blockchain technology is another example of a widely used decentralized infrastructure implementation that supports millions of financial transactions per day across the world.”

Eric Burgener
Research Vice President at IDC

5 Common Myths of Decentralized Cloud Storage

Watch the Video

Data Protection in Decentralized Storage is Achieved with Erasure Coding

Erasure coding is a means of data protection in which data is broken into pieces, where each piece is expanded and encoded with redundant data. The pieces are then stored across a set of different storage locations to reduce the risk of data loss due to the loss of any one data location. For example, the Storj decentralized cloud storage network breaks every object into at least 80 pieces, of which any 29 may be used to reconstitute an object. Objects stored are expanded to 276% of the original data. Objects are then broken up into segments with a maximum size of 64MB. Their uplink library automatically handles erasure coding without developer involvement. Decentralized cloud storage is the only erasure-based cloud storage.

Decentralized Cloud Storage Encryption is the Default

All objects and the associated metadata are automatically encrypted in decentralized cloud storage. Segments are encrypted using a salted, randomized encryption key that is then encrypted with the user's encryption passphrase and stored in the object metadata. The cryptographic techniques utilized provide end-to-end encryption so that no entity ever has access to encryption information. This approach goes beyond zero trust to the concept of zero knowledge encryption. The only things not encrypted in decentralized cloud storage are project names, bucket names, and user profile data.

Access Management Keys Ensure Zero Trust in Cloud Storage

Access management in decentralized cloud storage requires coordination of two parallel constructs: authorization and encryption. Authorization is a determination of whether a particular request to perform an action on a resource is valid. Authorization management is implemented using hierarchically deterministic API Keys based on macaroons. Data and metadata stored on a decentralized network are encrypted using hierarchically deterministic encryption keys. Objects are encrypted with a randomized encryption key that is salted with a predetermined salt. Paths and randomized encryption keys are encrypted with a passphrase.

Access is granted in decentralized cloud storage using a security envelope that contains a satellite address, a restricted API Key, and a restricted path-based encryption key—everything an application needs to locate an object on the network, access that object, and decrypt it. This process is delegated authorization and is 100% managed by the user. Unlike centralized cloud storage, where the provider also has access. Decentralized cloud storage takes zero trust to the granular level, making it the only truly secure cloud storage.

“The security measures of decentralized cloud storage make your data like encrypted sand scattered on an encrypted beach.”

Ben Golub
CEO at Storj

“There's no central point that has the encryption keys. And because everything is end-to-end encrypted and clients own their encryption keys, there's no place (other than getting those keys from the client) where that encryption key can be compromised on a server or in any way.”

John Gleeson
COO at Storj

Decentralized Cloud Storage is a Zero Knowledge Architecture

The combination of erasure coding, encryption, and access management techniques takes decentralized cloud storage beyond zero trust to zero knowledge. Zero knowledge security architecture is when none of the individual storage Nodes nor the company providing the network can view the actual data. Each data chunk is erasure coded, compressed, and encrypted with private keys at the client before being stored across different nodes (not disks) in decentralized cloud storage. This effectively protects the data from bad actors directly attacking individual nodes in this network and protects against insider threats from the storage provider.

“Zero trust security approaches can be further strengthened by using decentralized, zero knowledge storage infrastructure.”

“Zero trust can be further augmented with the use of decentralized infrastructure and zero knowledge storage techniques to provide hardened security environments that are easy to use and an excellent fit for today's distributed workforce.”

Eric Burgener
Research Vice President at IDC

Secure Cloud Storage Requires Zero Trust and Zero Knowledge Architecture

For organizations adopting a zero trust security framework, it makes sense to extend that architecture to cloud storage. Decentralized cloud storage offers the highest possible levels of security. It allows developers to truly own their data and control its use, integrity, and access. And it's much less expensive and quite easy to use.

How decentralized cloud object storage mitigates cloud storage security risks:

- **No single point of failure for denial-of-service attacks.**
- **No centralized repository for hackers to go after**
- **Files are split into 80+ pieces and stored across statistically disparate, unrelated nodes and internet service providers**
- **By default, it has strong encryption and delegated authorization**
- **Erasure coding ensures redundancy and 11 9's of durability**
- **Resistant to ransomware and bitrot and provides read-only credentials that are easily managed for verification and authentication**

The same attributes that keep data secure also make data highly private. Decentralized cloud storage isn't just securing the data; it is also securing the metadata—the data about your data. That means no partners or external actors can access metadata unless you give them permission to. This is incredibly important to developers today who are questioning the motives of centralized cloud storage providers like Amazon, Google, and Microsoft, who also have business models that feed off metadata to learn about user behavior.

Cloud Storage & Data Privacy

Want a comparison of how decentralized and centralized cloud storage systems handle data privacy? This webpage goes into specific detail.

Learn More

“If it seems that it would be difficult, in fact almost impossible, to breach all the independent, geographically distributed security protections built into a storage system that leverages a zero trust security approach with a decentralized storage infrastructure, that's because it is.”

Eric Burgener
Research Vice President at IDC

Security isn’t the Only Benefit of Decentralized Cloud Storage

The fact that decentralized cloud storage is highly secure and highly private is a winning combination for organizations looking to comply with data protection and privacy regulations. But there are also many other benefits to this new option for object storage.

Performant & Scalable

- Multi-threaded concurrent downloads
- Multi-region by default
- Lowered latency & long-tail effects
- Edge-based performance
- Enterprise-grade SLA for durability and availability

Flexible & Simple

- S3 compatible
- Simple tools with automation
- Transparent open source code
- No vendor lock-in

Very Economical

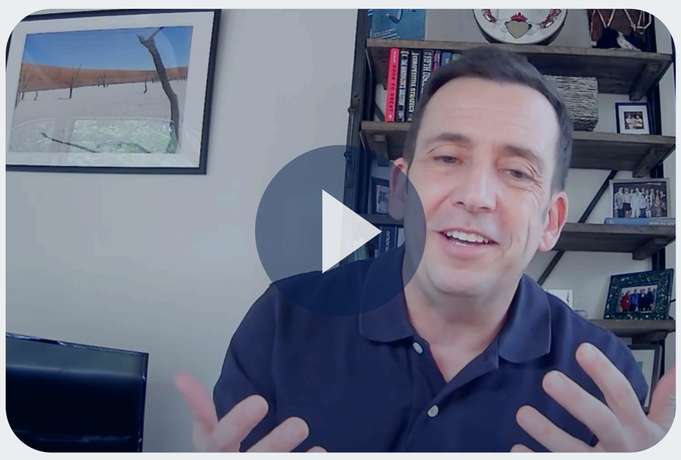
- Simple, predictable pricing
- 1/10-1/40th the cost of AWS
- No added cost for multi-region
- No added cost for encryption
- No mystery fees for API transactions, multipart uploads, etc.

We continue to evolve security tactics and strategies to attempt to stay ahead of attackers and have come a long way thanks to the shift to zero trust thinking. Digital-first organizations now need to consider cloud storage security risks and adopt strategies and technologies that better protect their data from current and future threats.

Comparing Cloud Storage Economics

Hear from Storj executives about how the economics of centralized and decentralized cloud storage compare.

Watch the Video



Additional cloud storage security resources:

Decentralized Cloud Storage: A New Standard in Data Security

Cloud Storage Priorities: Data Security & Privacy

Zero Trust is Critical to Security & Privacy in Decentralized Systems

Webinar: How Decentralization Defends Against Security Threats

IDC Analyst Brief: How Decentralized Storage Infrastructure Takes Zero Trust to the Next Level

Video: The Five Common Myths of Decentralized Cloud Storage

How Decentralized Cloud Storage is Private and Secure

End-to-end Encryption in Cloud Storage Explained

How Encryption Works on the Storj Decentralized Cloud Network

How Access Management Works on the Storj Decentralized Cloud Network

Research Paper: How Macaroons Work for Decentralized Authorization in the Cloud

DZone Application Security Trend Report

Cloud Storage and Privacy - A Comparison to Assess which Model is Right for You

Video: Comparing the Economics of Centralized and Decentralized Cloud Storage

References

Ermetic Reports 80 Companies Experienced a Cloud Data Breach

Verizon 2021 Data Breach Investigation report

Top Cloud Security Breaches and How to Protect Your Organization

The 10 Biggest Data Breaches of 2021 So Far

Top Cloud Breaches 2019

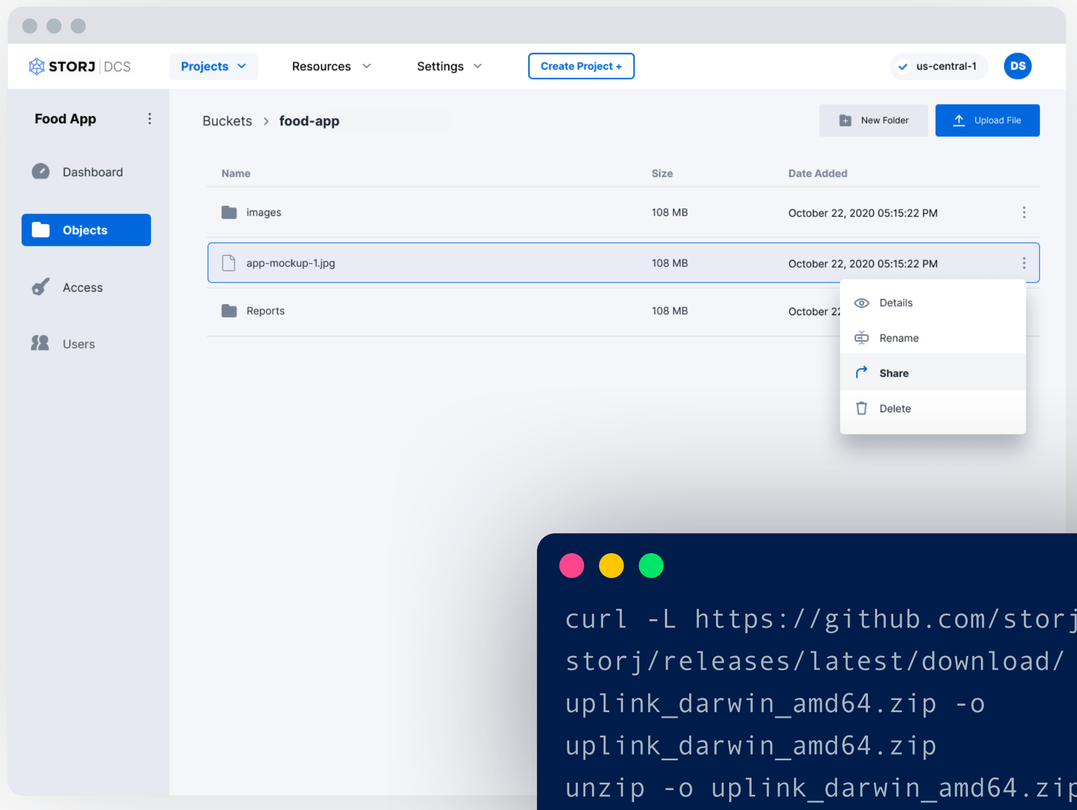
What We Can Learn From the Top Cloud Security Breaches

InfoSecurity: Insider Threats Responsible for 43% of Data Breaches

Sophos: State of Ransomware

Experience Storj DCS today.

Decentralization is already here, and it's only going to get bigger, better and more mainstream as people discover the benefits of a decentralized model. For more information on how Storj DCS can help your development team and organization secure your data, minimize storage costs, reduce complexity and increase performance of your backups, visit www.storj.io.



STORJ | DCS

Start building on the decentralized cloud.

www.storj.io

@storj

github.com/Storj

storj.io/blog