komprise

WHITE PAPER

How to Manage Cloud Costs and Accelerate Cloud Data Migration

With cloud adoption skyrocketing, IDC reports that managing cloud costs and accelerating cloud migrations have become top priorities. But 80% of businesses will **overspend** their cloud infrastructure budgets, according to Gartner—**due to a lack of cloud cost optimization**.

Storage is the second largest area of cloud spend, so getting a handle on those costs is critical. This white paper shows how an analytics-driven approach to data management can save 50% of cloud storage costs while simplifying cloud data migrations.

The Challenges of Managing Cloud Data: low visibility and complex pricing

Managing cloud data costs takes significant manual effort, multiple tools, and constant monitoring. As a result, companies are using less than 20% of the cloud cost-saving options available to them. "Bucket sprawl" makes matter worse, as users easily create accounts and buckets and fill them with data—some of which is never accessed again.

When trying to optimize cloud data, cloud administrators contend with **poor visibility and complexity** of data management:

HALF of the **175ZB** in 2025 will be in the **public clouds**.

—IDC

80% will **overspend cloud budget**

—Gartner

How can you know your cloud data?

- How fast is cloud data growing and who's using it?
- How much is active vs. how much is cold?
- How can you dig deeper to optimize across object sizes and storage classes?

How can you make managing data and costs manageable?

- It's hard to decipher complicated cost structures.
- Need more information to manage data better, e.g., when was an object last accessed?
- Factoring in multiple billable dimensions and costs is extremely complex: storage, access, retrievals, API, transitions, initial transfer, and minimal storage-time costs.
- There are unexpected costs of moving data across different storage classes (e.g., Amazon S3 Standard to S3 Glacier). If access isn't continually monitored, and data is not moved back up when it gets hot, you will face expensive retrieval fees

These issues are further compounded as enterprises move toward a multicloud approach and require a single set of tools, policies, and workflow to optimize and manage data residing within and across clouds.

©2020 Komprise, Inc. All rights reserved.

Multicloud Data Management and Cloud Cost Optimization

What is needed is radical simplification to take the pain out of managing cloud data and saving costs. **Komprise Intelligent Data Management for Multicloud** removes the complexity of managing cloud data within a cloud and across clouds with a single set of tools, policies, and data management functions for today's multicloud world. Consider all you can see, know, and do with your cloud data with Komprise.

Gain Visibility Across Cloud Accounts

First you need to get your true cloud picture—across all your cloud accounts and services—*no matter what vendors and cloud services you use.* Komprise provides instant visibility so you can know your data usage, growth, and costs so you can optimize your cloud data. Simply point Komprise at your cloud accounts and see how much data you have, what storage classes are being used, who's using the data, how it's growing, and data access probabilities. With this visibility, you can understand how your data is actually being used to develop a cost-saving strategy.

Forecast Savings and Plan Data Management Strategies

Once you gain visibility into your true cloud picture, you can understand your current cloud costs and establish a baseline. To better manage your data, Komprise lets you examine "what-if" scenarios to see exactly how much you'll save with different data management policies—all *before* you start moving your cloud data. Set policies on when cold data will get archived across storage classes, and Komprise instantly projects your company's savings.

Archive Based on Accurate Data Usage

Managing the lifecycle of your data in the cloud is key to ongoing savings. But that's only possible with insight from the right data. With Komprise, you can set object tracking and Object Lifecycle Management policies in the cloud that are based on when it was **last accessed** (last read or written) which is far more accurate than what the cloud tracks—**last modified** (or written).

Last-Accessed
The most accurate data usage pattern used by Komprise, vs. last *modified,* used by cloud providers.

The last-*modified* time doesn't identify the *most prevalent data usage pattern*—where data is created once and read frequently again and again making it hot data. Using last *modify* time can result in archiving hot data to lower storage classes, thereby reducing performance, causing disruption, and in some cases, breaking applications. In addition, it can actually increase costs incurred by frequently accessing the hot data from a lower tier that has much higher access fees.

Komprise archives based on *last-accessed* time of the objects, providing a more predictable decision on the objects that will be accessed in the future, which helps intelligently archive the data in a cost-effective and efficient way without disrupting users and applications. Decisions based on **access** enable Komprise to archive to the most cost-effective storage classes, including S3 Glacier and S3 Glacier Deep Archive, without the risk of increased costs from accessing hot data in cold storage.

Digging Deeper

Beyond savings, it's important to derive more value from your cloud data—harnessing its power across all your cloud providers. With Komprise Deep Analytics, you can automatically develop a virtual data lake of all your data and use it to search, tag, and operate on your data regardless of where it resides. You can quickly find a needle across your cloud haystacks and operate on it, and feed your AI and ML applications what they need and extract value from your data. Get deeper insight into how your data is being used and who's using it and what sorts of objects are being used. Develop custom queries using a simple UI to drive these queries and operate on the results.

S3-to-S3 and Replication Across Clouds

S3-to-S3 migrations are made easy with Komprise. You can bring data into the cloud and migrate data across clouds for risk abatement, improved resiliency, or to simply reduce storage costs. Komprise makes cloud-to-cloud migration and asynchronous replication easy with the fast, resilient Elastic Data Migration. **Elastic Data Migration** enables you to simply pick your source and your destination—even if they are on different clouds—and create migration or copy tasks. You can run hundreds of migrations in parallel. Komprise automates all the heavy lifting so you don't have to. It finds the most efficient way to migrate the data with multi-level parallelism that exploits the inherent parallelism of each data set. It adjusts to network unavailability and other issues and retries automatically so you don't have to. It reports on the migration status and performs multiple iterations to accommodate incremental changes.

©2020 Komprise, Inc. All rights reserved.

To read more on how Komprise Elastic Data Migration works, check out **this white paper**. You can also move certain data back and forth between on-prem and the cloud where needed to address the needs of your application.

Reduce cloud storage costs by more than 50%

Cloud providers offer a range of storage services. Generally, there are storage classes with higher performance and costs for hot and warm data, such as Amazon S3 Standard and S3 Standard-IA, and there are storage classes with much lower performance and costs that are appropriate for cold data, such as S3 Glacier and S3 Glacier Deep Archive. Data access fees and retrieval fees for the lower cost storage classes are much higher than that of the higher performance and higher cost storage classes.

To maximize savings, you need an automated data management solution that takes into account data access patterns to dynamically and cost optimally move data across storage classes (e.g., Amazon S3 Standard to S3 Standard-IA or S3 Standard-IA to S3 Glacier) and across multi-vendor storage services (e.g., NetApp Cloud Volumes ONTAP to Amazon S3 Standard to S3 Standard-IA to S3 Glacier to S3 Glacier Deep Archive).

While some limited manual data movement through Object Lifecycle Management policies based on modified times or intelligent tiering is available from the cloud providers, these approaches offer limited savings and involve hidden costs.

Komprise automates full lifecycle management across multi-vendor storage classes using intelligence from data usage patterns to maximize your savings without heavy lifting. As the following analysis shows, this results in 50%+ cloud storage cost savings.

Compare the Cost Savings

We compared the cost of 1PB of data on Amazon S3 Standard versus managing it via Komprise Intelligent Data Management.

Assumptions:

1. 1PB of data consisting of 1 billion objects with an average object size of 1MB
2. For Komprise, we assumed the data is stored across three storage classes: Amazon S3 Standard, S3 Standard-IA, and S3 Glacier Deep Archive
    a. 10% of the data is kept in Amazon S3 Standard.
    b. 20% of the data is in S3 Standard-IA.
    c. The remainder of the data is very cold and stored in S3 Glacier Deep Archive.
3. We assumed the following access patterns that are based on what we have seen from our existing customers:
    a. Hot objects in Amazon S3 Standard see 882 million accesses per month. This translates to 97% of the access being to the hot objects.
    b. Warm objects in S3 Standard-IA see 18 million accesses per month. This translates to 2% of the access being to the warm objects.
    c. Cold objects in S3 Glacier Deep Archive see 9 million accesses per month. This translates to 1% of the access being to the cold objects.
4. We assumed the following modification patterns:
    a. Hot objects in Amazon S3 Standard see 147 million writes per month. This translates to 98% of the modifications being to the hot objects.
    b. Warm objects in S3 Standard-IA see 3 million writes per month. This translates to 2% of the modifications being to the warm objects.
    c. Cold objects in S3 Glacier Deep Archive are not modified.

Results: 55% storage savings with Komprise

The table shows that Komprise **reduces the storage cost by 55%.** Note that these savings *include* the purchase cost of Komprise. By using last access times and the probability of data access, Komprise intelligently tiers the data across the higher performance storage classes, from Amazon S3 Standard to the S3 Standard-IA. Once the

©2020 Komprise, Inc. All rights reserved.

probability of access is low, Komprise intelligently archives the data to S3 Glacier Deep Archive. Since this is based on access probabilities, you can use Komprise across your buckets and automatically achieve these cost savings without disruption.

*Table 1. Cost savings from archiving data: 1PB, **1B objects** with average size of **1MB** per object*

Storage Class	No Archiving		Komprise Intelligent Data Management	
	% Data in Storage Class	Probability of Access	% Data in Storage Class	Probability of Access
S3 Standard	100%	100%	10%	97%
S3 Standard-IA			20%	2%
S3 Glacier Deep Archive			70%	1%
Monthly Cost	**$22.1K**		**$9.8K**	
% Savings			***55%***	

Summary

Companies are leveraging the cloud more than ever. IDC predicts that half of all the data will be in public clouds by 2025. Given the cloud's pay-as-you-go model, an analytics-driven, and automated data management approach is critical to keep costs down while enabling new operations and functions to extract more value from all the data, whether it's hot or cold.

Today, it is almost too easy to create new cloud accounts and buckets and fill them with data without regard to the ongoing costs that will be incurred. What is more, data usage patterns are in constant flux, which makes automated policies that change as access patterns change a crucial capability. Because cloud administrators are the custodians, not the users of the data, they don't know how the data's being used.

Komprise Intelligent Data Management for Multicloud provides unprecedented cloud data insight within and across clouds, continuously optimizes data storage to dramatically reduce costs based on actual data use, and mitigates risk by simplifying the migration of data between multiple cloud providers and between on-premises and the cloud. You can also get more from your cloud data by creating virtual data lakes that can be used to search, tag, and operate on all your data regardless of where it resides. This helps quickly establish AI and ML operations and ensures that you get the most from your hot and cold data.

With so much data in the cloud, an analytics-driven approach to cloud data management is a strategic move to get more from your multicloud data and save substantial costs.

Learn More

Learn more about how Komprise can help you better manage cloud costs and accelerate cloud data migrations. Contact **sales@komprise.com**

komprise

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

For media requests email
marketing@komprise.com
©2020 Komprise, Inc. All rights reserved.