aws storage

How to optimize storage costs using Amazon S3

Contents

Introduction

On-premises storage can be costly and difficult to manage. Upfront investments in physical hardware, maintenance contracts, expensive refresh cycles, and managed migrations during hardware upgrades significantly increase your total cost of ownership (TCO). Most organizations require multiple storage systems, creating data silos that add operational complexity and prevent you from gaining valuable insights.

To break free of expensive hardware refresh cycles, organizations are turning to Amazon Web Services (AWS) cloud storage. Amazon Simple Storage Service (Amazon S3) delivers on-demand storage infrastructure, pay-as-you-go pricing, and virtually unlimited scale, so you can tear down data silos to gain business insights, at significantly lower cost.

In this eBook we will demonstrate how to leverage Amazon S3 and its various storage classes to start reducing your storage costs today.

Key features

- Industry-leading scalability, data availability, security, and performance
- Designed for 99.999999999% (11 9's) of durability
- Stores data for millions of applications for companies all around the world
- Wide range of cost-effective storage classes
- Unmatched security, compliance, and audit capabilities

Common use cases

- Data lakes
- Big data analytics
- Backup and restore
- Disaster recovery (DR)
- Archive
- Hybrid cloud storage
- Cloud-native application data

Understanding your true storage TCO

When looking at cloud pricing, it is tempting to directly compare the cloud storage cost against the acquisition cost of storage hardware. However, that does not show you the full picture.
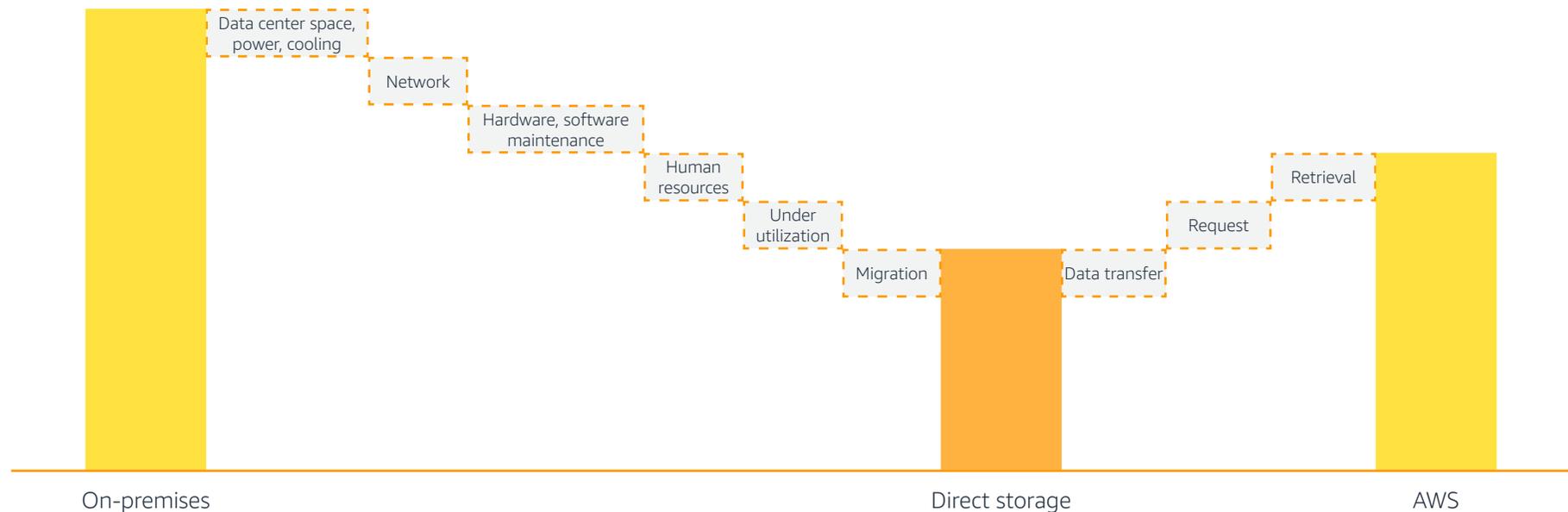

Traditional on-premises storage infrastructure includes several cost drivers you do not incur when running in the cloud. These include the cost of data center space, power, and cooling, the cost of building and maintaining a network, hardware and software maintenance expenses, and the headcount costs to operate this infrastructure.

Additionally, you have to account for the cost of under-utilization (e.g. any storage capacity you purchased but are not actively using). This inherently inflates the cost of data stored. In contrast, when using Amazon S3, as soon as you delete data you stop paying for its underlying storage. Switching to a pay-as-you-go pricing model ultimately allows you to optimize the cost of running your applications.

Amazon S3 is designed for 11 9's of durability

Operational performance is second only to security
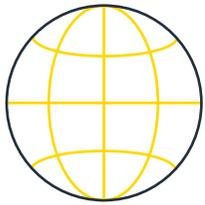
24 Regions

Geographic locations with multiple, physically miles separated and isolated Availability Zones

77 Availability Zones (AZs)

Objects are stored across a minimum of three AZs. Leveraging a multi-AZ approach protects your data in case of an AZ outage.
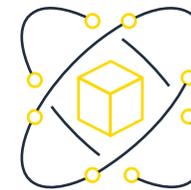
Highly durable

Designed to sustain concurrent device failures and data in the event of an entire AZ loss

AWS pricing principles

No upfront investment

Pay-as-you-go approach

Pay less by using more

Pay less as AWS grows

Decreasing prices and accelerating innovation
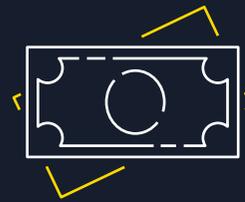
Decreasing storage prices

80% savings

2006   2019

Accelerating innovation

S3 Standard

S3 Glacier

S3 Standard-IA

S3 One Zone-IA

S3 Intelligent-Tiering

S3 GlacierDeep Archive

2006   2019
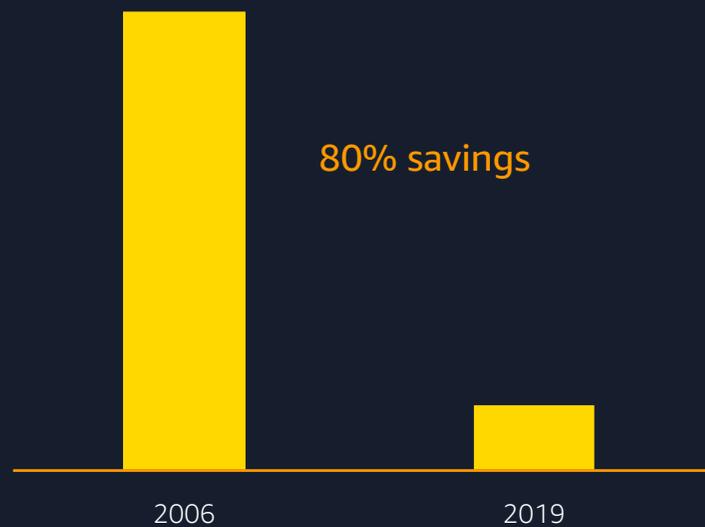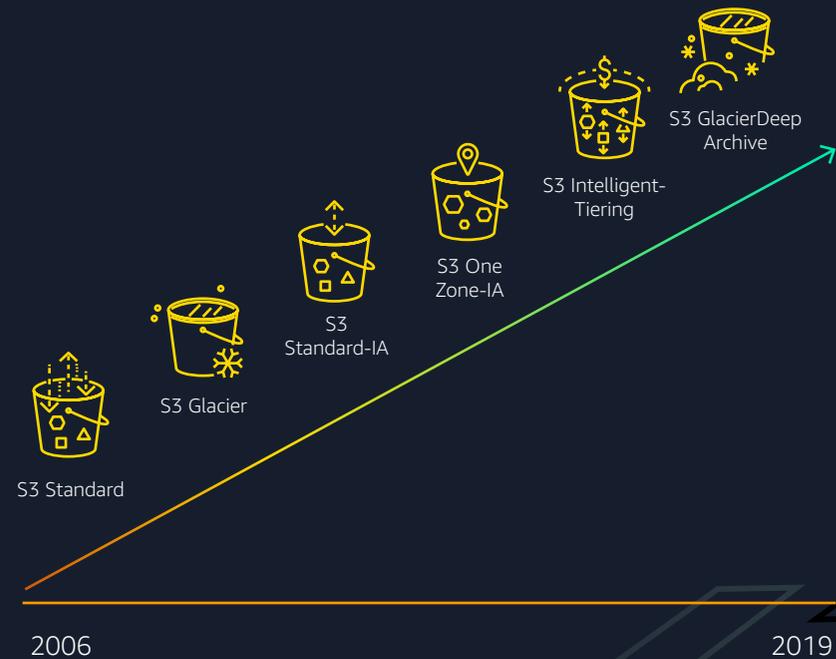
Four pillars of cost optimization for Amazon S3

When it comes to optimizing your storage costs, there are four primary mechanisms for doing so:
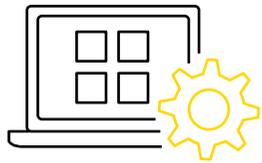
Application requirements

Define your performance and data access needs
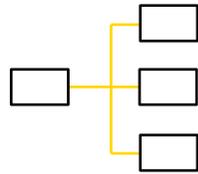
Data organization

Ensure your data have the proper buckets, object tags, and/or prefixes

Storage classes

Determine which storage class best meets your business needs
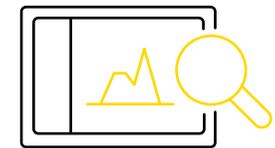
Monitor, analyze, and optimize

Manage your storage to continuously improve TCO

Pillar 1: Defining application requirements

It is important to understand your application and workloads performance and data access requirements, when moving workloads to AWS. If the application is for backup and archive, the needs are significantly different from an e-commerce site, or streaming media services.

Understanding how and when your data is acquired, accessed, and archived or deleted by users is critical to your storage cost management.

Below are some examples around the different types of applications and how their requirements differ.

Static website

Requirements:

- Random data access
- High availability
- High durability

Data analytics application

Requirements:

- Large capacity
- Frequent data access
- Lower availability

Electronic medical records

Requirements

- Long term data retention for 7-10 years
- Infrequent data access
- High durability

Pillar 2: Data organization

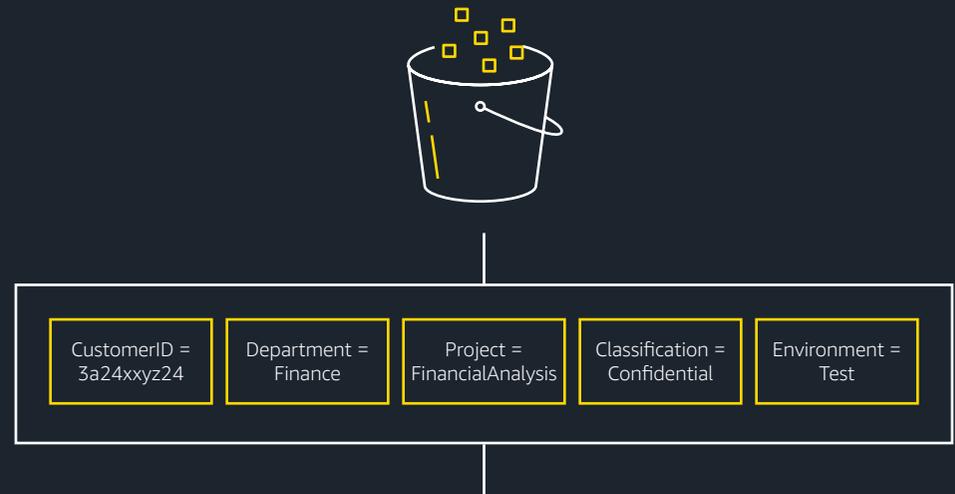As you initially move data to Amazon S3, it is important to organize your data to effectively optimizing cost. Amazon S3 has tools and methods to organize your data on a granular object level which allows you to control access and optimize cost.

This can be done through buckets, object tags, and prefixes:

- You can associate up to 10 tags with an object

- You can add tags to new objects when you upload them, or you can add them to existing objects

- (e.g.) You can grant an Identity and Access Manager (IAM) user permission to read-only objects with specific tags

- Object tags enable fine-grained object lifecycle management, in which you can specify a tag-based filter, in additional to a key name prefix, in a lifecycle rule

When using Amazon S3 Storage Class Analysis, you can configure filters to group objects together for analysis by object tags, key name prefix, or both prefix and tags. You can also customize Amazon CloudWatch metrics to display information by specific tag filters.

Organize data with object tags and prefixes

| CustomerID = 3a24xxyz24 | Department = Finance | Project = FinancialAnalysis | Classification = Confidential | Environment = Test |

- A single bucket can contain objects stored across any storage class
- Control access, analyze usage, manage lifecycle policies, and replicate objects
- Up to 10 mutable metadata tags (key value pair) per object
- Tag objects when created, later, or both

Organize your data from the beginning to enable easy cost optimization

Pillar 3: Choosing the right Amazon S3 Storage class

Amazon S3 offers a range of storage classes designed for different use cases, which support different data access levels at corresponding rates. Picking the correct storage class is a key component of any S3 cost optimization strategy. This enables you to leverage the performance and capacity you need to build highly scalable applications for virtually any use case through the most cost-efficient means.

You can transition data between S3 storage classes using S3 Lifecycle policies, or have Amazon S3 automate data movement and cost savings by using S3 Intelligent-Tiering.

S3 Standard	S3 Intelligent-Tiering	S3 Standard-IA	S3 One Zone-IA	S3 Glacier	S3 GlacierDeep Archive
Frequent		**Access frequency**			**Archive**
• Active, frequently accessed data	• Data with changing access patterns	• Infrequently accessed data	• Re-creatable, less accessed data	• Archive data	• Long-term archive data
• Milliseconds access	• Milliseconds access	• Milliseconds access	• Milliseconds access	• Select minutes or hours	• Select hours
• > 3 AZ	• > 3 AZ	• > 3 AZ	• 1 AZ	• > 3 AZ	• > 3 AZ
	• Monitoring fee per object	• Retrieval fee per GB	• Retrieval fee per GB	• Retrieval fee per GB	• Retrieval fee per GB
	• Min storage duration	• Min storage duration	• Min storage duration	• Min storage duration	• Min storage duration
		• Min object size	• Min object size	• Min object size	• Min object size

Frequent data access

Amazon S3 Standard (S3 Standard)

Amazon S3 Standard is the most common storage class, designed to provide highly durable, available, and performant object storage for frequently accessed data . Its broad usage encompasses use cases such as dynamic websites, content distribution, mobile and gaming applications, and big data analytics.

- Low latency and high throughput performance

- Designed for durability of 99.999999999% of objects across multiple Availability Zones

- Resilient against events that impact an entire Availability Zone

- Designed for 99.99% availability over a given year

- Backed with the Amazon S3 Service Level Agreement for availability

- Supports SSL for data in transit and encryption of data at rest

- S3 Lifecycle management for automatic migration of objects to other S3 Storage Classes

Unknown or changing data access

Amazon S3 Intelligent-Tiering (S3 Intelligent-Tiering)

S3 Intelligent-Tiering enables you to take a hands-off approach to data movement between storage classes. If data has not been accessed for 30 days, it will automatically be moved into the infrequent access tier without retrieval or tiering fees, providing immediate cost-savings (there is a small monthly monitoring and auto-tiering fee). Conversely, if data in the infrequently access tier is accessed, it will then be moved back into the frequent access tier.

- Same low latency and high throughput performance of S3 Standard
- Automatically moves objects between two access tiers based on changing access patterns
- S3 Lifecycle management for automatic migration of objects to other S3 Storage Classes

Amazon S3 Intelligent-Tiering

Optimizes costs by automatically moving data to the most cost-effective access tier.

S3 Intelligent-Tiering frequent access tier

When data is added to S3 Intelligent -Tiering, it starts out in the Frequent Access tier.

Access patterns are monitored to automate object movement between access tiers.

3 Intelligent-Tiering infrequent access tier

Objects not accessed for 30 consecutive days, are moved to the Infrequent Access tier.

Infrequent data access

Amazon S3 Standard-Infrequent Access (S3 Standard-IA)

S3 Standard-IA is for data that is less frequently accessed, but still requires rapid access when needed. This makes it ideal for reducing costs, while maintaining the high performance of Amazon S3. It is most commonly used for long-term storage, backups, and as a data store for disaster recovery files.

- Same low latency and high throughput performance of S3 Standard
- S3 Lifecycle management for automatic migration of objects to other S3 Storage Classes

Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA)

This storage class is near-identical to S3 Standard-IA. However, while other S3 Storage Classes store data in a minimum of three AZs, S3 One Zone-IA only stores data in a single AZ. This is valuable for storing secondary backup copies of on-premises data or data that can easily be recreated, as well as data replication from another AWS Region using S3 Cross-Region Replication.

- Performance of Amazon S3 Standard-IA stored in a single AZ at 20% of the cost
- Same low latency and high throughput performance of S3 Standard
- S3 Lifecycle management for automatic migration of objects to other S3 Storage Classes

Data archive

Amazon S3 Glacier (S3 Glacier)

Ideal for backup and archive, S3 Glacier provides secure, durable, and low-cost storage that are competitive with or cheaper than on-premises solutions. Objects can be uploaded directly to S3 Glacier, or moved into the service using S3 Lifecycle policies from one of the other classes.

- Low-cost design is ideal for backup and archives with varying access needs
- Configurable retrieval times, ranging from a few minutes to hours
- S3 PUT API for direct uploads to S3 Glacier, and S3 Lifecycle management for automatic migration of objects

Amazon S3 Glacier Deep Archive (S3 Glacier Deep Archive)

S3 Glacier Deep Archive is the lowest-cost S3 Storage Class, designed for long-term data retention. It caters toward organizations in highly-regulated industries, such as Financial Services, Healthcare, and the Public Sector, that are required to retain data for seven or more years to meet compliance requirements. It is commonly used for backup, disaster recovery, and a cost-effective alternative to magnetic tape systems (both on- and off-premises services).

- Lowest cost storage class of approximately $1 per terabyte/month
- Designed for long-term data retention (7-10 years), accessed 1-2 times a year
- Optimal alternative to magnetic tape libraries
- Retrieval time within 12 hours
- S3 PUT API for direct uploads to S3 Glacier Deep Archive, and S3 Lifecycle management for automatic migration of objects

Amazon S3 Storage Class performance chart

	S3 Standard	S3 Intelligent-Tiering*	S3 Standard-IA	S3 One Zone-IA†	S3 Glacier	S3 Glacier Deep Archive
Designed for durability	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)	99.999999999% (11 9's)
Designed for availability	99.99%	99.9%	99.9%	99.5%	99.99%	99.99%
Availability SLA	99.9%	99%	99%	99%	99.9%	99.9%
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	milliseconds	milliseconds	milliseconds	select minutes or hours	select hours

† Because S3 One Zone-IA stores data in a single AWS Availability Zone, data stored in this storage class will be lost in the event of Availability Zone destruction.

* S3 Intelligent-Tiering charges a small tiering fee and has a minimum eligible object size of 128KB for auto-tiering. Smaller objects may be stored but will always be charged at the Frequent Access tier rates. See the Amazon S3 Pricing for more information.

Pillar 4: Monitor, analyze, and optimize

Monitor

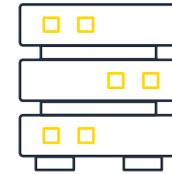with Amazon S3 inventory, Amazon CloudWatch, S3 server access logging

Understand access patterns

with Amazon S3 storage class analysis

Perform API actions

across thousands, millions, or billions of objects with Amazon S3 Batch Operations

Tier and expire storage

manually with Amazon S3 lifecycle policies

or

automatic cost optimization with S3 Intelligent-Tiering

1. Monitor

By monitoring your S3 usage, you can lower storage costs and growth over time. AWS Budgets gives you the ability to set custom budgets that alert you when your costs or usage exceed (or are forecasted to exceed) your budgeted amount. You can also use Amazon CloudWatch metrics to monitor your storage and request activity growth in real time and trigger an alarm when you reach a threshold.

2. Analyze

Effectively determining the best S3 Storage Class requires knowledge of your data access patterns. Amazon S3 Storage Class Analysis provides you with insights into your prior storage access patterns, so you can make informed choices about which S3 Storage Class is the best fit for your data.

3. Optimize

To ensure your objects are stored in the most cost-efficient S3 Storage Class throughout their lifecycle, you can configure S3 Lifecycle Policies. Lifecycle configurations are rule sets that define what actions Amazon S3 will apply to a group of objects; such as whether you will transition objects into colder storage class (e.g. Amazon S3 Glacier) or automatically delete objects that are no longer needed.

Alternatively, you can use S3 Intelligent-Tiering to automate the movement of your objects between storage tiers based on their access history.

Optimization withS3's building blocks

S3 Standard **S3 Storage Class Analysis** **S3 Lifecycle** **S3 Standard-IA**

- Storage Class Analysis classifies data as frequently or infrequently accessed
- Great for predictable workloads with S3 lifecycle
- Can be filtered by bucket, prefix, or object tag

Automatic optimization with S3 Intelligent-Tiering

S3 Intelligent-Tiering **Amazon CloudWatch** **S3 Inventory (Access Tiers)**

- S3 Intelligent-Tiering stores objects in two access tiers – frequent access and infrequent access
- Great for unpredictable or unknown access patterns
- CloudWatch shows bytes by tier and S3 Inventory

Amazon S3 tools to help you optimize storage costs

Amazon S3 has various features you can use to analyze and organize your data in ways that enable cost efficiencies, letting you optimize your storage for access patterns and cost.

Amazon S3 Lifecycle Management

To manage your objects so that they are stored cost effectively throughout their lifecycle, use Amazon S3 Lifecycle. An S3 Lifecycle configuration is a set of rules that define actions that Amazon S3 applies to a group of objects, you can either transition to another storage class, or delete expired objects

Amazon S3 Storage Class Analysis

Use S3 Storage Class Analysis to analyze data access patterns to help you decide when to transition the right data to the right storage class. After using S3 Storage Class Analysis to monitor access patterns, you can use this information to configure S3 Lifecycle policies to make the data transfer to the appropriate storage class.

Amazon S3 Pricing and TCO Calculators

Configure a cost estimate that fits your unique business needs with Amazon S3 by using the pricing calculator. To compare the cost of running your applications in an on-premises or colocation environment to AWS, use the AWS Total Cost of Ownership (TCO) calculator.

Putting it all together

Understand your application requirements
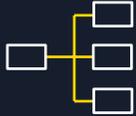
Use tags and prefixes to organize your data

Cost optimize across all storage classes (per object, tag, prefix)

Use S3 building blocks to monitor, analyze, and manage storage

Store in S3 Intelligent-Tiering for automated cost savings

Archive to S3 Glacier Deep Archive for the lowest storage price

Getting Started: Start leveraging these best practices to lower your Amazon S3 storage costs

For additional insights into cost optimization using Amazon S3, **visit our website**. There you will find videos highlighting the various storage classes, a breakdown of Amazon S3 Intelligent-Tiering, and customer case studies.

Additional Amazon S3 cost optimization resources

Webinar: Cost optimization guidelines for Amazon S3

Go more in-depth on the pillars of cost optimization, how to optimize costs through data placement, how to most effectively leverage Amazon S3 Intelligent-Tiering, and the optimal design patterns for your Amazon S3 storage.

Watch now »

re:Invent: Guidelines and design patterns for optimizing costs in Amazon S3

Take a deeper dive into cost-optimization guidelines, design patterns, and discover how various customers have been able to improve storage cost efficiency.

Watch now »

Webinar: Amazon S3 Glacier Deep Archive

Take a deep dive into Amazon S3 Glacier Deep Archive for best practices on how to leverage the lowest cost object storage available on the cloud.

Watch now »

Infographic: Optimize costs with S3 storage classes

Learn how to manage and reduce storage costs by using the S3 storage classes.

Learn more »