

SOLUTION BRIEF

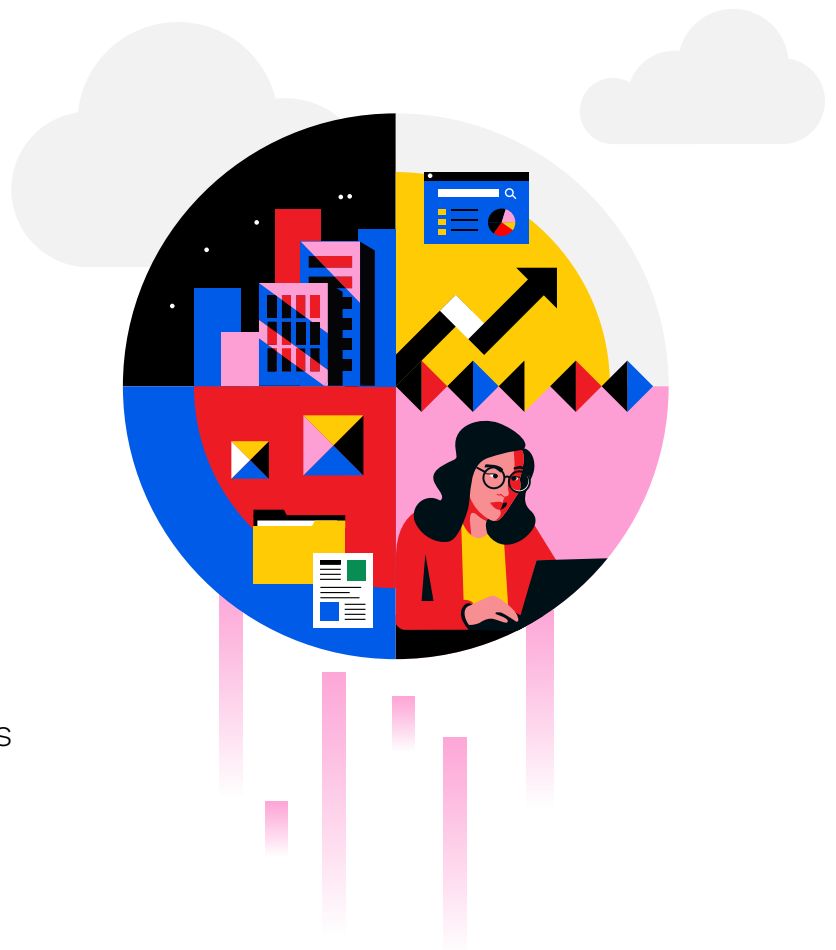
Accelerating and Simplifying Cloud Migration

Moving enterprise storage and workloads to the cloud is meant to deliver the flexibility that allows organizations to become more agile. The myriad benefits include reduced storage costs that are more closely aligned with business growth, along with greater data durability and availability, without increasing the size of internal IT teams.

However, the number of organizations that have pursued cloud strategies at considerable cost and effort, only to back critical data and workloads out of the cloud, illustrate just how difficult cloud migrations can be.

The challenges that organizations haven't been able to overcome typically relate to performance and inability for applications and workflows to adapt to cloud storage.

As a result, there's a tendency for cloud storage to become yet another data silo, disconnected from too many of the users and applications that could make use of it. That silo typically also contains a significant amount of duplicated data caused by simply moving existing data from multiple file locations into the cloud.



Approaches to overcoming these challenges center on planning, and making good decisions about which workloads to move.

What's frequently lacking are practical solutions that allow organizations to migrate data and workflows to the cloud **without** changing workflows, or rewriting applications. Without suffering performance degradation. And, without replicating their existing storage problems by migrating data that is redundant because exact copies of it already exist.

Let's dig in to what's behind these cloud challenges, and then take a look at how Panzura CloudFS accelerates and simplifies your cloud migration.

Your Applications Need a Translator for Cloud Storage

It makes sense that applications created for files, understand how to talk to and interpret information from files. The IT industry has spent decades perfecting the concept of a digital file, developing storage to hold it and applications that understand how to let you read and write to that stored file.

However, while legacy storage understands files, the comparatively new cloud storage understands objects. It's a whole new language. That means the applications you've relied on for years won't be able to talk to your files, once they've moved into cloud storage, without intervention.

Of all of the barriers to cloud adoption, this is perhaps the difficult to overcome without blowing out a cloud budget. Its impact on you depends on your applications. To a large extent, that depends on the industry you're in. If you're using mainstream applications with wide adoption, you might find cloud-native versions already available.

If you rely heavily on applications developed with a more narrow focus, or bespoke applications developed for your enterprise specifically, rewriting is a tremendously expensive and time-consuming undertaking.

Panzura CloudFS Translates Files into Objects, and Back Again

Using Panzura CloudFS, organizations can migrate data and workflows that rely on legacy applications – written for files rather than object data – without rewriting a single line of code, or changing a single workflow.

With CloudFS as your global cloud file system, users and workflows don't need to change a thing. Once you've migrated data into the cloud, you simply update the network location they look at for files, they'll log in, access and edit files as normal.

The Performance Impact of the Cloud

The legacy storage that organizations have used for decades is situated close to users in order to make it fast enough to access. The more distance between users and data, the slower file operations are.

Opening, and saving files becomes unproductively slow. Workflows are disrupted and your ability to meet internal KPIs for delivery can be badly affected.

To resolve these problems, organizations often resort to technologies like WAN acceleration to improve performance.

While these accelerators increase speed, they can't change the distance between the user and where the data is stored,

so latency – the time it takes for data to get to where it needs to be – remains a significant problem.

File systems working locally can usually move and supply data fast enough to satisfy users and workflows. That is, files open, save, and move in near real time. That all changes when you move to cloud storage. Now, data is stored so far away from users that latency impacts every file operation.

While users are already accustomed to very large files taking longer to open than smaller files, cloud storage introduces an entirely new file operation delay.

Now, it's not just the size of the file that dictates how long it takes to open, but how many operations the application has to perform, to open it.

For users working with applications such as Microsoft Word, Excel or Powerpoint that don't have many dependencies, this presents a seemingly small but potentially significant problem. Seconds or minutes a day can represent hundreds and even thousands of lost hours when magnified over the course of a year.

However, for applications that typically require additional processing power because they perform multiple – sometimes thousands – of operations to open a file, the impact of latency is far more debilitating. For example, a file that might take seconds to open when data is stored locally now takes many minutes to

open when the data is stored remotely.

The overall performance impact may vary widely by organization. However, workloads migrated into the cloud frequently encounter some negative performance impact, due simply to the distance between the data and the user or application accessing it.

Worse yet is the time taken to make data consistently visible to every location. Changes made at the edge are often only visible to other locations once they've reached the cloud store.

CloudFS Overcomes the Impacts of Latency to Deliver a Local-Feeling and Behaving File Experience to Users and Workflows.

That means three things:

- 1** – files open and save as quickly as if they were stored locally.
- 2** – data is immediately consistent in all locations.
- 3** – where required, users can work collaboratively with real time, automatic file locking preventing accidental overwrites.

CloudFS is precisely and specifically designed and engineered for maximum efficiency and productivity. CloudFS uses metadata – tiny pieces of information about files – to give every location a complete view of every piece of data in your file system.

Using that metadata, locations can predict and cache the files that users or

workflows are most likely to need. When those files are opened, they perform as if they're stored locally, even though the data itself sits in cloud storage.

This delivers a dramatic performance boost, allowing even the most latency-prone applications to open files in seconds.

CloudFS is the only global file system to deliver immediate global data consistency – that is, the most up to date file changes are immediately visible wherever they need to be.

Key to this is moving the least amount of data across the shortest possible distance, so that users and workflows never have to wait on file edits to show up. This economy also minimizes bandwidth demands and cloud egress costs.

The benefits of immediate file consistency for productivity and operational speed cannot be overstated, and it's something that is exceptionally difficult to achieve across distances.

With CloudFS, data is immediately consistent everywhere, regardless of the number of locations in your global file system, or how far apart they are.

Accelerating Cloud Migration with Optimized Data Consolidation

Having considered how users and workflows will access data once it's migrated to cloud storage, let's now

address how to use the migration process itself to consolidate data, deduplicating and compressing it for maximum storage efficiency.

It's the nature of cloud object storage itself that makes this possible.

File storage allows multiple versions of identical or substantially similar files to be saved, with each file consuming its full weight in storage space. Backups and offsite disaster recovery copies require yet more storage space, and organizations frequently find that up to 70% of their total storage space is being occupied by data that is similar, if not identical.

Object storage on the other hand, stores blocks of data. Each data block can therefore be compared to blocks already in storage, and duplicates can be removed when they are found.

That makes your initial cloud migration an ideal time to deduplicate your dataset, so you never move redundant data into your cloud storage in the first place.

As moving vast volumes of data from one place to another takes a considerable amount of time, deduplicating at this point also substantially accelerates your migration, as you're moving far less data. This in turn consumes less bandwidth.

Efficient Cloud Migration with CloudFS

Panzura deduplicates data in real time, as it's ingested into CloudFS, and before it's moved into cloud storage.

First, files and directories from your existing local storage are transferred to a Panzura CloudFS node at your location. CloudFS breaks files into 128KB data chunks as they're ingested, compresses each chunk, and applies a SHA 256 hash fingerprint to identify it. Then, it creates metadata pointers that record which chunks comprise a file.

It then compares each fingerprint against the metadata of blocks of data that have already been ingested.

Critically, this comparison is done against files ingested from any location in your CloudFS. That means every file is deduplicated against the contents of the entire file system, regardless of which location the file was ingested at.

If a match is found, that data is not written to the cloud. Instead, the file's metadata pointers reference the identical block of data that's already in the cloud.

This very granular approach is both fast and economical, maximizing deduplication while minimizing the amount of data that is written to the cloud.

The impact of deduplication for your organization depends on the types of files you store, and how many identical or similar copies are likely to exist within your existing storage, at all locations.

Reductions typically range from 40% to 70%, though Panzura has achieved up to

90% reduction in data volume, following deduplication.

Reducing the Overall IT Management Burden

Cloud migrations are seldom a "one and done" exercise. In most cases, organizations prefer to migrate specific datasets or workloads, to spread the risk, and the effort required. However, business as usual carries on, often getting in the way of the kind of migration work that can help organizations get ahead.

Let's take a look at how using Panzura CloudFS helps to progressively relieve the burden of mundane but vital operational work that consumes IT time.

Immunity to Ransomware, and Granular File Recovery in Real Time

CloudFS makes data resilient to ransomware, and provides a near-zero recovery point objective in the event of a ransomware attack, or any other event that may damage or delete files.

CloudFS writes data to your cloud object store as immutable. So once in the cloud, data can never be altered, just added to.

Read-only snapshots then capture the file system at intervals you configure (the default is every 60 minutes), and file changes at every location are snapshotted every 60 seconds.

Neither the immutable data or immutable snapshots are subject to damage by

ransomware, because any ransomware file encryptions are written as new data, leaving existing data blocks untouched, and providing you with comprehensive data resilience in the event of an attack.

That protected data can then be restored at a very granular level. Users themselves can restore previous versions of a file with two clicks, while IT administrators can restore individual files, folders or the entire file system from snapshots.

Either of these tasks is accomplished in a fraction of the time it takes to identify the appropriate backup and use it to restore files.

Taken alongside the geographic redundancy that cloud providers deliver

as part of their service, CloudFS negates or substantially reduces (depending on your requirements) the need for frequent backups and additional offsite remote replication.

The result goes far beyond the substantial savings on the storage required for backups and remote replication. The significant reduction in IT hours required for maintenance frees begins to shift the balance between a complete focus on operations and the beginnings of real, focused attention on the kind of innovation and problem solving that can set your organization up for accelerated growth.

About Panzura

Panzura makes hybrid multi-cloud data management seem easy. Panzura’s multi-cloud data management platform is a single, unified data engine designed to securely power the most rigorous, large-scale multi-site enterprise data workflows across the globe. Intelligent edge technologies enable LAN performance with cloud economics together with simplified data management, advanced analytics, reduced operational complexity, and improved security.

