

# Global Namespace vs Global File System

## What's The Difference and Why Does it Matter?

### Why a Global Namespace?

It's easy to see the appeal of a single control plane to access and manage data no matter where it lives. Imagine having one place to get visibility into data across all your silos, identify hot and cold data, and plan and execute data migrations and data tiering across all your storage and cloud locations? And what if this same system allowed your users to search for relevant data across storage silos and feed AI/ML pipelines and create automated data workflows?

These are the many advantages of a global namespace for enterprise data storage. As unstructured data volumes continue to expand exponentially, data silos proliferate and IT budgets remain relatively flat, many organizations are interested in simplifying data visibility and managing data across various silos, and a global namespace can offer this.

However, it's important to note that a global namespace does not require a global file system (GFS), despite vendors often claiming this to be the case. A GFS sits in front of the data and serves the appropriate files, thus acting as a controller. While a GFS is useful in certain collaboration scenarios where simultaneous editing of large files is needed across geographically disparate locations that can share data without violating data privacy issues, the broader use case of data visibility across silos and simpler data management does not require a GFS.

This paper will summarize the differences between a global namespace and a GFS and review the benefits that Komprise Intelligent Data Management delivers sitting outside of the hot data path, without the overhead of a global file system.

### Comparing a Global Namespace with a Global File System

A global namespace delivers data visibility across heterogeneous file storage vendors, clouds and locations. It delivers a consistent way to search, find and use your data no matter where the data lives. A global file system addresses a different need: enabling geographically distributed access to the data managed within one storage vendor's file system. While a global file system gives access to the data it manages across locations and hence is often mistakenly marketed as the means to a global namespace, using a GFS in scenarios where what you need is global data management and visibility creates unnecessary lock-in, performance issues and high costs. We will explain why using the two common architectures of global file systems: storage-centric and metadata-based or "virtual".

### Global File System: Two Options

A global file system provides a consistent way to access the data or metadata residing in that file system from many locations, and where multiple users in different locations may be working on copies of the same file. It also provides a consistent way to access, configure and administer the file system. The two types of global file systems are:

- **Storage-centric:** Stores the data and provides access to it using a single mount that fronts all data requests and is always in the hot data path. By "fronts all data" we mean that all data and metadata requests are channeled

through this mount. Some vendors extend this notion to keep the bulk of the data as proprietary blocks in the cloud. In the case of a “cloud storage” GFS, you need to recognize that access to your data always requires licensing the GFS even when the bulk of your data may be in the cloud, which may unnecessarily add costs.

- **Metadata-based:** Also known as a virtual global file system, this type of GFS channels data and metadata access from existing storage file systems. The benefit of this approach is that it works across multiple storage vendors. However, there is a heavy price for this as all access must pass through the metadata-based controller. This slows down performance if it is implemented fully in software or increases costs substantially if it requires dedicated hardware. As with the storage-centric GFS, it is in the hot data path and manages data access even though it is not storing any data blocks.

**Summary:** A storage-centric GFS does not provide a truly global namespace. It can only provide visibility into data residing on that vendor’s storage system. A metadata-centric GFS can provide a global namespace across multi-vendor storage systems, but it must do so by fronting all data access, which will negatively impact performance and scalability.

## Global Namespace

A global namespace delivers unified visibility, management and data workflows across heterogeneous storage without being in the hot data path. There is no “controller” maintaining access to data. This is the only scalable way to see all your data in one place:

- It moves, tiers and manages data across different storage vendors.
- Users can directly access any data copied or tiered from the global namespace at the secondary location, resulting in zero disruption or performance impact.

As you can see, a global namespace does not require the heavy overhead of a global file system and understanding the differences can ensure you don’t get caught flat-footed in a heterogeneous storage environment. To review, a GFS delivers one siloed view of your data—but not the entire view unless you are only using one storage vendor, or you are willing to control all data access through a single vendor. While using a single vendor to control all your data might sound like a simpler solution, it may lead to unnecessary costs and lack of flexibility. For example, when using a global filesystem, even if all your data is in the cloud you access it via the GFS, and you then must pay high licensing costs in perpetuity to the GFS vendor to access your own data. Secondly, when using a GFS to manage and move data across different storage or to feed data pipelines, your organization will likely experience performance issues, poor scalability and unnecessary lock-in. This is because a GFS is an extraneous layer sitting in front of the hot data path.

## Komprise: Unified Unstructured Data Management - Not in the Hot Data Path

Komprise Intelligent Data Management delivers a global namespace that works across any storage—from the data center to edge to the cloud. The SaaS platform can manage all data in storage (with automated, policy-driven actions) and it never sits in the hot data path. Komprise is even more powerful with its unified metadata catalog that allows users to search, enrich and feed data pipelines from heterogeneous storage and clouds. Komprise Transparent Move Technology (TMT™) ensures there is no user disruption when IT tiers files to secondary storage using Komprise. That means users can access the files as usual from the original location or they can access them natively from the secondary location—such as object storage in the cloud. Hence, it enables multiple ways to access the data without getting in front of the hot data path and without lock-in. Just as you can access a website directly through its URL or through a Google Search, similarly, a global namespace allows direct access to the data from each storage location. The top unstructured data management challenge for enterprise IT is moving data without disrupting users and applications, according to [The Komprise 2023 State of Unstructured Data Management](#).

## When to Use a Global File System vs a Global Namespace

A global file system enhances the inherent value of a storage solution when employees need to actively collaborate in use cases such as engineering collaboration and design. But since 80% of data is cold and not actively accessed and typically less than 5% of data requires active collaboration, a global namespace is a better solution for data management, data tiering and feeding data to AI/ML applications with the best performance.

	Storage-Centric Global File System	Metadata-Centric Global File System	Global Namespace Through Komprise
Approach	In front of all data and controls access to all the data it stores	In front of all metadata and data and controls access to all the metadata it manages	Not in front of all data or metadata and analyzes, moves, manages and provides access
Is it a Global Namespace?	No – only allows access to data stored on that vendor’s system	Yes - but by fronting and adding a control layer in front of all data. All data must be migrated to the virtual global file system	Yes – across vendors, file, object, clouds without being in front of hot data
Visibility Use case – you want to get analytics on data usage, data growth and see data across locations and silos	Limited – only its data	Limited – not main use case	Yes – visibility into usage, growth, costs, deep analytics, reports
Management - Data Migrations you want to migrate data across vendors, across locations	Not a focus	Limited -fronts data path	Yes – 25x faster with Hypertransfer
Management - Data Tiering you want to transparently tier cold data to the cloud or to another vendor	Limited but block-level, lock-in, needs file system for ongoing access	Limited, in the hot data path, creates lock-in, requires ongoing licensing to access your own data	Yes - Flexible, policy-based, no lock-in with transparent, native access to your data
Data Search, Tagging – you want to search, find and use data no matter where it lives	Not a focus	Not a focus	Yes – Deep Analytics. Enrich data with ML, perform management based on search queries
Data Workflows – you want to feed AI, Big Data and enrich data	Not a focus	Not a focus	Yes – customizable data workflows to AI, Big Data
Collaboration Use Case – you want to modify the same data from different locations and require local caching for performance reasons	Yes – for its data, but may create data privacy issues wrt GDPR and other privacy laws	Yes – but fronts hot data path and may create data privacy issues wrt GDPR and other privacy laws.	No
Data Management Performance	File storage, not data management	Adds another software layer, creates bottlenecks	Fast by leveraging parallelism at every level
Data Management Scalability	File storage, not data management	May require dedicated infrastructure which limits scaling & increases costs	Scales linearly with no central bottlenecks
Data Management Cost	Not a focus	Hidden costs from fronting all data access	Saves 70% costs by right placing data
Lock-In	Creates lock-in	Creates lock-in	No lock-in
Best Suited For	Distributed storage across locations	Commodity storage that requires external controllers	Data visibility, analysis, management and data movement across storage, clouds

# Analyze and Manage Data with a Global Namespace, not a Global File System

Managing data growth requires an unstructured data management solution that delivers a global namespace, not a global file system. It is important to understand your use case and pick the best solution that fits your needs. Unstructured data management should sit outside the hot data path and bring visibility to all data, not just data that it fronts. It should right-place data while putting users in full control of their data. Contact Komprise to learn how a global namespace can cut costs, support departmental data services and help extract greater value from your unstructured data.

## Appendix I: Questions to ask your vendor when considering investing in a global namespace

1. Do you provide visibility across all storage vendors or does all data need to be stored in your file system?
2. Do you provide a global namespace?
3. Do you require a global file system to front our data for visibility across all data?
4. Is your technology in the hot data path?
5. Do we need an ongoing license with your company to access all our data?
6. If you're in the hot data path, what performance benchmarks do you have to show the impact on hot data performance?

## Appendix II: Making the Wrong GFS Decision – A Case Study

Here is a recent customer situation that shines a light on the importance of asking the right questions up front and separating the marketing pitch from the product reality.

### Scenario

Customer was pitched on the idea of being able to swap storage in and out easily, without having to make any changes to the environment, such as creating aliases and DNS, etc.

### Proposed Solution

Customer embarked on a global namespace project and decided to implement a metadata-based GFS. They wanted their users to be able to access their files and no matter what storage they go to in the future, the namespace stays the same. No changes. No disruption.

### Results

The customer came to realize that the approach they chose required implementing a new software layer in front of their file systems. This device must sit in front of the hot data on NetApp, Pure Storage, Isilon, etc., to be able to do the conversion. They managed to successfully implement a solution and it worked well for a while. But over time, the network became slower and slower because this new layer becomes a bottleneck. Even worse is how difficult it is to move from a global file system sitting in front of all your data. The customer has now implemented analytics-driven data management from Komprise to provide global visibility and easy management across vendors. They are no longer experiencing any of the performance issues they had with a GFS.

## About Komprise

Komprise is the leader of unstructured data management and mobility software that frees enterprises to easily analyze, mobilize, and monetize the right file and object data across clouds without shackling data to any vendor. With Komprise Intelligent Data Management, you can cut 70% of enterprise storage, backup and cloud costs while making data easily available to cloud-based data lakes and analytics tools.

Learn more at: [www.komprise.com](http://www.komprise.com)

---



**Komprise, Inc.**  
1901 S. Bascom Ave. Suite 500  
Campbell, CA 95008  
United States

For more information:  
Call: 1-888-995-0290  
Email: [info@komprise.com](mailto:info@komprise.com)  
Visit: [komprise.com](http://komprise.com)

For media requests email  
[marketing@komprise.com](mailto:marketing@komprise.com)  
©Komprise, Inc. All rights reserved.