



Transparent Move Technology™ (TMT)

Leverage the Full Power of the Cloud
without Disrupting Users with Komprise TMT.

[White Paper]

TABLE OF CONTENTS

Executive Summary	3
Introduction	3
Key Considerations When Tiering Data to the Cloud	4
Komprise Transparent Move Technology	6
Availability and Scalability	10
TMT in Action	10
Summary	11

EXECUTIVE SUMMARY

Explosive data growth requires a re-think of how data is managed. Storage capacity is running out, Data is continuing to grow explosively, and we now have several storage options such as the cloud where data can be actively used, not just passively archived. But to do this, you need a new way to manage your data—one that's focused not just on storage efficiency but how to extract value from the data stored in the cloud.

This white paper describes how the patented Komprise Transparent Move Technology™ (TMT) goes beyond storage tiering to analyze, migrate, tier and replicate data across multi-vendor storage and clouds while enabling native use of the data at each layer. This is possible without disrupting users and without vendor lock-in.

INTRODUCTION

A few decades ago, data storage options were primarily disk and tape. Since tape was offline, data had to be on disk before it could be actively used. So, it made sense for disk-based storage to hierarchically tier rarely-used cold data to lower disk tiers or to tape and to bring it back as needed. Data management considerations focused almost exclusively on the storage hardware. Storage teams had little visibility into the nature of the data they managed and were forced to treat all data the same. Critical business data as well as low value or temporary data were both actively managed the same way.

This blind approach is not economically viable anymore. We are experiencing explosive growth of unstructured data, leading to a proliferation of data storage silos. We now have many classes of storage where data can be actively used, including flash, disk, private cloud object storage, public cloud file storage, and public cloud object storage. Most enterprise IT organizations today want a hybrid cloud architecture where the data they move to the cloud should leverage all the native cloud capabilities.

The proliferation of data and active storage options requires a data management solution that is independent of the storage in which the data resides. An independent data management solution analyzes and moves data across multi-vendor storage and cloud. Like storage tiering, modern data management should enable moved data to be transparently accessible without disrupting users from the original location. But unlike storage tiering, a modern data management solution should enable moved data to be directly accessible in the cloud without requiring access through the original file system first. Beyond being able to leverage the full compute power and elasticity of the cloud, you can also take advantage of cloud native services and applications without locking your data into a proprietary storage technology.

Komprise TMT powers Komprise Intelligent Data Management, software that delivers key data metrics across all storage and automatically moves data according to user-defined policies—without any changes to user or application access. Komprise TMT delivers transparent data tiering and allows you to use any of the compute capabilities of the cloud on your data.

KEY CONSIDERATIONS WHEN TIERING DATA TO THE CLOUD

When moving data from file-based Network Attached Storage (NAS) to cost-efficient object-based storage, such as the cloud, here are some things to consider:

1. Preserve transparent file access from the source.

NAS storage environments use file protocols such as NFS and SMB. When moving data from a NAS, it is important to ensure users and applications can still access the moved files from the NAS as before. Moved data should be presented as files, even if they are stored as objects in the cloud.

2. Enable data access as objects in the cloud.

Moving data to the cloud is not just about cost efficiencies. It's also about leveraging the elastic compute power of the cloud. Yet, storage-based data tiering solutions and cloud storage gateways treat the cloud like a hard disk, converting data into proprietary blocks, which means you cannot directly access your data as cloud objects. With storage-based tiering and cloud file storage systems, the cloud becomes a digital landfill of cold data blocks. Instead, imagine rather if you could access your data as objects in the cloud and use any of the analytics, AI and cloud compute functions provided by the cloud directly on your data without additional bridge software or license fees.

With storage-based tiering, the cloud becomes a digital landfill of proprietary data blocks.

3. Your data will outlive your storage.

The lifecycle of data and storage are out of sync. On-premises storage hardware typically has a supported lifecycle of 3 to 5 years, while your data typically lives for decades. With the cloud, new options for how to store and process data are continually emerging. Locking data into a single storage architecture or vendor no longer makes business sense.

4. Eliminate performance degradation of hot data.

If data is moved through a proprietary mechanism, then an agent, stub or virtualization layer is needed in front of hot data. This is not desirable as the performance of hot data is impacted. Solutions such as cloud gateways, with the concept of global file systems, create an overlay in front of data adding complexity and a performance penalty vs native data access. Data movement should occur in the background, without impeding hot data access and without requiring proprietary agents or stubs.

5. Cut backup costs, rehydration and cloud egress costs.

Proprietary data tiering solutions such as storage “pools” and cloud gateways create storage efficiency but do not reduce third-party backup, DR or access costs. Since the bulk of the costs are in the active backup and management of data, these solutions miss 80% of the savings opportunities. They also result in high cloud egress and retrieval costs, all of which minimize storage savings. Data management solutions that move data in standard form can shrink backup costs by 80% and provide access to moved data without rehydration or unnecessary fees while leveraging native features of object storage for data protection.

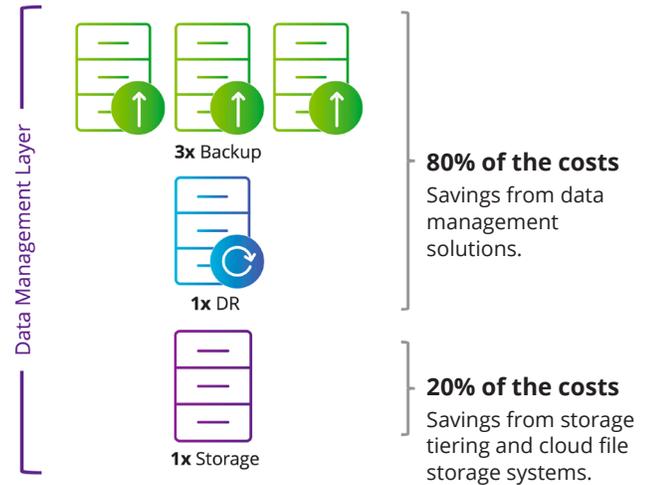


Figure 1: A storage-centric approach misses the bigger opportunity for cost savings that a data management solution can provide.

Understand Your Options: Gateways, Storage Array Tiering, Komprise TMT

Gateways

- All data must be migrated from the existing NAS to the gateway appliance, creating a new storage silo and disrupting users.
- Hot data is cached locally and **all data (both hot and cold) is tiered to the cloud** in proprietary format.
- Data cannot be accessed from the cloud natively and cloud-native applications are not accessible.
- Since the data is tiered to the cloud in proprietary format, data must now be accessed through the gateway. This causes disruption to users and applications.
- This solution provides only storage cost savings because the backup and DR footprint is not reduced. Access by these tools causes the entire file that is moved to be copied.

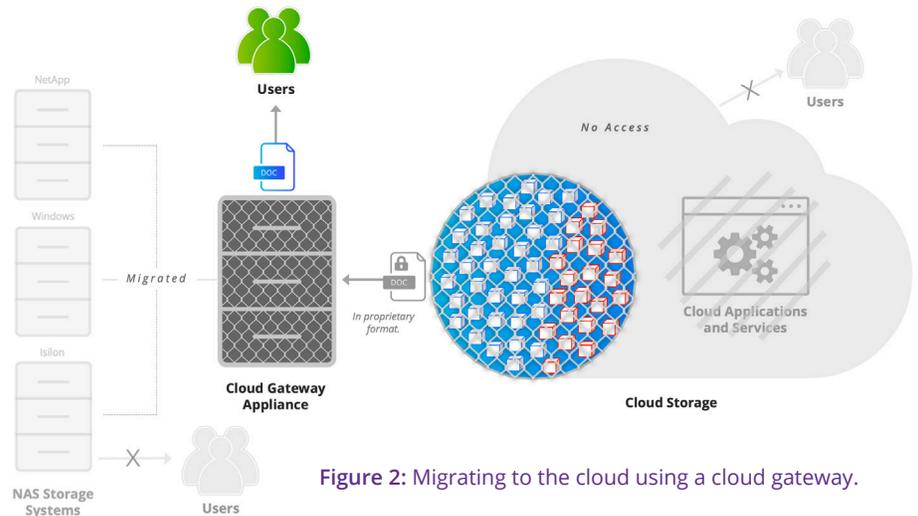


Figure 2: Migrating to the cloud using a cloud gateway.

Storage Array Tiering

- Creates a proprietary file system in the cloud. Users can use a “pool” feature where each NAS creates its own silo in the cloud.
- Data cannot be accessed from the cloud natively and cloud-native applications and services are not accessible.
- Users must access data from the local NAS only because the data in the cloud is stored in the vendor’s proprietary format.
- The data is sent back from the cloud in a block format, which contains all or part of the file requested. The block is then processed by the local storage array and then becomes a file.
- This solution provides only storage cost savings because the backup and DR footprint is not reduced. Access by these tools causes the entire file that is tiered to be copied.

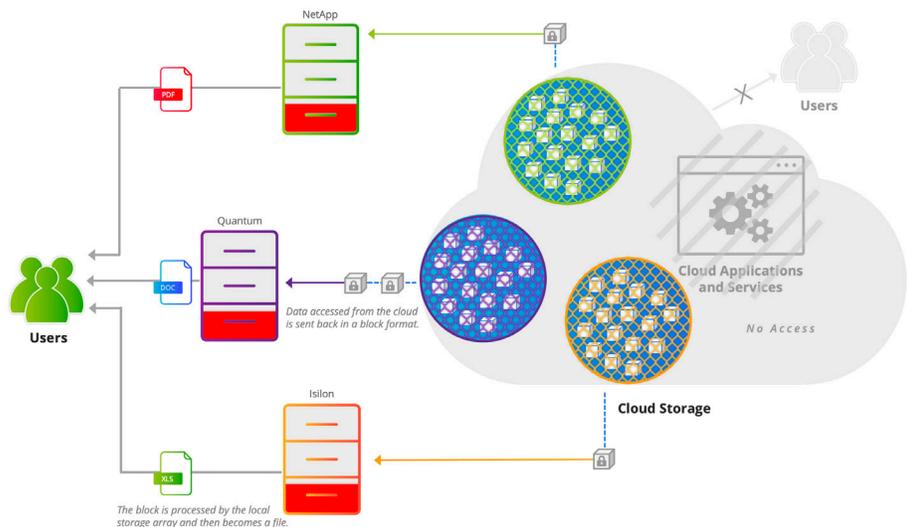


Figure 3 : Moving cold data to the cloud with storage array tiering.

Komprise TMT

- Komprise TMT seamlessly extends the existing NAS to the cloud and transparently moves cold files from the NAS to the cloud in their native format using TMT.
- There is no data lock-in and users are also able to leverage cloud-native applications and services without disrupting performance.
- With TMT users can access the tiered data from the source NAS, exactly as they did before, or directly from the cloud.
- There's no vendor lock-in, so the tiered data can be moved from one NAS to another without rehydration.
- TMT reduces both the backup and DR footprint, providing significant cost savings.

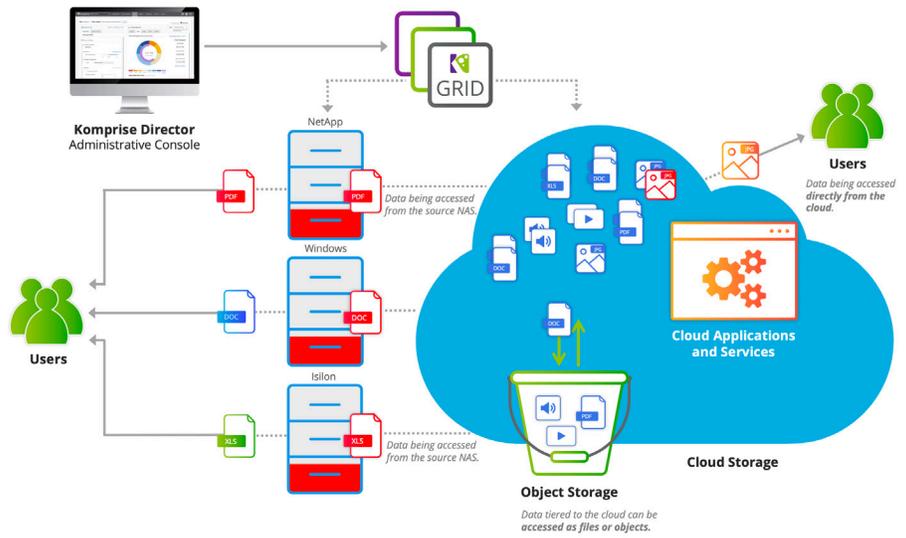


Figure 4 : Moving cold data to the cloud with Komprise TMT.

Komprise eliminates these data movement challenges with a patented technology that transparently provides file-object duality across NAS and cloud using standard protocols without getting in the path of hot data.

KOMPRISE TRANSPARENT MOVE TECHNOLOGY

Komprise TMT tiers, archives, migrates, replicates and extracts data using these principles:

1. Transparent access to moved files from original NAS.

Komprise TMT moves files by policy to a secondary storage of your choice such as the cloud. It leaves behind industry-standard symbolic links that are dynamic and resilient, called Komprise Dynamic Links. These links look like the original file and preserve the original file permissions and attributes. Users and applications can open and access the moved files from their original location exactly as before, without any changes.

Komprise works across multi-vendor NAS since it follows standard NFS and SMB constructs. A benefit of moving a file with TMT as opposed to proprietary blocks with storage tiering is that TMT reduces backup and DR costs by shrinking the data footprint on primary storage whereas storage tiering only provides storage efficiency. The cold, infrequently-accessed files are moved and protected in the cloud so your backup application only needs to protect the hot data. This is on top of the dramatic savings of using object storage in the cloud.

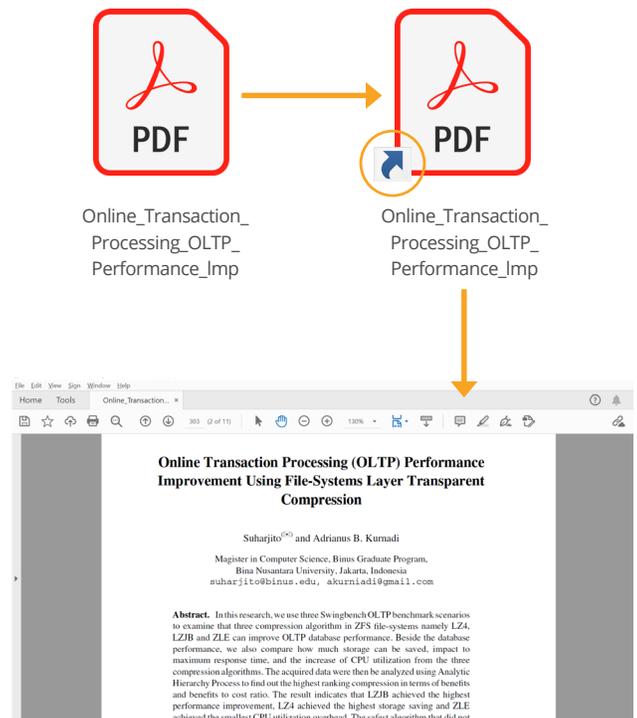


Figure 5: Moved files appear to the user just as they did before and open like a normal file on the desktop.

2. Preserve file-object duality, so moved files can be natively accessed as objects in the cloud.

When Komprise moves a file to object storage such as Amazon S3, it writes the entire file as an object. This means you can directly use an S3 browser to access your moved data. Any AWS cloud service such as RedShift or third-party services on AWS can operate on your data without going through Komprise or the original storage file system. You can turn Komprise off at any time or switch your storage vendors or clouds without any repercussions. This is a major difference from storage-based tiering or “pools” solutions that do not move the entire file but rather move blocks to cheaper storage such as the cloud. With block or storage-based tiering solutions and cloud file storage gateways, you cannot directly access the moved data in the cloud without first going through the storage technology—which adds cost in licensing and/or capacity and it prevents use of native cloud tools to make use of and extract value from your cold data.

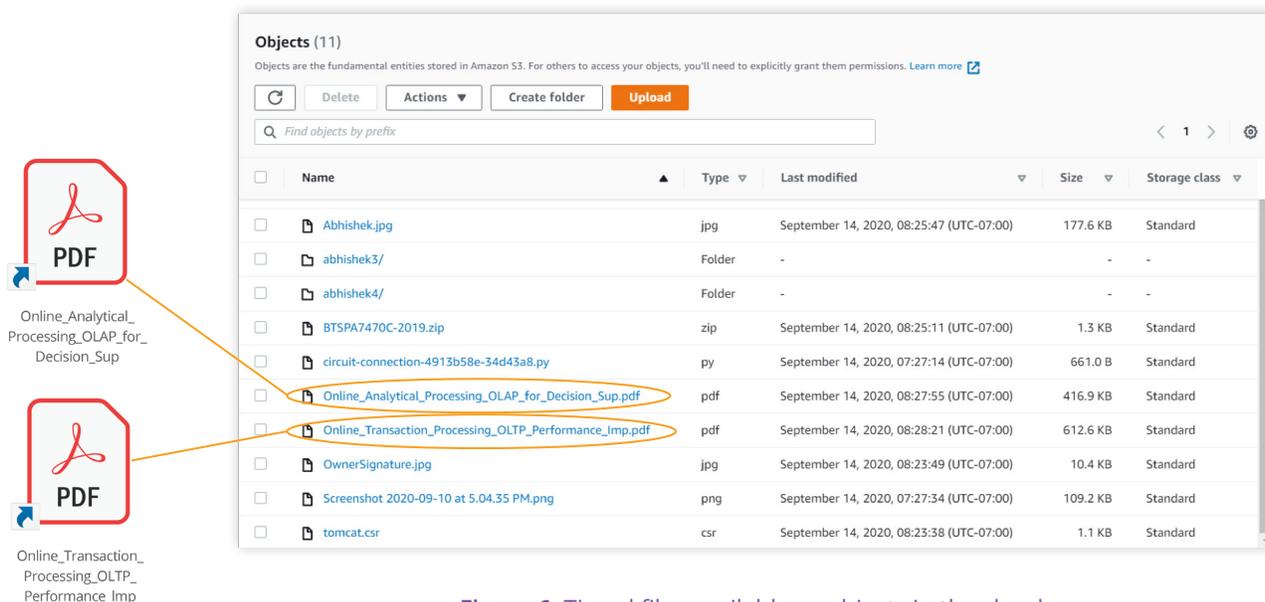


Figure 6: Tiered files available as objects in the cloud.

3. No obstruction of hot data.

Storage tiering, global file storage gateways and data virtualization solutions move data in a proprietary form and are in front of the hot data path. While some use cases may require this, often the performance impact to hot data and the cost of licensing this solution to access your data in the cloud is undesirable. Komprise moves data using standard protocol constructs, so it is not in the hot data path. Komprise is only called when cold files are accessed, which happens rarely. This improves performance of primary storage and allows you to run data analytics in the cloud versus on your file storage.

4. Move data without proprietary interfaces.

Proprietary interfaces such as stubs or agents are brittle and problematic to manage. If stubs are deleted, they can leave data orphaned, creating havoc for users and applications. Instead, as mentioned earlier, Komprise maintains access to moved data using Komprise Dynamic Links. Proprietary solutions require a central database to maintain the reference from the stub to the actual storage location, whereas Komprise Dynamic Links contain all the file system information a user or application requires in a simple, scalable format.

Figure 7 shows the logical workflow when a user or application recalls cold data that has been moved by Komprise. As shown in the figure, when a user clicks on a Komprise Dynamic link, the user's file system creates a file system request and sends it to the Komprise Grid. The Komprise Grid is a fault-tolerant, load balanced collection of two or more virtual appliances called Komprise Observers. The request is dynamically bound to a Komprise Observer, which parses the request, determines the location of the file and fetches the file, caches it locally to reduce egress costs and returns the data via a file system response. All of this occurs unbeknownst to the user or the application accessing the file via the link while Komprise validates and maintains the Dynamic Links.

5. Enable file access in the cloud to minimize rehydration.

Komprise TMT moves data at the file level with all the file metadata fully preserved at the target so the moved data can be accessed as files from both the source and from the target. Targets can be public or private clouds or other NAS storage. For example, you can move data from an Isilon SMB share to Amazon S3, users can access it from the Isilon as they always have or view the files in the cloud through a Komprise Observer. You can use files in the cloud for any file-based applications without having to bring data back to the original NAS. This has the added benefit that your big data analytics projects don't tax your primary NAS.

1. User or app calls file that has been moved per plan resulting in file system request.
2. Komprise Dynamic Link redirects the request to the Grid.
3. Komprise Grid resolves the link and fetches the data requested from its current location. It may opt to pre-fetch additional data segments.
4. Komprise Grid uses a file system response to present the data and metadata back to the user or app with high performance and perfect fidelity.

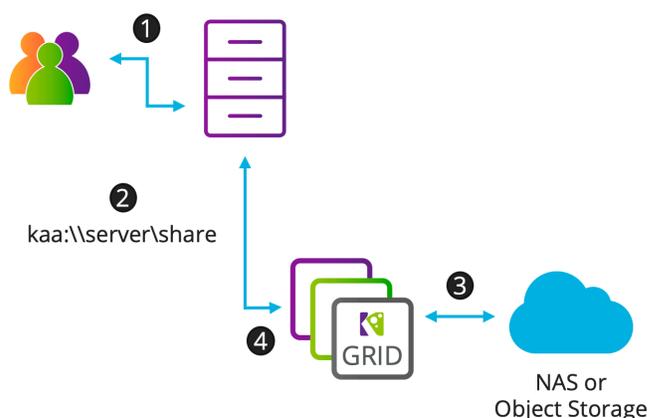


Figure 7: The logical workflow when a user or application recalls cold data that has been moved by Komprise TMT.

6. Provide full permissions and access control on moved data.

Komprise ensures that the original permissions and access control on files are fully enforced when a moved file is opened from the original source. If the data is rehydrated back to the NAS, it returns with identical permissions, access control lists (ACLs), and attributes—nothing changes. Since users aren't affected by moving the data, control and management of the data is back in IT's hands.

Komprise TMT moves data at the file level:

- All the file metadata is fully preserved at the target.
- Moved data can be accessed as files from both the source and from the target.

7. Optimize recall of cold data to minimize rehydration and egress.

When reading archived data stored in the cloud from the original storage location there are several considerations:

- a. **Performance:** Komprise will cache the file locally so that any subsequent reads are local and therefore low latency. Large files are “streamed” from the cloud, meaning that users can start reading the file during the transfer.
- b. **Cost:** By caching read objects, Komprise reduces egress fees.
- c. **Flexibility:** Data that has been untouched for years can become critical for a new project. Komprise gives you the power to set recall policies based on repeated access to promote data back to primary or do a proactive “bulk recall” of entire directories or shares that are about to become active. Data management with Komprise enables ongoing analysis of data to fine tune to your evolving requirements.

8. Establish a layered defense against ransomware.

Komprise customers can use immutable cloud storage via AWS S3 Object Lock to provide another layer of defense against ransomware. This means any file data archived or replicated to the cloud via Komprise can be protected against deletion or alteration for the term of the retention set by the customer. Solutions using proprietary data formats, such as storage tiering and cloud file storage gateways, cannot take advantage of Object Lock protection.

9. Provide systematic, continuous tiering for maximum savings.

Since TMT does not disrupt users and applications, it allows IT to deploy tiering across their storage devices using policies. IT can “set it and forget it”. IT does not have to regularly scan their file systems, find old files and have discussions with their end users to acquire permission to tier cold data. This not only saves IT from significant manual overheads, it ensures that the cold data is continuously tiered, resulting in maximum savings.

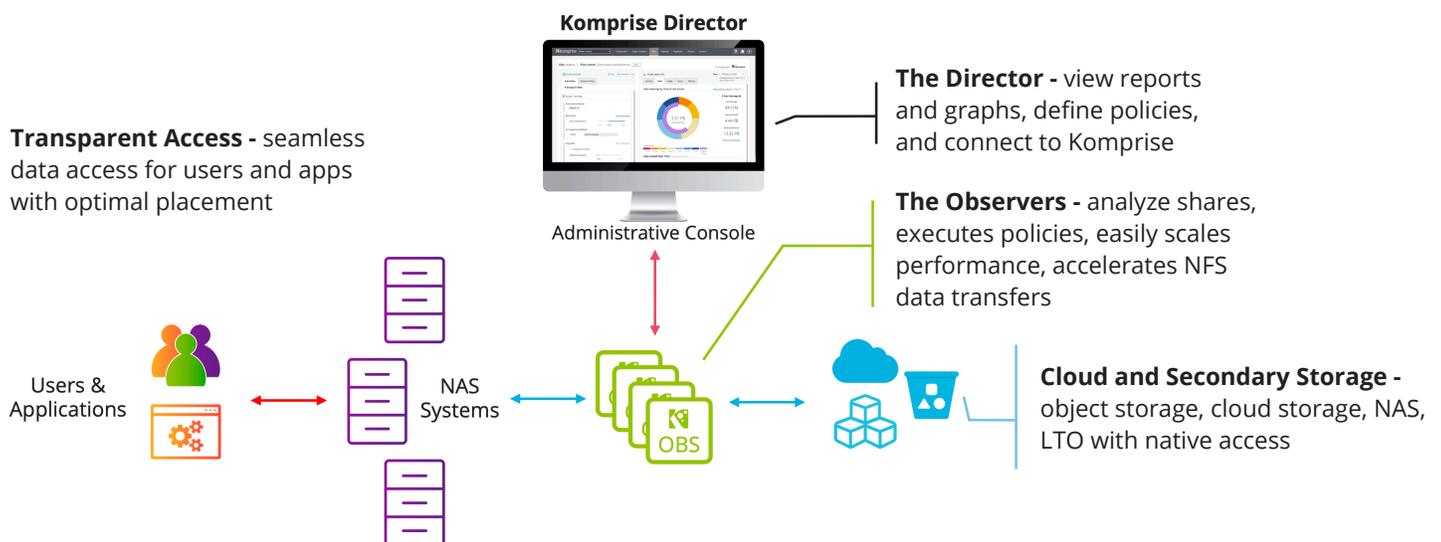


Figure 8: The logical workflow when a user or application recalls cold data that has been moved by Komprise TMT.

AVAILABILITY AND SCALABILITY

Komprise TMT is a capability delivered by the grid of Komprise Observers. Komprise uses a distributed, scale-out architecture designed for today's massive data growth. It runs as a hybrid cloud service with a "grid" of one or more Komprise virtual appliances, called Observers, deployed adjacent to your data, as shown in Figure 4. The Komprise Elastic Grid uses a cooperative, distributed algorithm, which enables scalability, load balancing and fault tolerance for high availability.

Observers analyze data across on-premises NAS storage, securely move and migrate data by policy and provide transparent file access to data that's been moved. Another virtual appliance, called a Director, runs in the cloud and functions as the management console. The Director can manage Observers over multiple sites and cloud deployments. Resiliency and high availability are built-in without requiring any dedicated infrastructure. To learn more about the Komprise architecture, read the [Komprise Architecture Overview White Paper](#).

TMT IN ACTION

Komprise TMT is vendor agnostic and works with most common storage devices in use today. In the table below, you can see how TMT has helped customers from the [genomics industry](#) to [higher education](#) tackle their rising data growth costs by migrating and tiering their cold data to more efficient storage.

Customer	Use Case	Benefits
Pfizer	Cold Data Strategy <ul style="list-style-type: none"> • Needed to reduce data consumption through externalization of cold data to low cost, cloud storage • Analytics to ensure right data moved AWS with no disruption to users/apps and cloud native access 	<ul style="list-style-type: none"> • Externalized 1PB of data in first year, 90-day ROI • Saving 75% on storage by continuously moving cold data to Amazon S3 as it ages. • Native cloud access with tagging enables data lakes and big data analytics for researchers • Reallocation of funds, continuous data insight, improved automation
Pacific Biosciences	Massive Data Management <ul style="list-style-type: none"> • Manage data across different NAS & secondary storage (NetApp E-Series, SpectraLogic) transparently • Create policies that run continuously to ensure maximum savings 	<ul style="list-style-type: none"> • Manage massive data growth within flat budgets • Able to move data as it becomes cold. IT no longer must ask permission since there is no user disruption. Enables IT to deploy archiving company-wide for maximum savings
Cadence Design Systems	Future-proof with Cloud & Create Cost Efficiency <ul style="list-style-type: none"> • Use Komprise to identify cold data across on premises NAS and transparently move to cloud—cut 60%+ costs 	<ul style="list-style-type: none"> • 60%+ of data was cold and not recently accessed, moving this to cloud cut storage, DR, and backup costs • Created a path to the cloud
Leading California University	Reduce DR Costs <ul style="list-style-type: none"> • Put DR copy in the cloud and eliminate on premises NAS mirror 	<ul style="list-style-type: none"> • Cut 50% of NAS costs by eliminating half the footprint • Instant recoverability in the cloud • Fit with their cloud-first initiative

SUMMARY

Given the unabated growth of unstructured data, IT leaders are investing in hybrid cloud strategies. To successfully leverage the cloud and its services, data needs to be moved such that users have native access to it once there, which isn't possible when using a proprietary storage file system or gateway to move that data. This requires an independent data management solution that works across multi-vendor storage and clouds using standard protocols. Komprise Transparent Move Technology is a core capability of Komprise Intelligent Data Management. With TMT, you can tier data across NAS and cloud while preserving file-object duality. Enterprises can cut 70%+ of data storage and backup costs while transforming to the cloud, while positioning themselves to fully leverage the compute power of the cloud to monetize data and generate new value through AI/ML, big data and cloud analytics.

Using Komprise TMT, you can take control of rampant data growth with a strategic approach, allowing you to:

- Move data without stubs or agents.
- Minimize impact on users and applications of tiered data.
- Retain full data access from source or target.
- Avoid vendor lock-in to storage devices or to Komprise.
- Eliminate obstruction to hot data and enable faster recall of cold data.
- Get more value from your data by leveraging native cloud capabilities.
- Deliver efficient ransomware protection by using expensive protection on hot data only and storing cold data in immutable object storage.
- Dramatically reduce storage costs by reducing backup and DR footprint.

ABOUT KOMPRISE

Komprise is the industry's only multi-cloud data management-as-a-service that frees you to easily analyze, mobilize, and access the right file and object data across clouds without shackling your data to any vendor. With Komprise Intelligent Data Management, you are able to know first, move smart, and take control of massive unstructured data growth while cutting 70% of enterprise storage, backup, and cloud costs.

To learn more about Komprise Intelligent Data Management visit [Komprise.com/product](https://www.komprise.com/product)



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
Visit: [komprise.com](https://www.komprise.com)

For media requests email
marketing@komprise.com
©2021 Komprise, Inc. All rights reserved.