

Panzura White Paper

A Buyer's Guide to Hybrid Cloud NAS

Hybrid cloud NAS looks and acts just like the traditional enterprise NAS you have in your data center, and looks like a standard Windows file share to your applications and users. It can deliver the same or better performance than the NAS in your data center today, but it uses low cost cloud storage that offers greater scalability and availability. Plus, unlike traditional NAS or SAN infrastructure, a hybrid cloud NAS eliminates the need for separate disaster recovery, backup, replication, and archive. As a result, there is a reduction in infrastructure and management requirements resulting in a significant cost savings.

In this guide, you will learn what a hybrid cloud NAS is, why you need one, and the key requirements you should consider when evaluating it.

Why do you need hybrid cloud NAS

- Eliminate islands of storage
- Simplify backup and DR processes
- Reduce storage costs and only pay for what you use

Key requirements for hybrid cloud NAS

- Global file system
- Global deduplication
- Integrated Snapshot
- Policy driven caching
- Immediately consistent file system for data integrity



A Buyer's Guide to Hybrid Cloud NAS

What is Hybrid Cloud NAS?

Hybrid cloud NAS functions like the enterprise NAS in your data center and appears to applications and users like a standard Windows file share. However, unlike traditional NAS which is limited to a single location, hybrid cloud NAS spans offices around the globe. It's more than just a gateway or a file distribution solution. It's a "NAS in the sky" that uses the cloud to simplify infrastructure and reduce costs, but with local NAS-like performance. Plus, by combining advanced snapshot techniques with the availability and reliability of cloud storage, backup and disaster recovery process can be eliminated.

Why Do You Need Hybrid Cloud NAS?

With the right capabilities, hybrid cloud NAS creates a solution with the economics, capacity, and scale of cloud storage with the features, performance, and flexibility of local storage.

Hybrid cloud NAS can provide significant benefits, including:

- **Eliminates inefficient local NAS and file servers at each office.** A hybrid cloud NAS makes cloud storage appear like a local file share for every user.
- **Eliminates need for local backups or cumbersome tapes.** Backup and archive are no longer needed when hybrid cloud NAS is deployed for primary storage. Amazon has 11 9s of durability and can survive the loss of two cloud data centers. When combined with global snapshots and Windows previous versions the result is a simple cloud and cache model that eliminates the need for all on premise unstructured data equipment and operations.
- **Built-in disaster recovery.** Hybrid cloud NAS uses the cloud as the authoritative data source to consolidate unstructured data. Since cloud providers use redundant copies of data and multiple data centers, data is always recoverable.
- **Pay as you go.** Only pay for the storage you need, taking advantage of the falling price of cloud storage, and instantly scaling your cloud instance to meet your needs.

Selecting hybrid cloud NAS raises a lot of questions. Does it use storage space efficiently and de-duplicate files? Can it support my diverse application set? Will my data be secure? Available? Backed up? Does it take advantage of cloud economics and still provide the performance of local filers?



Key Requirements of Hybrid Cloud NAS?

Here are some key requirements to look for when selecting Global hybrid cloud NAS:

1. Global File System

File operations on hybrid cloud NAS – even simple things like directory browsing – must be as responsive as a local NAS. This can only be accomplished with a single file system across all sites and the cloud with a global namespace. When users navigate files and folders in the network drive, it has to “feel” like all those files are there. Look for a system with a global namespace.

A global namespace (GNS) is a universal file directory that makes files on multiple servers or storage arrays appear to be in the same structure. The namespace could exist within a single data center, or may span multiple locations (less common, as this is much harder to do). It looks to the user as if all the files are on the same file server even though the data may be spread across hundreds of different nodes.

Traditional storage architectures were never designed for this. They use a local database on each storage system to store and manage metadata, but this doesn't scale across sites.

2. Deduplication

Deduplicating data reduces the amount of data stored and sent over your network, which reduces the overall storage footprint. Hybrid cloud NAS should de-duplicate data before sending it to the cloud and then only store each unique block from a file once in the cloud and at every site.

Here's what you should expect for hybrid cloud NAS deduplication capabilities:

- **It's inline.** Data is de-duplicated inline when files are created or changed. If a file block already exists, metadata references are created so the data doesn't need to be written.
- **It scales to petabytes of data.** Deduplication operations should minimize lookups for unique data, making it efficient even for large volumes of data.
- **Global deduplication and compression mean less data is sent over your network in the first place.** Only new, unique file data is sent to the cloud when it's created.

3. Integrated Snapshots, Backup, and Disaster Recovery

Hybrid cloud NAS needs to provide the same levels of availability and redundancy as a traditional NAS. The system should accomplish this in two ways:

- Snapshots of the entire file system, rather than just at the volume or folder level
- Leverage the cloud for backup and disaster recovery (DR)

Global snapshots are far more efficient and easier to manage than the volume-level snapshots found in many file systems. Essentially, global snapshots are a “time machine” for your entire file system.

Here's what you should look for when evaluating the snapshot capabilities of hybrid cloud NAS:

- **Snapshots should be captured on end points, but aggregated across the entire file system.**
- **Snapshots should be performance efficient.** They should be created using redirect on write (rather than copy on write), so there's no I/O impact as data blocks are never overwritten for new snapshots.
- **Snapshots should be space efficient and leverage deduplication.**
- **Snapshot management should be simple,** without the need to specify per volume or controller snapshots. Some Hybrid Cloud NAS or gateway solutions have volume-based snapshots.
- **You should be able to configure schedules and retention policies so you can keep snapshots as long as you want.**

In a typical distributed environment, backup and DR is done at each site. Even cloud gateway solutions are usually mapped from each site to the cloud. A true hybrid cloud NAS uses the cloud as an authoritative data source. Since cloud providers use redundant data copies and multiple data centers for system durability, your data is always recoverable. You can consolidate all your file archiving, backup and DR to have integrated backup and disaster recovery across all sites.

4. Policy-Driven Caching and Fast Sync

Users expect fast access, so hybrid cloud NAS caches active data locally, with the entire data set stored in the cloud. Machine learning based on file usage, "pinned folders," and configurable policy-based caching is employed to make sure the right files are cached while less used files recede into the cloud.

To assure fast performance, metadata should always be stored in flash (SSD) storage.

When it comes to caching, look for these 3 things:

- **Cache hot data on each controller based on read and write frequency**
- **Provide policy-based caching based on file types, folders, and other criteria**
- **All metadata stored in flash**

5. Data Integrity

You need to avoid version conflicts and file corruption. Every storage solution and file system has various mechanisms to ensure data integrity, but there are different approaches and important trade-offs in any system design. In a distributed file system, architecture decisions have massive implications for data integrity, and its opposite, data corruption.

There are two critical parameters to consider:

- Where does the authoritative file data live?
- How does a node behave if it's partially or completely disconnected?

You want a hybrid cloud NAS solution that ensures data integrity by:

- **Keeping a clear separation between the authoritative copy of data in the cloud and the local cache copy at each site.**
A “transactionally consistent” file system can maintain file integrity even if there’s a hardware or power failure – without falling back on a file system check or earlier file version.
- **Verify before upload.** Running checksum, dedupe, and compression on data before it’s sent to cloud to avoid sending any corruption to the cloud.

These capabilities will let you extend your corporate data into the cloud so you can take advantage of the cloud to run applications – like high-performance computing, CAD, antivirus, file auditing, search, indexing, and rendering – the same way they would run in your own data center.

Hybrid cloud NAS uniquely combines the performance of local NAS storage with the scalability, economics, and reliability of the cloud to let businesses consolidate their data center, easily manage their growing data, and eliminate cumbersome backup and DR processes.



Panzura, Inc. | 695 Campbell Technology Pkwy #225, Campbell, CA, USA | 855-PANZURA | www.panzura.com
Copyright © 2017 Panzura, Inc. All rights reserved. Panzura is a registered trademark or trademark of Panzura, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.