AWS re:Invent

NOV. 29 – DEC. 3, 2021 | LAS VEGAS, NV

STG347

Best practices for cost optimization with Amazon S3

Alex Crabbe (he/him)
Principal Product Manager
Amazon S3

Jessie Felix (he/him)
Senior Product Manager, Technical
Amazon S3

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

Define workload requirements

Organize your data

Analyze, act, and measure

S3 Intelligent-Tiering for automatic storage cost savings

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS storage customers



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How are customers using Amazon S3?

Compliance records

Analytics

Geospatial or lunar imagery

IoT sensor data

Medical images and records

Data lakes

Customer call center records

Media primary files

Digital record preservation

Home-recording video

Mobile sync and storage



Origin storage for CDN

Seismic and reservoir simulation data

Pharmaceutical study data

Durable backups

DNA sequences

Surveillance video/closed-circuit television

ML training data

Media assets

Financial transaction records

Website hosting

Log files

Meteorological and environmental research

User-generated content

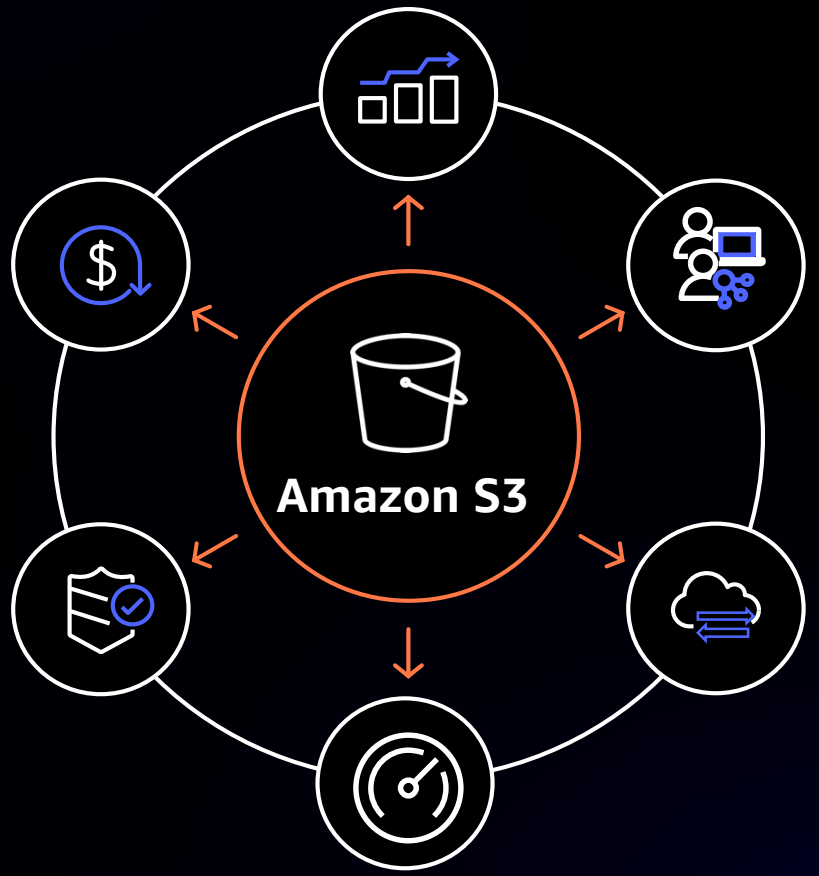
Autonomous vehicle data

Oil and gas topography

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

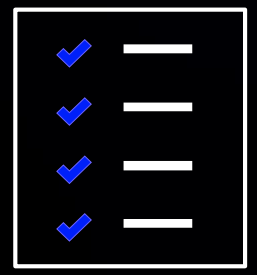
Operating at scale in Amazon S3



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

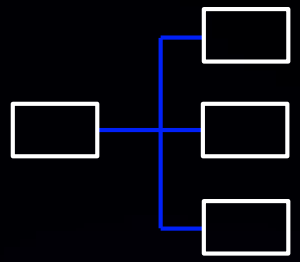
Steps to optimizing storage costs using Amazon S3

1



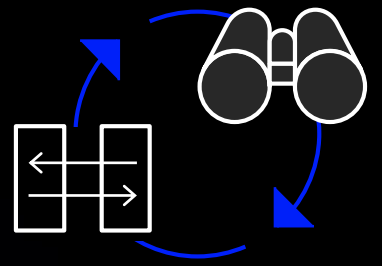
Define your workload requirements

2



Organize your data

3

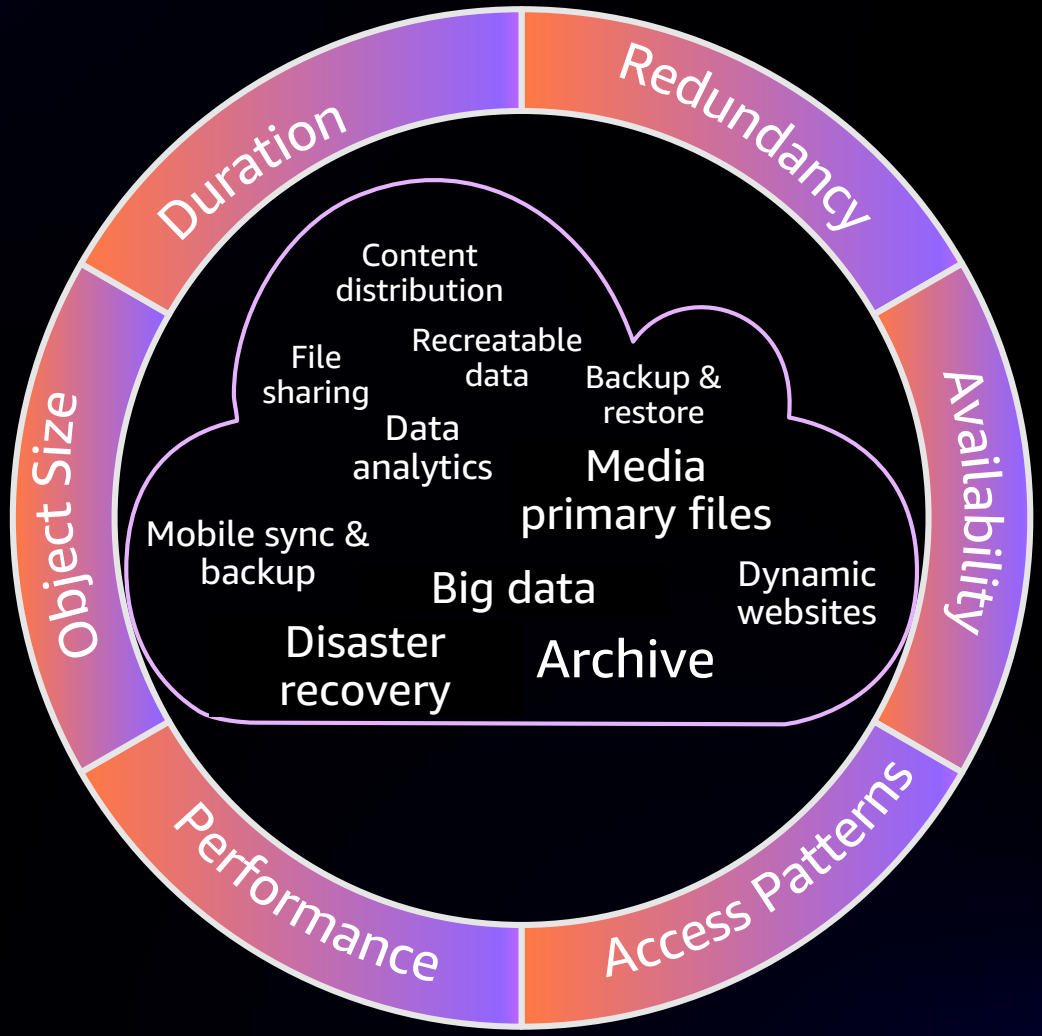


Analyze, act, and measure

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

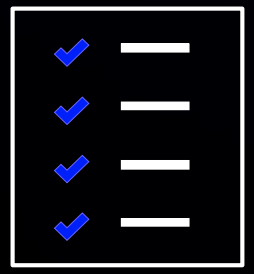
Define your workload requirements



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

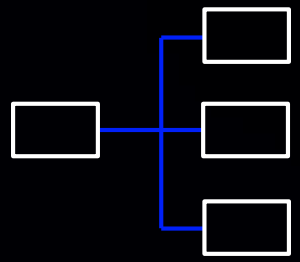
Steps to optimizing storage costs using Amazon S3

1



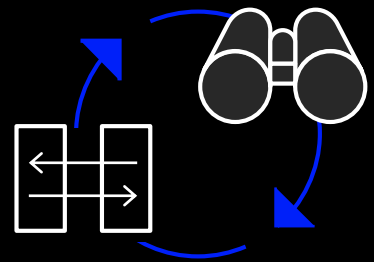
Define your workload requirements

2



Organize your data

3



Analyze, act, and measure

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Organize your data and manage at scale

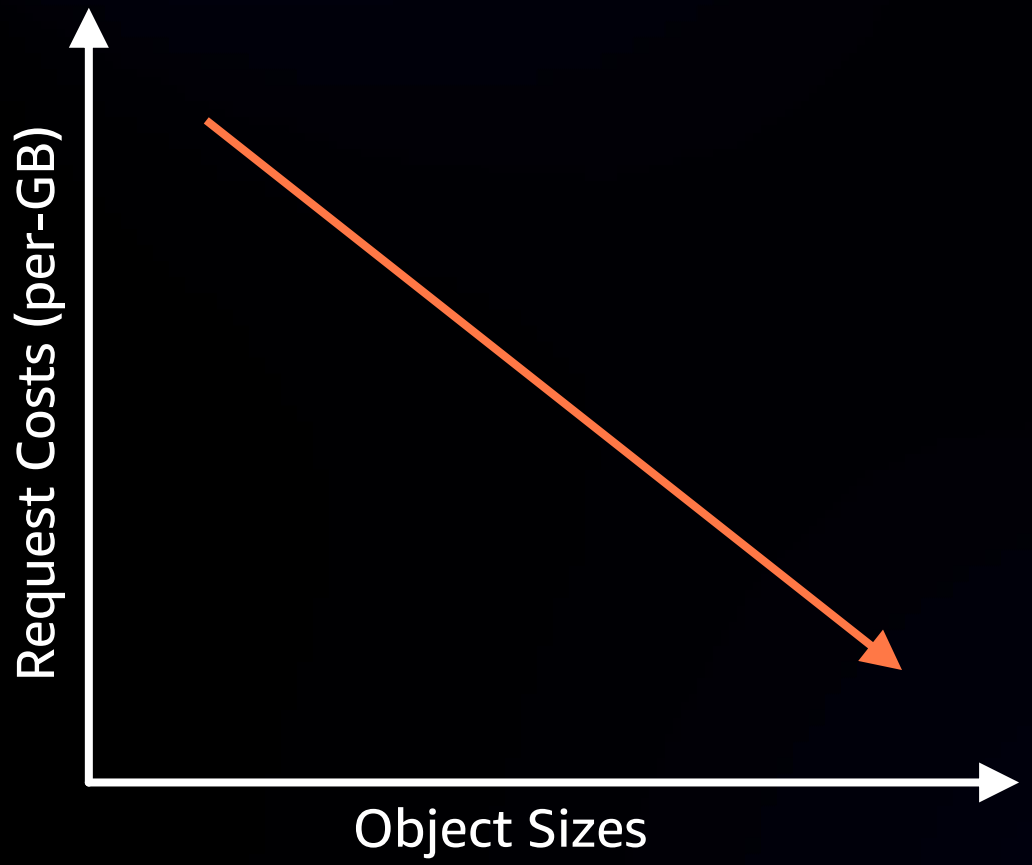


Account

Buckets	Prefixes	Object tags
Business Unit 1	Unit1/internal/usage/ Unit1/internal2/billing/usage Unit1/internal3/free/usage	Department = Finance Team = Business Reporting Classification = Confidential
Business Unit 2	Unit2/model1/inference Unit2/model2/inference Unit2/predictions/prod	Department = Marketing Team = Data Science Classification = Confidential

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Relationship between object size and total request costs



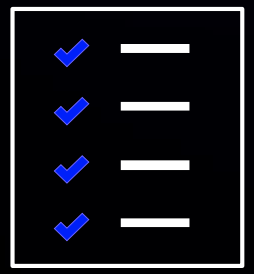
- Batch data on ingest
- Use tools like Amazon EMR to batch existing data into larger objects

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

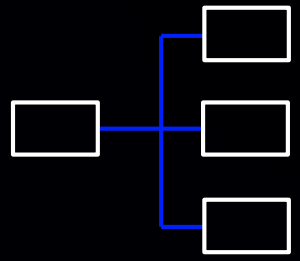
Steps to optimizing storage costs using Amazon S3

1



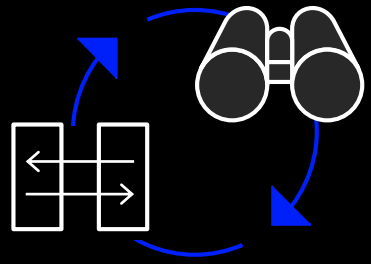
Define your workload requirements

2



Organize your data

3

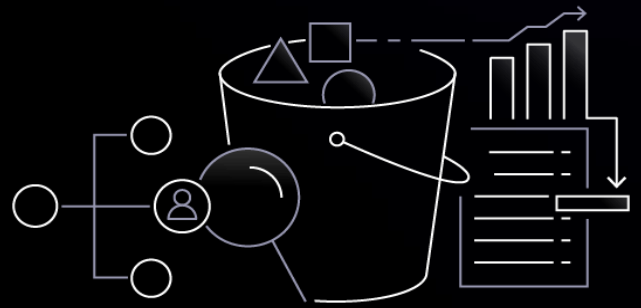


Analyze, act, and measure

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

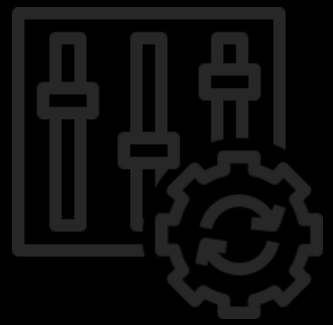
Analyze your Amazon S3 storage

S3 Storage Lens



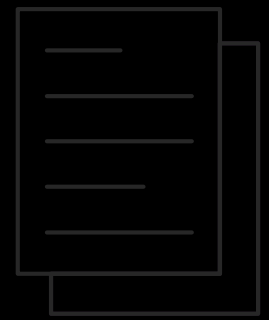
Organization-wide S3 usage and activity dashboard for cost optimization

Storage Class Analysis



Bucket-level analysis of retrievals for predictable workloads

S3 Inventory



Object-level analysis for analytics and auditing

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon S3 Storage Lens overview

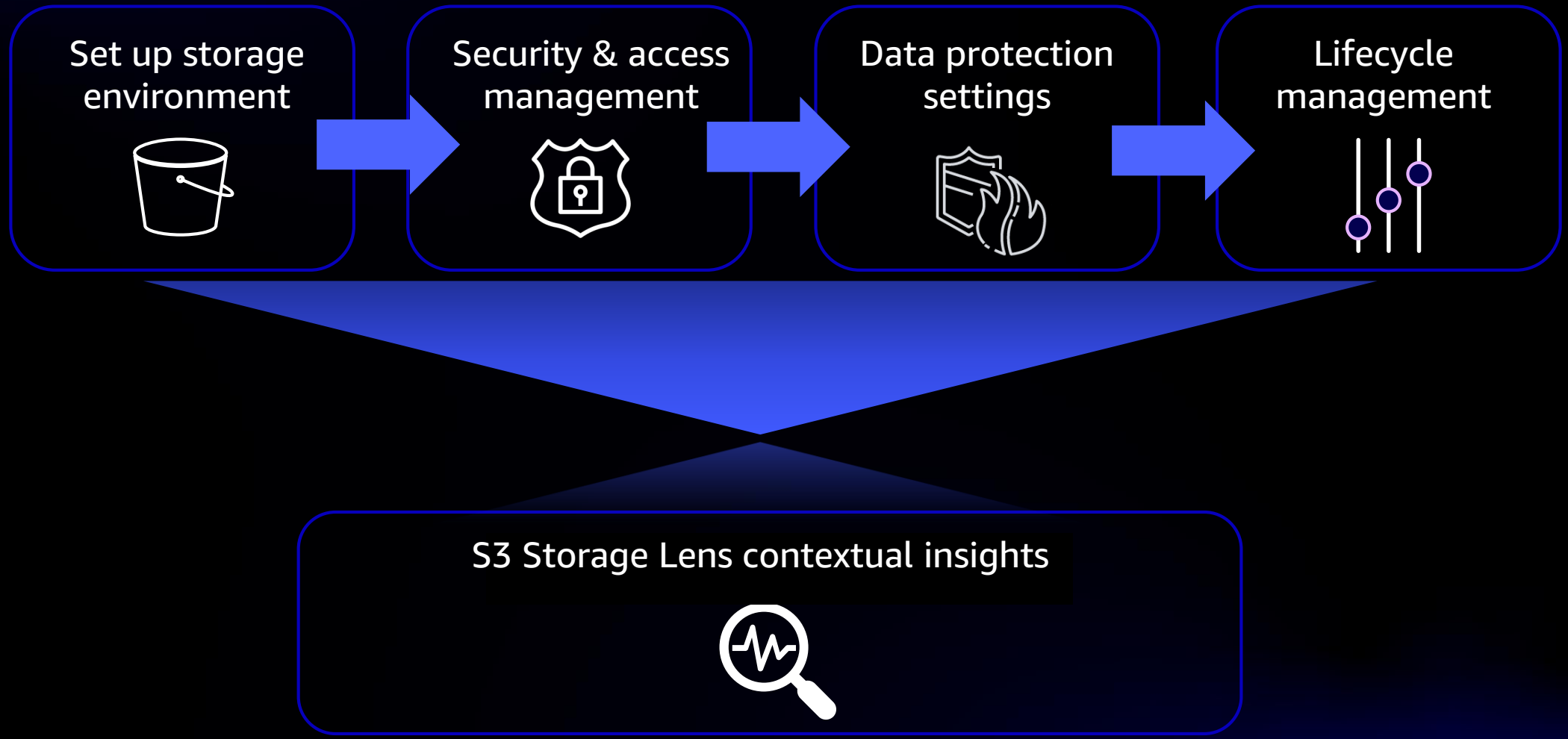


Launched November 2020

- Organization-wide visibility
- 29 storage usage & activity metrics
- Drill down by Region, storage class, bucket, and prefix
- Recommendations for cost efficiency & data protection best practices
- Integrated S3 console experience

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

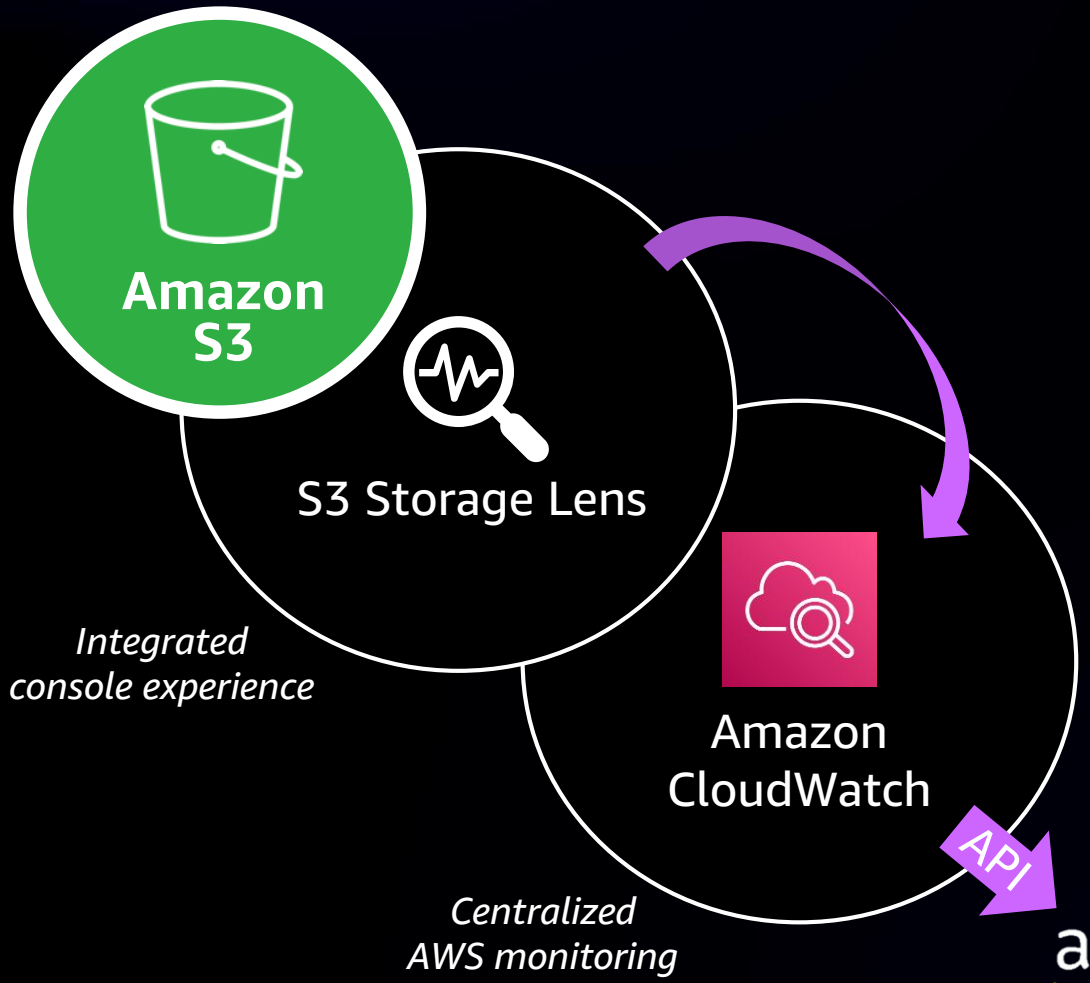
Contextually integrated experience



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Introducing Amazon CloudWatch publishing for S3 Storage Lens

New



- Unlock capabilities of CloudWatch for S3 Storage Lens metrics
  - Centralized, cross-service dashboards
  - Alarms and triggered actions
  - Anomaly detection
- Access metrics via CloudWatch API
  - Connect AWS observability partners

aws partner network

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why we built Amazon CloudWatch publishing for S3 Storage Lens

New



Customize dashboard expressions to get new metrics and chart styles



Configure alarms on S3 Storage Lens metrics to trigger SNS notifications



Build custom monitoring applications using S3 Storage Lens APIs

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

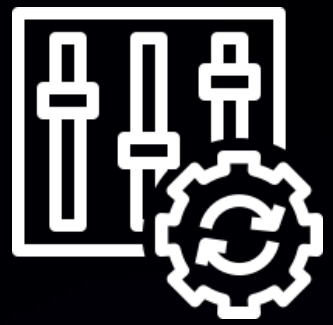
Analyze your S3 storage

S3 Storage Lens



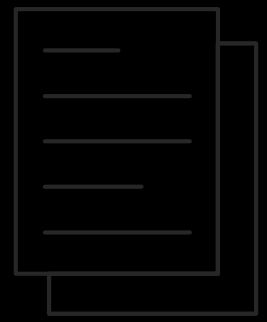
Organization-wide S3 usage and activity dashboard for cost optimization

Storage Class Analysis



Bucket-level analysis of retrievals for predictable workloads

S3 Inventory



Object-level analysis for analytics and auditing

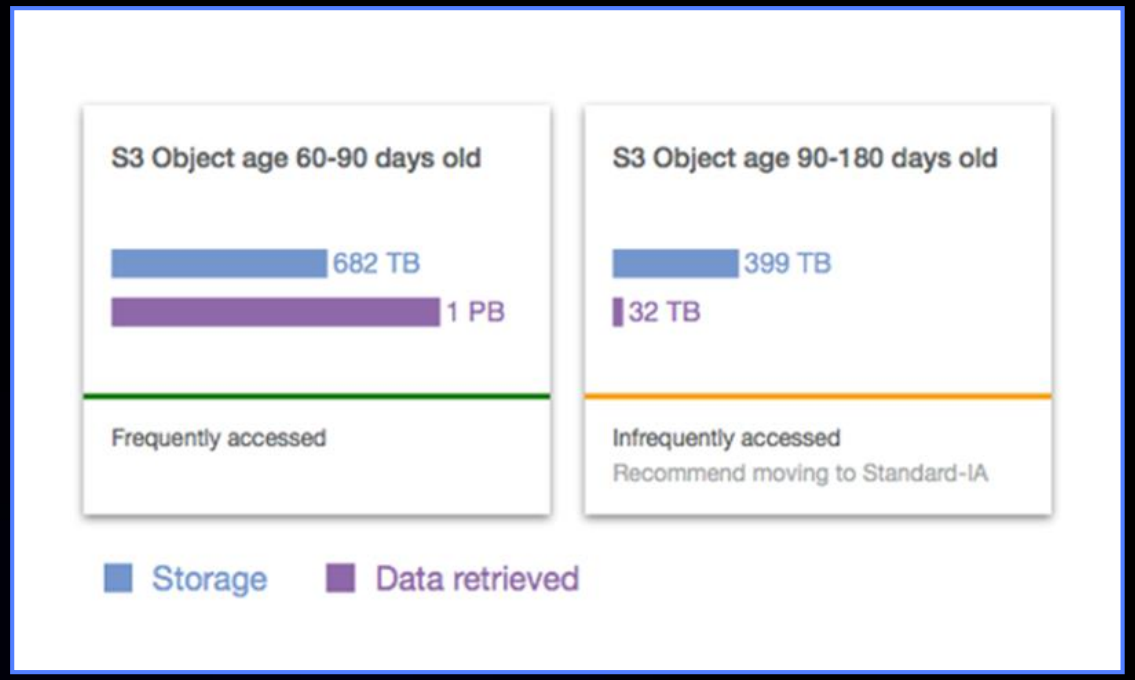
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

S3 Storage Class Analysis for known or predictable access patterns

Monitors access patterns

Classifies data as frequently or infrequently accessed

Can be filtered by bucket, prefix, or object tag



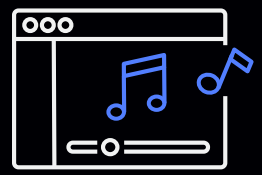
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

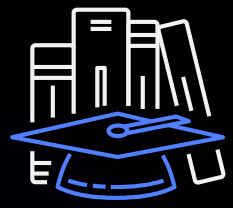
Data with predictable access patterns



Medical records



Media streaming



Learning resources



Backups

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

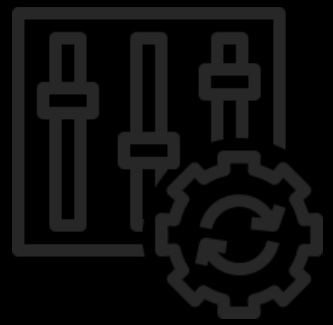
Analyze your S3 storage

S3 Storage Lens



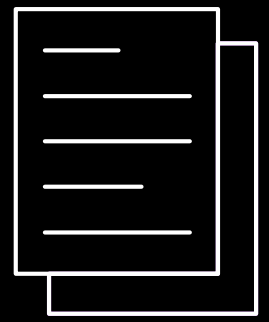
Organization-wide S3 usage and activity dashboard for cost optimization

Storage Class Analysis



Bucket-level analysis of retrievals for predictable workloads

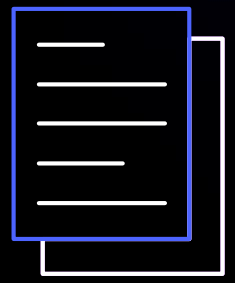
S3 Inventory



Object-level analysis for analytics and auditing

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Using Amazon Athena and S3 Inventory for object-level analysis



Inventory report





Amazon Athena

Use SQL query on your S3 Inventory list

No Minimum
Capacity Charge Per Object

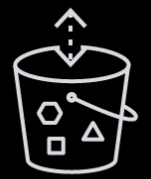


S3 Standard



S3 Intelligent-Tiering

128 KB Minimum
Capacity Charge Per Object



S3 Standard-IA



S3 One Zone-IA

40 KB Minimum
Capacity Charge Per Object



S3 Glacier Flexible Retrieval (formerly S3 Glacier)

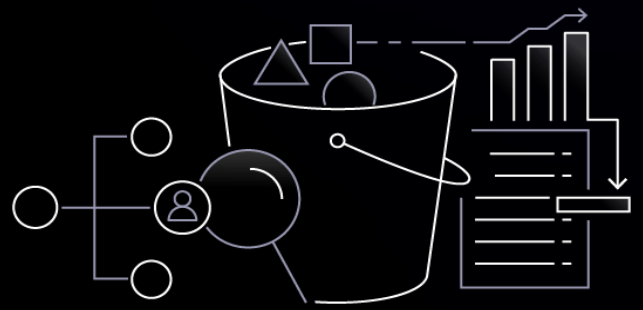


S3 Glacier Deep Archive

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

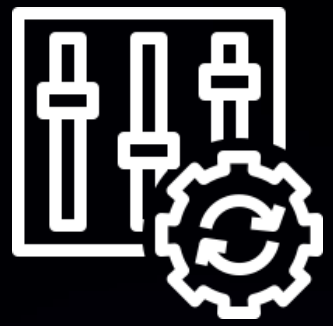
Analyze your S3 storage

S3 Storage Lens



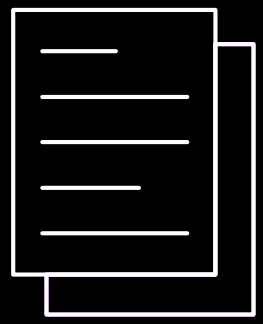
Organization-wide S3 usage and activity dashboard for cost optimization

Storage Class Analysis



Bucket-level analysis of retrievals for predictable workloads

S3 Inventory



Object-level analysis for analytics and auditing

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Define your **performance requirements**
- Organize and categorize data to **match your business logic**
- Aggregate small objects to **reduce request costs**
- Use Storage Lens APIs to **monitor anomalies**

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

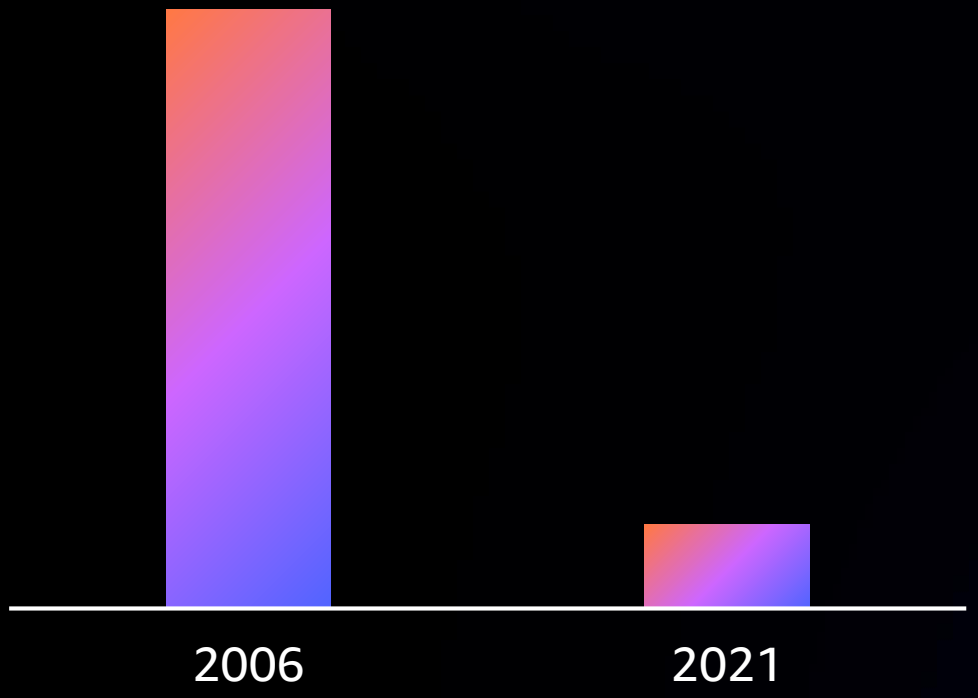
Using S3 storage classes

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

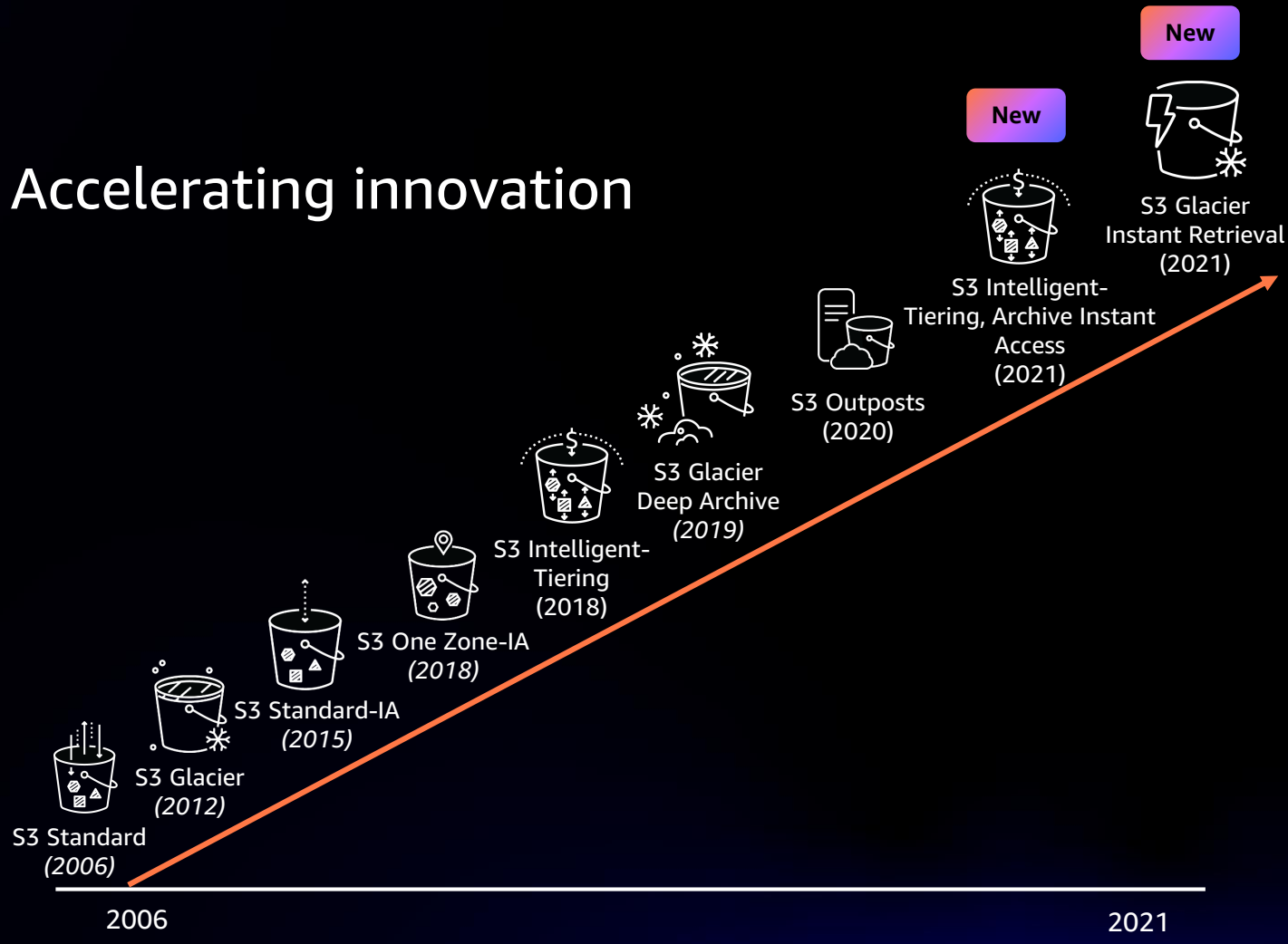
Amazon S3 storage classes

OPTIMIZE YOUR STORAGE COST BY UTILIZING ALL AMAZON S3 STORAGE CLASSES

Decreasing storage prices



Accelerating innovation



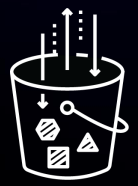
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

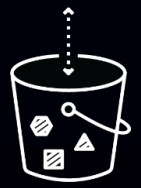
Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering

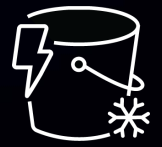


S3 Standard



S3 Standard-IA

New



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

AWS Region ≥ 3 Availability Zones

Data with changing access patterns

- Milliseconds access
- No retrieval fees
- Object monitoring charge
- Opt in Async Archive tiers
- **Archive Instant Access tier** New

Frequently accessed data

- Milliseconds access

Infrequently accessed data

- Milliseconds access
- Minimum storage duration
- Minimum object size

Rarely accessed data

- **Milliseconds Access**
- **Minimum storage duration**
- **Minimum object size**

Archive data

- Retrieval options from minutes to hours
- Minimum storage duration
- Minimum object size
- **Free bulk retrievals** New

Long term archive data

- Retrieval in hours
- Minimum storage duration
- Minimum object size

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

On-premises data

- Milliseconds access
- Encrypted with SSE-S3

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA

New



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

Data with changing access patterns

- Milliseconds access
- No retrieval fees
- Object monitoring charge
- Opt in Async Archive tiers
- **Archive Instant Access tier** New

AWS Region ≥ 3 Availability Zones

Frequently accessed data

- Milliseconds access

Infrequently accessed data

- Milliseconds access
- Minimum storage duration
- Minimum object size

Rarely accessed data

- **Milliseconds Access**
- **Minimum storage duration**
- **Minimum object size**

Archive data

- Retrieval options from minutes to hours
- Minimum storage duration
- Minimum object size
- **Free bulk retrievals** New

Long term archive data

- Retrieval in hours
- Minimum storage duration
- Minimum object size

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

On-premises data

- Milliseconds access
- Encrypted with SSE-S3

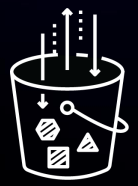
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

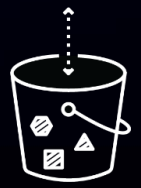
Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA

New



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

AWS Region ≥ 3 Availability Zones

Data with changing access patterns	Frequently accessed data	Infrequently accessed data	Rarely accessed data	Archive data	Long term archive data
• Milliseconds access • No retrieval fees • Object monitoring charge • Opt in Async Archive tiers • **Archive Instant Access tier** New	• Milliseconds access	• Milliseconds access • Minimum storage duration • Minimum object size	• **Milliseconds Access** • **Minimum storage duration** • **Minimum object size**	• Retrieval options from minutes to hours • Minimum storage duration • Minimum object size • **Free bulk retrievals** New	• Retrieval in hours • Minimum storage duration • Minimum object size

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

On-premises data

- Milliseconds access
- Encrypted with SSE-S3

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA

New



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

AWS Region ≥ 3 Availability Zones

Data with changing access patterns	Frequently accessed data	Infrequently accessed data	Rarely accessed data	Archive data	Long term archive data
• Milliseconds access • No retrieval fees • Object monitoring charge • Opt in Async Archive tiers • **Archive Instant Access tier** New	• Milliseconds access	• Milliseconds access • Minimum storage duration • Minimum object size	• **Milliseconds Access** • **Minimum storage duration** • **Minimum object size**	• Retrieval options from minutes to hours • Minimum storage duration • Minimum object size • **Free bulk retrievals** New	• Retrieval in hours • Minimum storage duration • Minimum object size

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

On-premises data

- Milliseconds access
- Encrypted with SSE-S3

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering

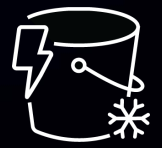


S3 Standard



S3 Standard-IA

New



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

AWS Region & 3 Availability Zones

Data with changing access patterns

- Milliseconds access
- No retrieval fees
- Object monitoring charge
- Opt in Async Archive tiers
- **Archive Instant Access tier** New

Frequently accessed data

- Milliseconds access

Infrequently accessed data

- Milliseconds access
- Minimum storage duration
- Minimum object size

Rarely accessed data

- **Milliseconds Access**
- **Minimum storage duration**
- **Minimum object size**

Archive data

- Retrieval options from minutes to hours
- Minimum storage duration
- Minimum object size
- **Free bulk retrievals** New

Long term archive data

- Retrieval in hours
- Minimum storage duration
- Minimum object size

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per GB
- Minimum storage duration
- Minimum object size

AWS Outposts

On-premises data

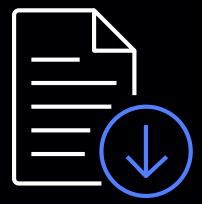
- Milliseconds access
- Encrypted with SSE-S3

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Why we built S3 Glacier Instant Retrieval



Archive data stored for several years



Very low rate of access



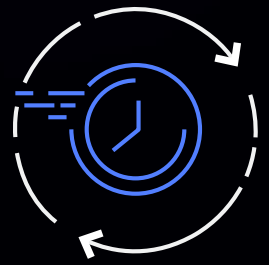
Immediately accessible when requested

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

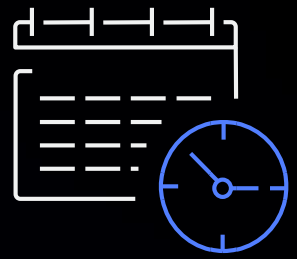
Introducing S3 Glacier Instant Retrieval



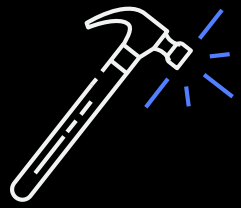
$0.004 per GB-month



Milliseconds access



Designed for 99.9% availability



Designed for 99.999999999% durability

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lifecycle data with predictable access patterns



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use lifecycle policies to transition objects to another storage class

Lifecycle rules take action based on object age; example:

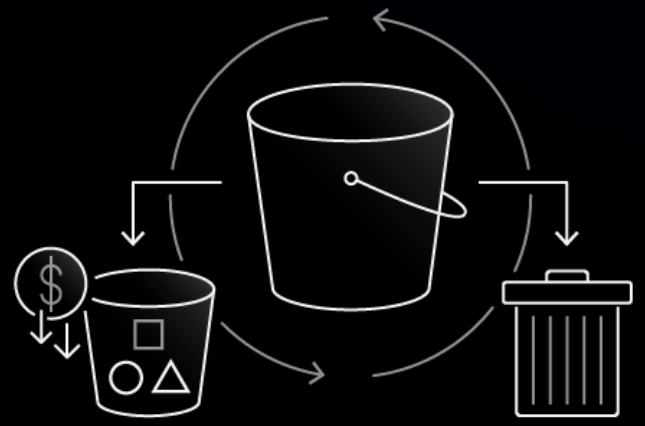
1. Move objects older than 90 days to S3 Glacier Instant Retrieval
2. Move objects older than 365 days to S3 Glacier Deep Archive



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

S3 Lifecycle filters and actions

New

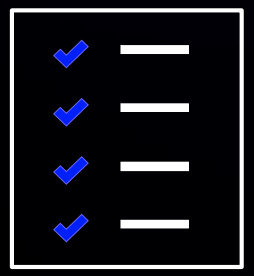


- Even **more granular** S3 Lifecycle configurations
- **Object size filters** to optimize transition costs and save more on storage
- **Number of noncurrent versions** to optimize storage spend for versioned buckets

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

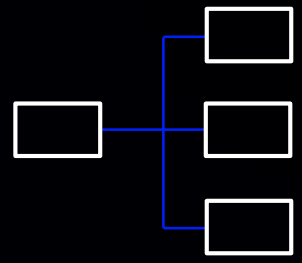
Steps to optimizing storage costs using Amazon S3

1



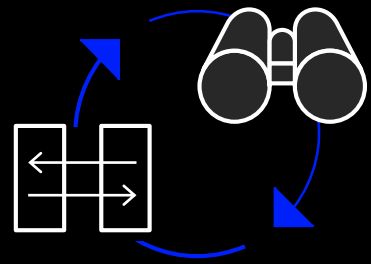
Define your workload requirements

2



Organize your data

3



Analyze, act, and measure

aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Use S3 Lifecycle policies and storage class building blocks to optimize data with **known** or **predictable** **access patterns**
- Fine-tune your S3 Lifecycle policies to **optimize transition request costs**
- Use S3 Glacier Instant Retrieval for datasets accessed **once per quarter**

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

S3 Intelligent-Tiering

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Since the launch of S3 Intelligent-Tiering, customers storage cost savings now exceed

$250,000,000

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Since the launch of S3 Intelligent-Tiering, customers storage cost savings now exceed

$250,000,000

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What is S3 Intelligent-Tiering?

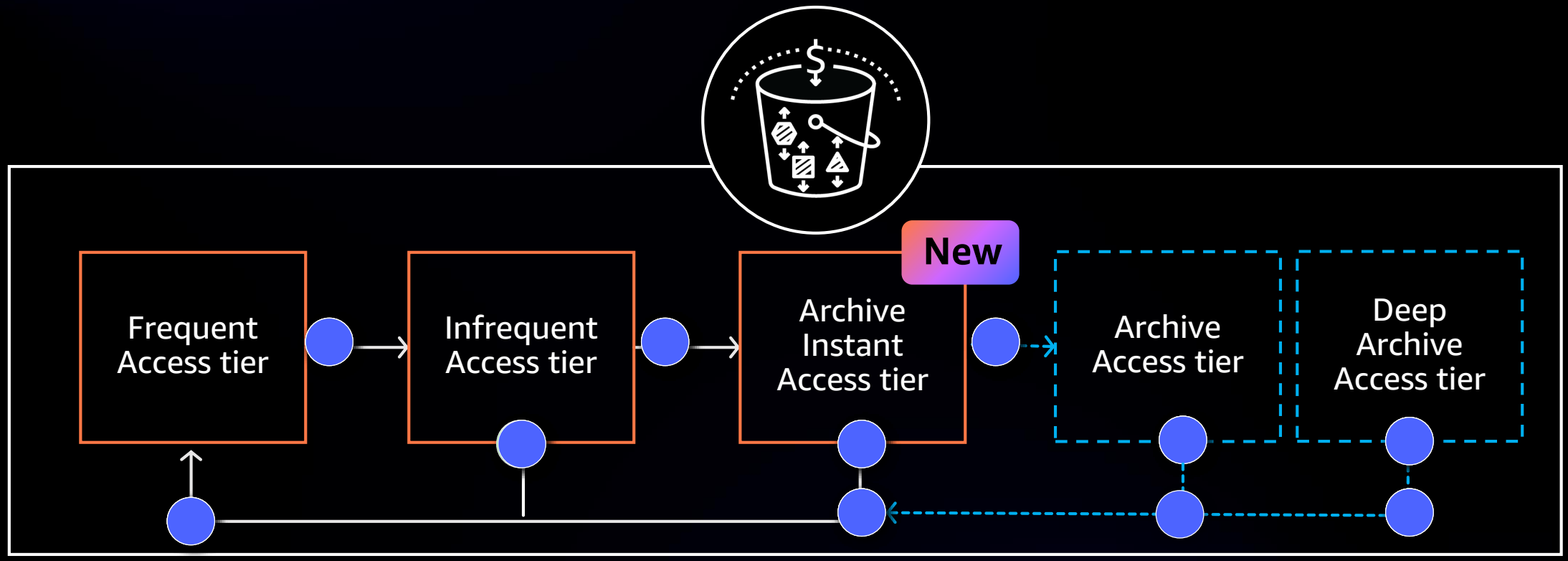


New

- Only cloud storage that **delivers automatic storage cost savings**
- Moves objects between **three access tiers** for a small monthly monitoring and automation fee
- **New Archive Instant Access tier** delivers up to 68% lower cost, **without any impact on performance**
- No operational overhead, no lifecycle fees, and no retrieval fees
- Designed for 99.9% availability and 99.999999999% (11 9s) of durability

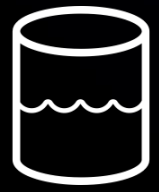
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Use S3 Intelligent-Tiering for data with unknown or changing access patterns



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

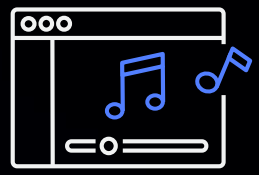
Data with unknown access patterns



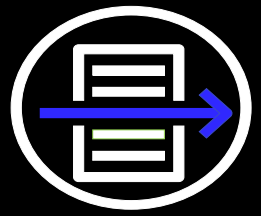
Data Lakes



Data Analytics



Content Sharing

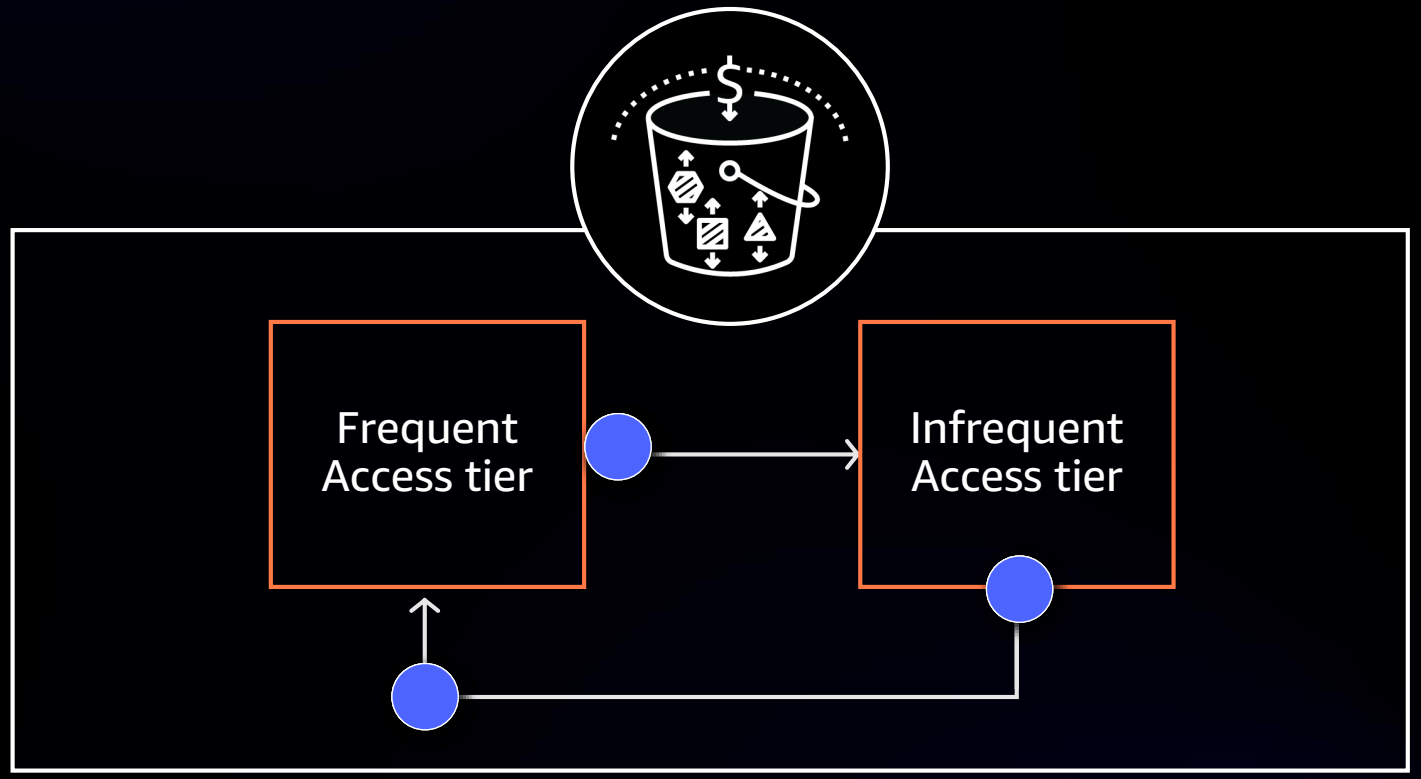


Long Tail Data

ACROSS EVERY INDUSTRY

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

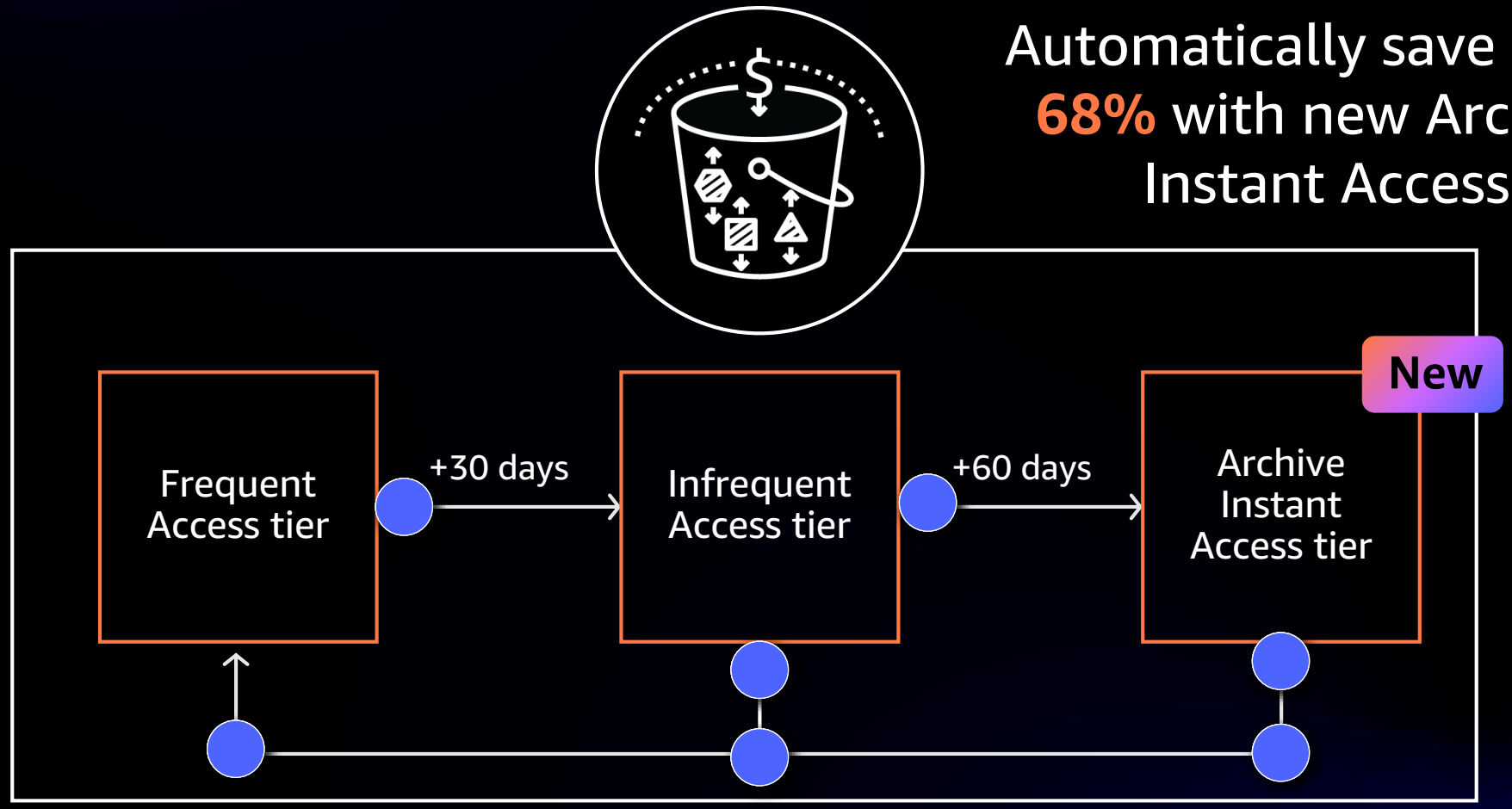
Before: S3 Intelligent-Tiering automatically optimized cost in two access tiers



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Now: S3 Intelligent-Tiering automatically optimizes cost in three access tiers

Automatically save up to **68%** with new Archive Instant Access



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

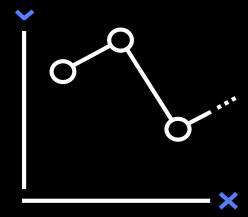
Why we built the Archive Instant Access tier



Petabytes of data stored for indefinite periods of time



Subsets of data become rarely accessed for long periods of time



Data needs millisecond retrievals when access patterns change

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

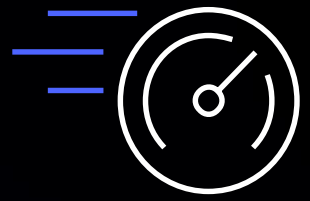
S3 Intelligent-Tiering: Archive Instant Access



Automatic savings for rarely accessed data



68% lower cost than the Infrequent Access tier



Millisecond access and high throughput performance

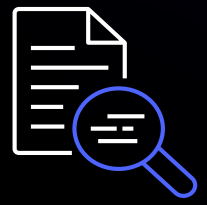


Designed for 99.999999999% durability

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

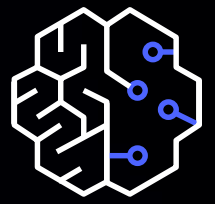
S3 Intelligent-Tiering pricing enhancements

New



Small objects not monitored or auto-tiered

Use S3 Intelligent-Tiering without needing to analyze object size distributions



Optimizing for short-lived objects

Use S3 Intelligent-Tiering without needing to analyze what the average life of objects

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Upload directly to S3 Intelligent-Tiering

Specify INTELLIGENT_TIERING in the PUT API request header

```
PUT /my-image.jpg HTTP/1.1
Host: myBucket.s3.<Region>.amazonaws.com (http://amazonaws.com/)
Date: Wed, 1 Dec 2021 17:50:00 GMT
Authorization: authorization string
Content-Type: image/jpeg
Content-Length: 11434
Expect: 100-continue
x-amz-storage-class: INTELLIGENT_TIERING
```

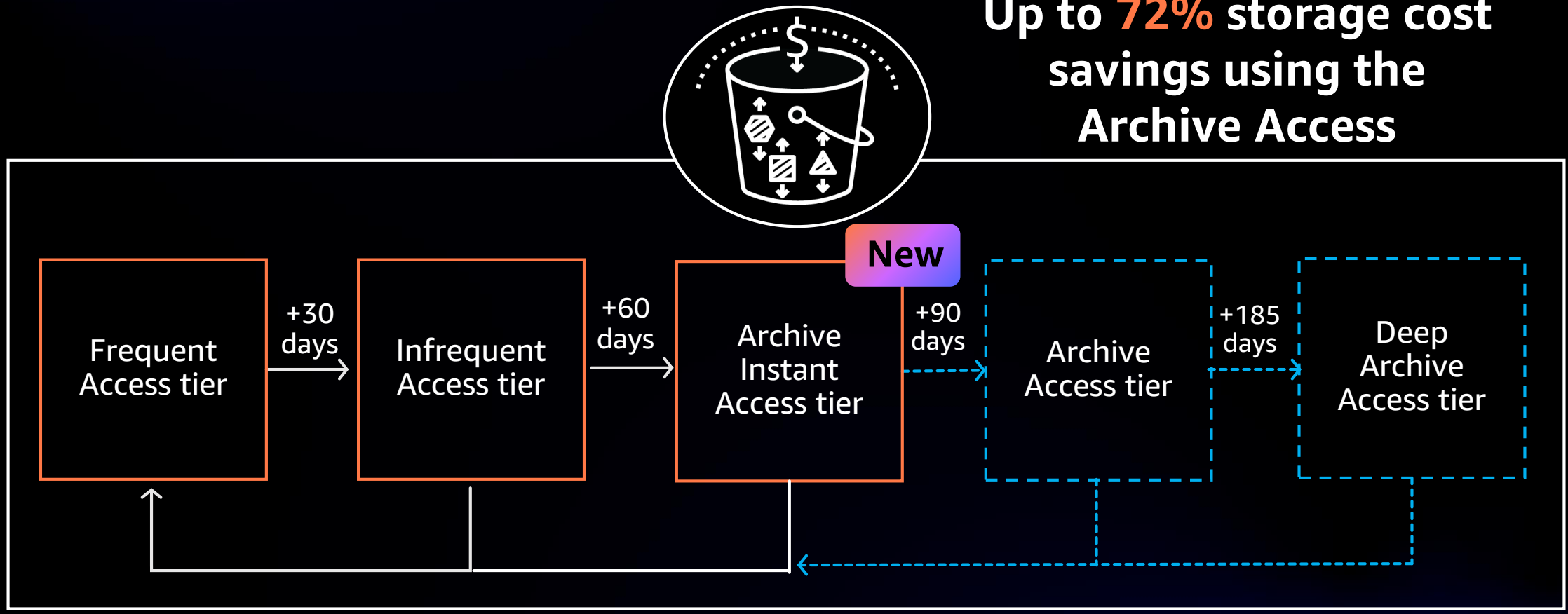
© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

S3 Intelligent-Tiering to automate data archiving

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Get the lowest storage cost for data that is not being accessed

Up to 72% storage cost savings using the Archive Access



Milliseconds access (automatic)

Minutes to hours (optional)

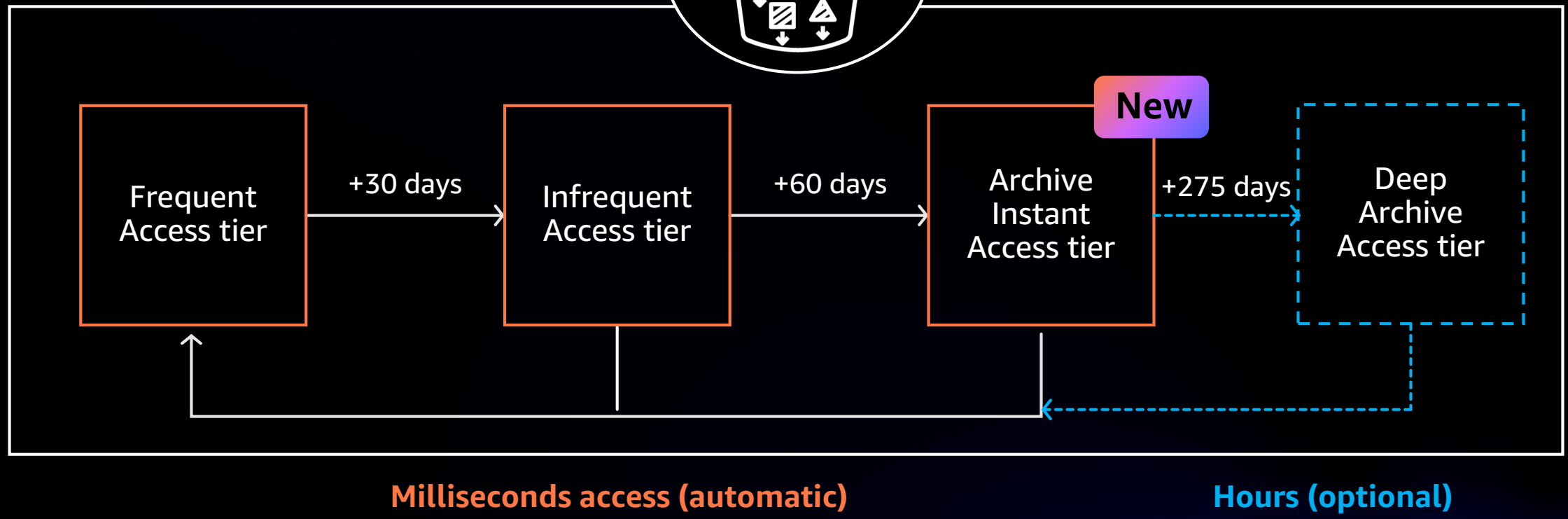
aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Get the lowest storage costs in the cloud automatically

Up to 95% storage cost savings using Deep Archive Access

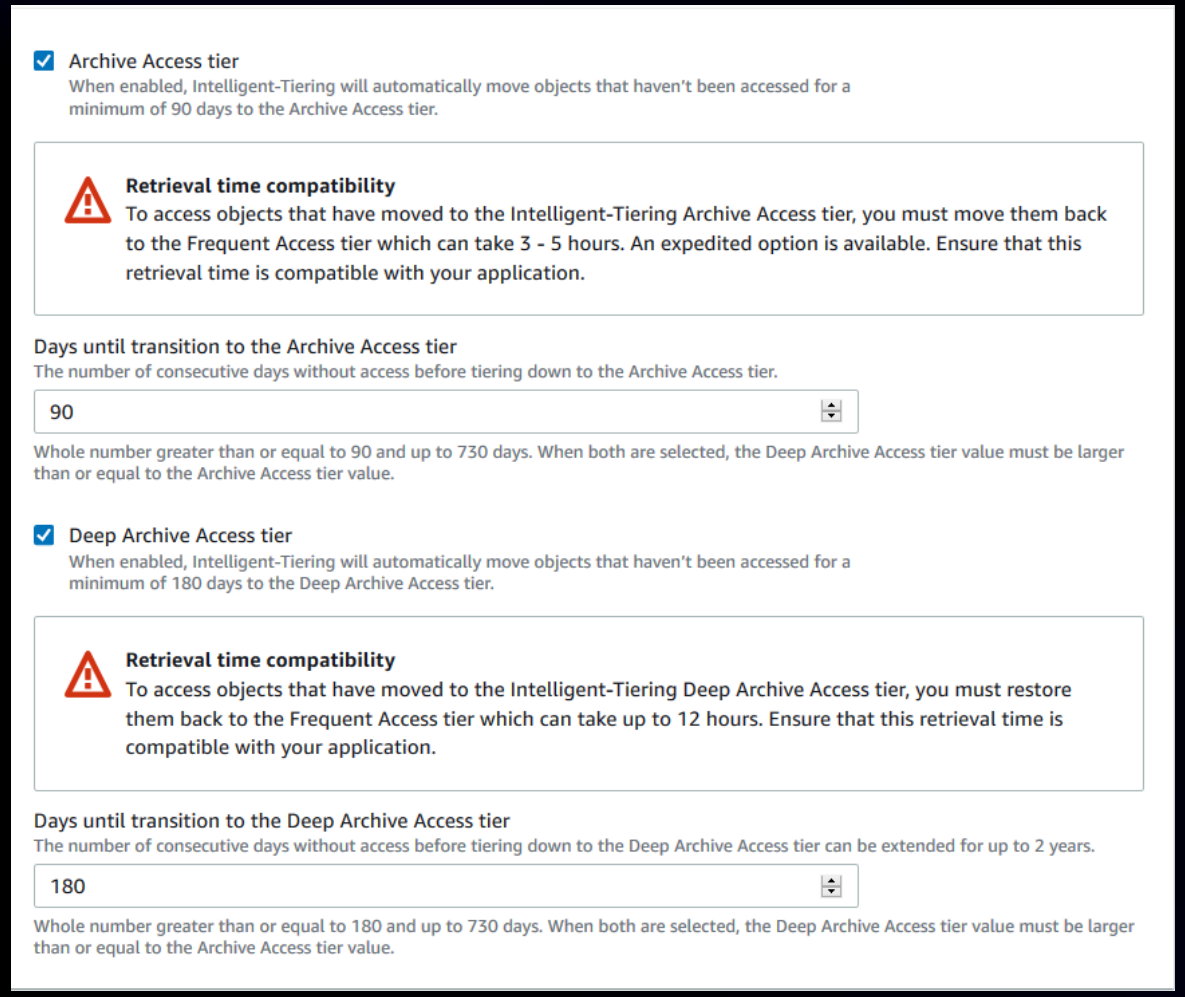




aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Archive data when you are ready



Configure the Archive Access or Deep Archive Access tiers, or both

90-day minimum days of consecutive no access to Archive Access tier

180-day minimum days of consecutive no access to Deep Archive Access tier

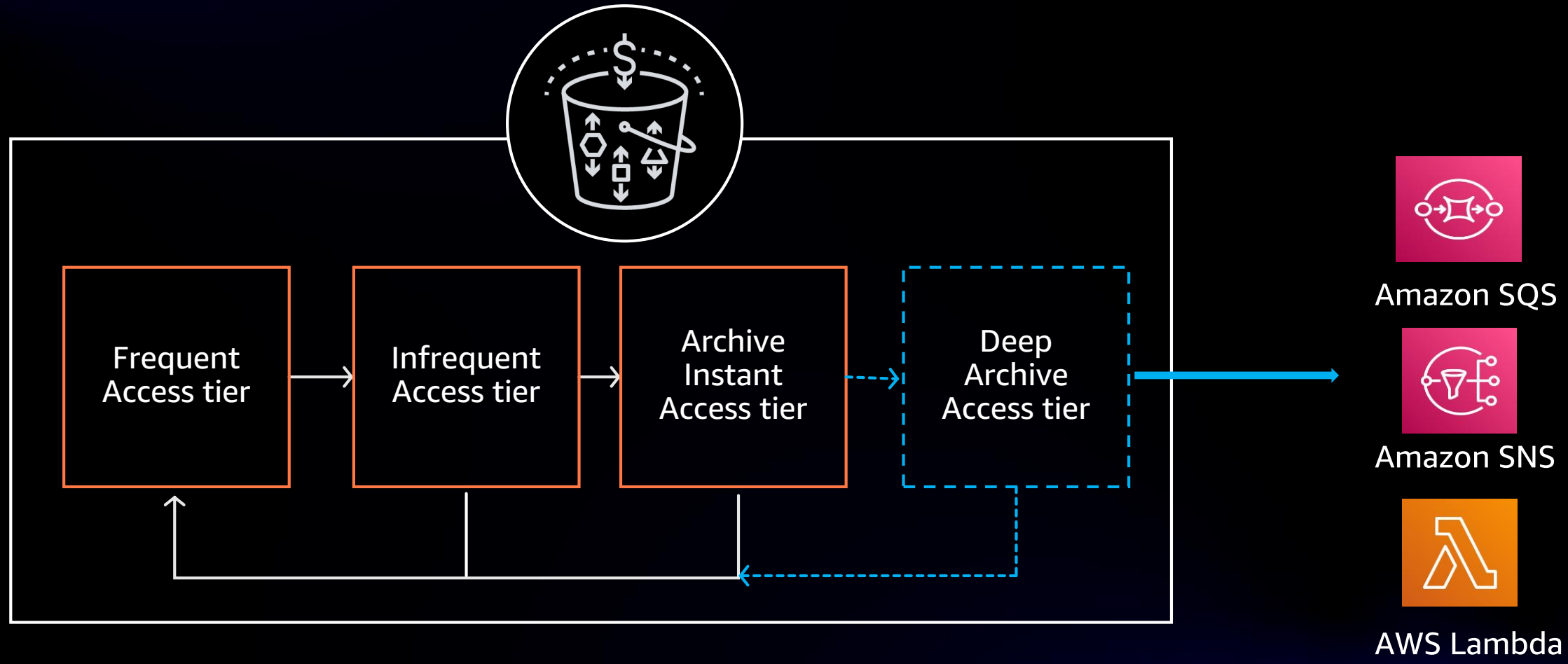
Extend the last access time for archiving up to two years

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon S3 Intelligent-Tiering

New

ENABLE EVENT NOTIFICATIONS TO TRACK OBJECTS MOVED TO ASYNC ARCHIVE



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

S3 Intelligent-Tiering continuous optimization

S3 Intelligent-Tiering (2018)

S3 Intelligent-Tiering - Access tiers in S3 Inventory (2019)

AWS Storage Gateway adds S3 Intelligent-Tiering (2019)

S3 Intelligent-Tiering Async Archive Access tiers (2020)

S3 Intelligent-Tiering Pricing optimizations (2021)

New

S3 Intelligent-Async Archive Events (2021)

New



S3 Intelligent-Tiering Archive Instant Access (2021)

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

"Since the launch of S3 Intelligent-Tiering in 2018, we've automatically saved ~30% per month on our storage costs without any impact on performance or need to analyze our data. With the new Archive Instant Access tier, we anticipate automatically realizing the benefit of archive storage pricing, while retaining the ability to access our data instantly when needed."

Kalyana Chadalavada

Head of Efficiency @ Stripe



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

"We use S3 Intelligent-Tiering because access patterns are often unpredictable. With the new Intelligent Tiering Archive Instant Access Tier, we'll further reduce cost for many Petabytes of Autonomous Vehicle data while keeping it immediately accessible for the use of our developers."

Yanor Barros
Director of Cloud Infrastructure Development @Mobileye



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

"We wanted to find a way to quickly optimize storage costs across the largest and fastest growing buckets across the enterprise. Because the storage usage patterns vary widely across our top buckets, there was no clear-cut rule we could safely apply without taking on some operational overhead. The S3 Intelligent-Tiering storage class delivered automatic storage savings based on the changing access patterns of our data without impact on performance. We look forward to S3 Intelligent-Tiering's new access tier which will allow us to realize even greater savings without additional effort."

Jerzy Grzywinski

Director of Software Engineering @ Capital One



aws

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Key takeaways

- Use S3 Intelligent-Tiering by default for datasets that have **unknow or changing access patterns**
- S3 Intelligent-Tiering automatically moves data to new Archive Instant Access tier, with **no impact on performance**
- Activate the asynchronous archive access tiers to automatically get the **lowest storage cost in the cloud**
- Use S3 Intelligent-Tiering archive events to **automatically update your application** when objects are not immediately accessible

© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Alex Crabbe
alcrabb@amazon.com

Jessie Felix
jessfeli@amazon.com



© 2021, Amazon Web Services, Inc. or its affiliates. All rights reserved.