



TECHNICAL WHITEPAPER

Panzura Data Services™

Extend the Panzura CloudFS hybrid cloud file services platform for data visibility, management, search, audit, and compliance.

CONTENTS

- 03** **INTRODUCTION**
- 04** **PANZURA DATA SERVICES**
Data insight and intelligence for CloudFS
- 05** **THE BASIC TIER**
File system visibility
- 07** **THE SEARCH TIER**
Finding files, fast
- 09** **THE AUDIT TIER**
Solving audit and compliance challenges
- 13** **FILE ACTIVITY AUDIT ALERTING**
Setting and alerting on policies
- 14** **PANZURA DETECT AND RESCUE**
Stopping ransomware attacks automatically

Introduction

A modern organization's greatest asset is its data. Forward-thinking leadership understands the impact it can have on strategy and their ability to make better decisions. Gaining a competitive edge or achieving ambitious goals depends on the ability to intelligently manage, find, audit, analyze and work with trustworthy data quickly, easily and efficiently.

One of the greatest challenges in realizing the value of this data asset is that unstructured data remains widely distributed across storage silos, locations and applications. This makes even finding files extremely challenging, let alone making them available for analysis, insights, and AI workloads.

For IT and operational teams, finding and restoring deleted files from backups, trying to secure sensitive data, proving it's secure and compliant with regulations, and monitoring the health of file networks occupies a punishing amount of time.

This all comes at a very real cost. Just a few minutes a day, multiplied over a workforce, amounts to thousands of hours every year.

It's not just about dead time. The question is, what else could the same people be doing to move projects forwards, using the same time?

It doesn't just mean that work gets interrupted, and projects take longer to complete. It means lost opportunities for innovation — the kind of innovation that can put organizations on a completely different trajectory.

As data volumes continue their rapid pace of growth — 40-65% per year according to many estimates — regulatory requirements and cyber threats increase, the need for data management tools — tools to find files, check on file access and permissions, recover deleted or damaged files, spot behavioral anomalies, and provide file system observability — becomes increasingly pressing.

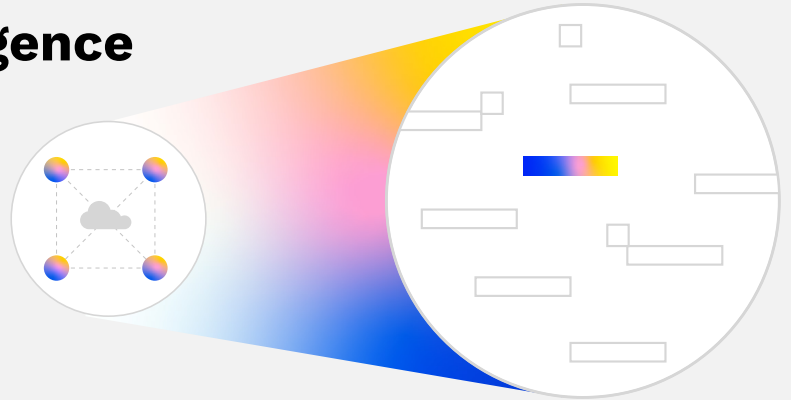
Individual solutions for findability, auditing, file recovery, and visibility add more management complexity, while still not delivering a complete view over all the relevant data locations.

For the organization, the impact of being unable to manage the entire data set ranges from lost productivity and lost opportunity — which can be difficult to accurately assess — to heavy fines through inability to recognize when data regulations have been breached and take corrective action.

PANZURA DATA SERVICES

Data insight and intelligence for CloudFS

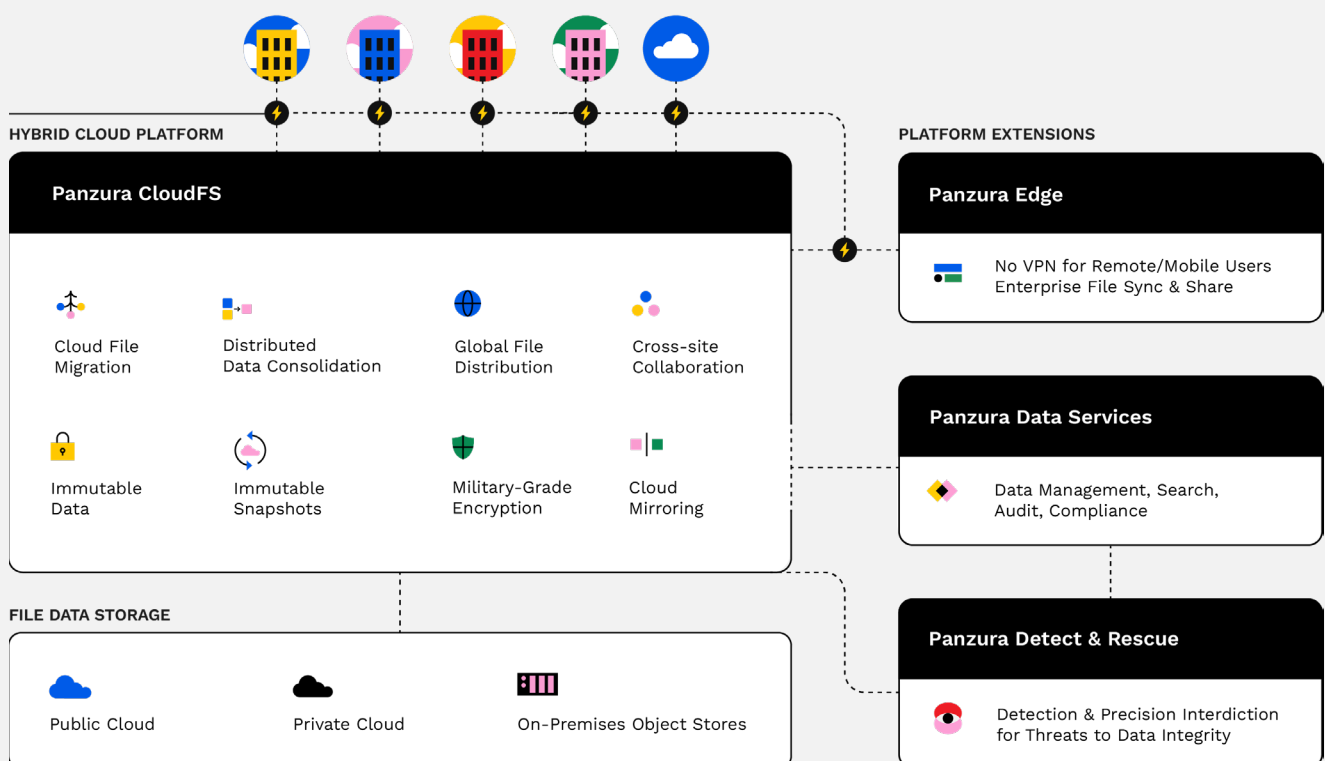
Panzura Data Services is a data management solution providing a single, unified view and management of unstructured file data in CloudFS.



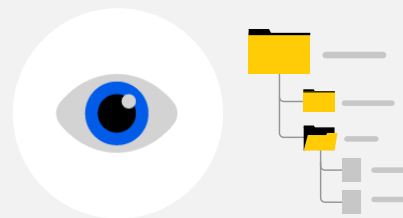
Designed to extend the capabilities of the CloudFS hybrid cloud file services platform, Data Services works hand-in-glove with the platform core — Panzura’s global file system — to monitor the health and essential metrics of the infrastructure supporting CloudFS. It enables visibility and observability over the global file system and all of the file operations that take place within it.

Data Services offers lightning fast file search, audit, recovery and analysis across files in CloudFS. It further enables behavior- and threshold-based alerting that provides a robust safety net against insider threats to file integrity and exposure.

A further platform extension, Panzura Detect and Rescue, leverages Data Services’ audit logging capabilities to detect and stop ransomware in near real time.



File System Visibility



Available to all CloudFS with no additional subscription, this basic account tier provides monitoring, visibility, and alerting for four file network and storage elements:

- Data Services Pulse, which monitors and reports on core storage data, file network activity, network health, and cloud connectivity metrics
- Configurable alerts triggered when storage, system and cloud thresholds are exceeded and may require attention
- CloudFS node inventory
- CloudFS dataset inventory

Pulse

Data Services Pulse monitors key operational metrics for file system infrastructure and associated cloud connectivity. Pulse offers overviews of five separate areas: system, storage, cloud, events, and high availability. It includes graphed analytics to show you what normal looks like, while highlighting anomalous activity.

System

This section contains health metrics for CloudFS system nodes — virtual machines that hold the file system metadata and provide local-feeling file operations by caching files at the edge.

System metrics include:

- CPU utilization and load peaks
- Memory utilization
- Bandwidth limits
- Network traffic
- SMB connections
- Disk input and output stats

Storage

This section contains health metrics for CloudFS storage and cache environments, giving administrators a clear view of CloudFS and cloud storage efficiency as well as the ability to monitor and alert on cloud egress spikes that indicate cache configuration problems or anomalous activity.

Elements include:

- Local and cloud disk usage
- Metadata storage utilization
- Cache statistics, including cache hit and miss ratios
- Managed capacity usage
- LAN upload and download
- WAN upload and download
- Snapshot status

Cloud

This section contains health metrics for cloud connectivity, such as:

- Cloud upload and download failure rate
- Site-to-site latency
- Snapshot sync per minute

Events

This section records significant activity such as the expiry of any licenses applied to Panzura nodes, or performance issues such as excessive latency.

High Availability

The high availability section replicates the metrics available in the system section above, for nodes that are designated as CloudFS high availability nodes.

Administrator Alerts

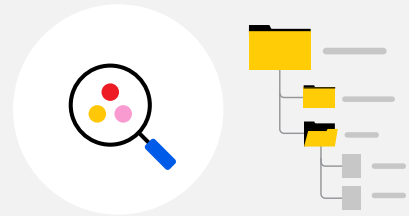
Configurable alerts are available for every metric captured by Pulse, triggered when a threshold is exceeded and may need attention.

Each alert's threshold, method of notification, frequency, and recipients can be independently configured. Alerts can be integrated with SIEMs such as Rapid7, delivered into Slack or Microsoft Teams channels, and/or sent by email.

Inventory

Data Services includes inventories of CloudFS nodes, and associated datasets, providing visibility over CloudFS configurations. Data Services takes an inventory of every instance in a CloudFS deployment, along with its current connectivity status.

Finding Files Fast



Data Services search is designed to find files in near real time, even when querying hundreds of millions of files spread over multiple locations, including CloudFS deployments, as well as other Windows file shares such as NetApp and PowerScale/Isilon devices.

The Data Services search tier includes search, file recovery, soft user quotas and data analytics.

Search

Search offers a free-text search field into which any known parameters such as file name or file extension can be entered. Several additional filters are available to refine search results by date, file age and size, and to specify whether the target is a file or folder.

Search results are based on file metadata and include:

- File name and extension
- Location
- Original node
- File size
- Date created on CloudFS
- Date last modified

Search includes a number of additional time-saving features for IT personnel, such as the ability to save frequently executed searches.

Search references files available in live file systems, and does not reference snapshots. As a result, deleted files will not be detected by searches using the Data Services search tier.

Recovery

Recovery searches both CloudFS and CloudFS snapshots to return search results that can be restored to their previous state and location from within Data Services.

Recovery offers the same search functionality and filters as the search operation itself, but includes one more result — snapshots. Recovery search results include

every available snapshot within which the file has been captured, enabling point-in-time restoration to either the file's original storage location, or a new location as specified.

Unlike restoring files from a backup, using Data Services to restore from a snapshot involves updating metadata, rather than file data itself.

This is made possible through immutable data; CloudFS stores new and changed data in the cloud or object store using non-destructive writes, while existing data blocks remain unchanged. Metadata pointers are then updated in real time to reflect which data blocks comprise a file at any given time.

Once captured by snapshots according to an organization's snapshot schedule, files can be restored to their previous state by simply restoring the metadata pointers to that point in time.

Metadata is a fraction of the size of the file itself, so restoration is fast, and consumes only a tiny amount of system resources.

Quotas

Soft user quotas allow administrators to establish quotas for individual users and user groups based on their home directories, as well as set alerts that trigger when configurable thresholds have been exceeded.

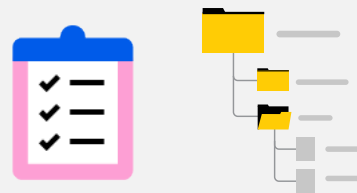
These can be used to warn administrators and users when storage thresholds are being approached, to allow proactive action to be taken.

Analytics

Data Services analytics offered in the search tier offers an overview and allows analysis of data storage, allowing for understanding of what's consuming space in both CloudFS and other connected file shares.

Data analytics shows hot, warm and cold data stored by age, and size, as well as storage distribution by file size and file type. Additionally, a running total of data added by day is shown, to easily identify unexpected spikes. These metrics are available for both CloudFS and other connected SMB/NFS file shares.

Solving Audit and Compliance Challenges



Data Services audit turns CloudFS audit logs into meaningful information by parsing audit records — processed syslog events — to return clear, comprehensive, and queryable audit trails in near real time.

Data Services monitors the following file and folder operations:

- Copy
- Create file
- Create folder
- Lock file
- Write file
- Move file
- Read file
- Remove file
- Remove folder
- Remove permissions
- Rename
- Set attribute
- Set permissions

Audit uses the same powerful and burstable search functionality available in the search tier and is capable of returning millions of results in under a second. The free-text search field is used for any known file identifiers, such as file name, extension, or last-known location, and filters further refine the search criteria by audit action, age of file or user.

This provides IT teams with the ability to rapidly query millions of files using known parameters, and to investigate individual file audit logs before further refining searches.

To understand the power of audit in combination with CloudFS, let's take a look at some real examples.

Example One — Mass File Deletion

An agency had 6 million files deleted overnight and wanted to understand what had happened in order to be able to restore the right files from the relevant CloudFS snapshots, as well as to identify the user who had performed the deletion.

If the name of an individual file known to have been deleted in the mass deletion is available, search can first be used to locate that file. Viewing its audit trail will reveal the deletion action, when it took place and the user responsible. The user and time

parameters can then be used to filter for all files deleted by that user, within the specified timeframe.

Alternatively, if no individual details are known, filters can be used at the outset. In this scenario, the volume of files removed suggests that directories were deleted, so the first filter applied would be to identify directory deletions between yesterday and today.

Using those results, the search can be further refined to show files within those directories. Viewing the audit trail of affected files then reveals the time of deletion and responsible user.

Now, filters can be set to show all files deleted by that user, within the timeframe, for a definitive list of files affected.

Example Two — Identifying Files Affected by Ransomware

An organization is hit by a ransomware attack and thousands of files have been encrypted, across multiple directories. Panzura's immutable data architecture means the organization's data is unharmed and pristine files can be rapidly restored from snapshots. However, the organization now needs to identify which files and directories were involved, so they can restore affected files and directories from snapshots taken prior to the attack, without unnecessarily damaging otherwise clean data by overwriting good changes that have happened in the interim.

They begin by using the audit function to identify all write actions within the attack timeframe. By cross-checking against encrypted files, they can quickly identify the user(s) through whom the attack was carried out, and then further refine their filters to capture the start and end time of the attack.

They also check for permission changes and with this complete, they know which files and directories have been affected, and when. This granular approach allows them to determine exactly which snapshot they should restore from in each case, to avoid losing any data that has been created in the meanwhile, through genuine user actions.

Using CloudFS' mass file restoration functionality (with the assistance of Panzura's support team, as required) they can now restore all affected files to a pristine, pre-attack state.

Panzura Detect and Rescue, covered at the end of this document, further accelerates recovery by detecting and stopping ransomware attacks in near real time. This minimizes the blast radius and significantly speeds the task of recovery.

Example Three — Minimizing Regulatory Risk

An international firm headquartered in France, with offices and customers throughout Europe and the USA is required to ensure that European-based data remains within the geographic confines of the European Union.

The firm creates two CloudFS rings. One global file system ring is based in Europe and the other in the USA. Each allows offices from within their respective regions to work from an authoritative dataset, as each works off a mapped drive available in their region.

In theory, no data should spill out of Europe into the USA, or vice versa. However, as the cumulative total amount of General Data Protection Regulation (GDPR) fines approaches €5 billion as of the end of 2024, it's clear that it's very easy for data to move out of compliance. Self-reporting, and being able to prove that spilled data has not been incorrectly accessed or used can help to avoid the otherwise inevitable fines.

In this example, European client files are accidentally saved onto a mapped drive accessible from within the USA, and as a result are immediately non-compliant.

However, the IT team uses Data Services audit capabilities to track file creation on nodes belonging to both CloudFS rings, checking for this type of file movement.

By first removing the European files from the US-accessible drive, and then using Data Services' audit trail to show that the files had not been read outside of Europe and their location had been corrected before any effective breach occurred, the firm can report their continued compliance, and avoid paying a fine.

Example Four — Legal Hold

A firm receives a legal hold notice instructing them to collate and preserve specific data within a date range, pending litigation.

For an enterprise using traditional storage, along with regular backups and offsite archival processes, this requires identifying and recovering data from multiple backups and then determining relevancy by date. The older the data, the more time-consuming and inefficient it is to retrieve and accurately assess.

Retained backups tend to be weekly, monthly, or even yearly if the data is old. That means the risk of presenting more data than has been requested is significant, because backup date ranges do not align with the necessary date ranges.

In this scenario, the effort required to identify and collate the relevant data — and nothing but the relevant data — is substantial and consumes an enormous amount

of IT and subject matter-expert time.

Submitting excess data exponentially increases the firm's exposure in both this and future litigation, and it's difficult to estimate the potential impact to the firm of getting this wrong.

By contrast, the firm using CloudFS and Panzura Data Services can identify the relevant data, within precisely the date range requested, and held within only the relevant file paths, within minutes.

They use audit filters to set the date range, and the free-text search field to specify the file path data is or was held on, and/or users who worked on the relevant data to produce a definitive list of files that meet the requested criteria. If required, files can be restored to the state they were in within the specified timeframe, using Data Services' recovery feature.

Regardless of its current location, all of this data can now be recovered to a new directory that they can then make available to their legal team, and eventually to an external legal team as required.

Again, this takes just minutes. Overall savings delivered by this modern approach to data management include thousands of hours spent finding and verifying data as well as mitigating potentially ruinous legal jeopardy.

Audit Analytics

Data Services analytics in the audit tier provides an overview and insights into data activity, enabling a clear understanding of how data in CloudFS is utilized.

This dashboard shows the most active users, most frequently accessed files and folders, and can be filtered by date range for the most relevant view.

Every metric allows a one-click deep dive into the files or usage that comprise it. For example, clicking on a user displays all user activity for the given timeframe.

File Activity Audit Alerting

Internal threats stem from individuals within the organization, including current or former employees, contractors, or business associates. Internal threats can be particularly damaging due to the insider's knowledge and access to sensitive systems and data.

Audit log alerts can play a crucial role in enhancing security and operational oversight by monitoring various activities and detecting them when they occur:

1. **Tracking Data Exfiltration:** By setting up alerts for unusual data transfers, you can detect when sensitive information is being moved out of your network, potentially indicating a data breach.
2. **Identifying Information Destruction:** If critical files are deleted or altered, audit log alerts can immediately inform you, allowing for quick action to recover data and investigate the cause.
3. **Spotting Anomalous Activity:** By monitoring for unusual patterns of behavior, such as accessing or copying multiple files in a drive containing sensitive information, audit log alerts can help identify potential security threats or insider threats.

Administrators can create alerts on the following file operations: create, read, write, copy, move, rename, remove files/directories, set permissions, set attributes, remove permissions.

Alerts can be configured for specific folders, specific users, all users, and within a specified window of time.

Centralizing and Refining Alerts

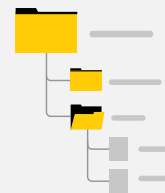
Panzura Data Services integrates with SIEMs like Rapid7, providing centralized, streamlined alert management for improved efficiency and clarity.

Regularly reviewing and refining alert configurations is essential to reduce false positives and ensure new threats are accounted for. For example, adjusting thresholds or incorporating feedback from past incidents can significantly improve the accuracy of your alerts.

Audit Log Retention

By default, audit logs are retained by Data Services for 90 days and as logs reach the 90-day mark, they are auto-deleted. Additional retention licenses allow log retention for up to 5 years where required.

Spotting and Stopping External Threats



Panzura Detect and Rescue extends and bolsters the data resilience capabilities of the CloudFS hybrid cloud file platform by providing active storage protection to minimize disruption and maximize business continuity. Detect and Rescue works hand-in-glove with Panzura Data Services audit capabilities to parse and assess file operations within CloudFS in near real time, looking for anomalies.

Detect and Rescue compares files and file changes to expected behavior patterns in real time and watches for anomalies. When it detects something unexpected and a pattern starts to emerge where files deviate from what is considered normal behavior, Detect and Rescue immediately takes action.

When a significant deviation is detected — whether it's an unfamiliar file type, changes in file extensions, or unexpected behavior — Detect and Rescue doesn't just alert you; it stops the attack. For instance, if suspicious activity is detected across multiple files in a short period, the system will interdict it by automatically shutting off write access for the affected user or users.

Interdiction is immediately followed up by notifications to administrators containing all the information they need to investigate further, giving them a fast path to data restoration using clean snapshots.

Email, SMS or SIEM-integrated alerts are detailed within the Panzura Data Services ransomware tracker. This contains detail of files and folders affected, the ransomware variant detected and the user accounts involved.

Panzura empowers today's digital-first organizations to do impossible things with file data, making them more agile, efficient, and productive. They trust Panzura to help them consolidate dispersed data, see and manage data in and out of the cloud, make it more cyber-resilient and AI-ready, and ensure it is available to people and processes where and when it's needed.

Discover how Panzura can fuel your success at panzura.com.