



# Optimizing storage costs using Amazon S3

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

Andrew Kutsy (he/him)  
Product Manager, Amazon S3  
AWS

# Agenda

Patterns of cost optimization

---

Using Amazon S3 storage classes

---

Analyzing your storage using S3 Storage Lens

---

S3 Intelligent-Tiering for automatic storage cost savings





# 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

Digital record preservation

Media primary files

Home-recording video

**Mobile sync and storage**

Origin storage for CDN

Seismic and reservoir simulation data

Durable backups

Pharmaceutical study data

DNA sequences



## ML training data

## Media assets

Surveillance video/closed-circuit television

Financial transaction records

Website hosting

Log files

Meteorological and environmental research

## User-generated content

Oil and gas topography

Autonomous vehicle data



# Operating at scale in Amazon S3





# Redefining cost optimization with S3

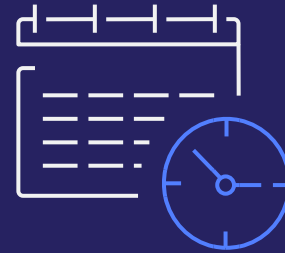
# The S3 Glacier Instant Retrieval storage class



\$0.004 per  
GB-month



Milliseconds  
access



Designed for  
99.9%  
availability



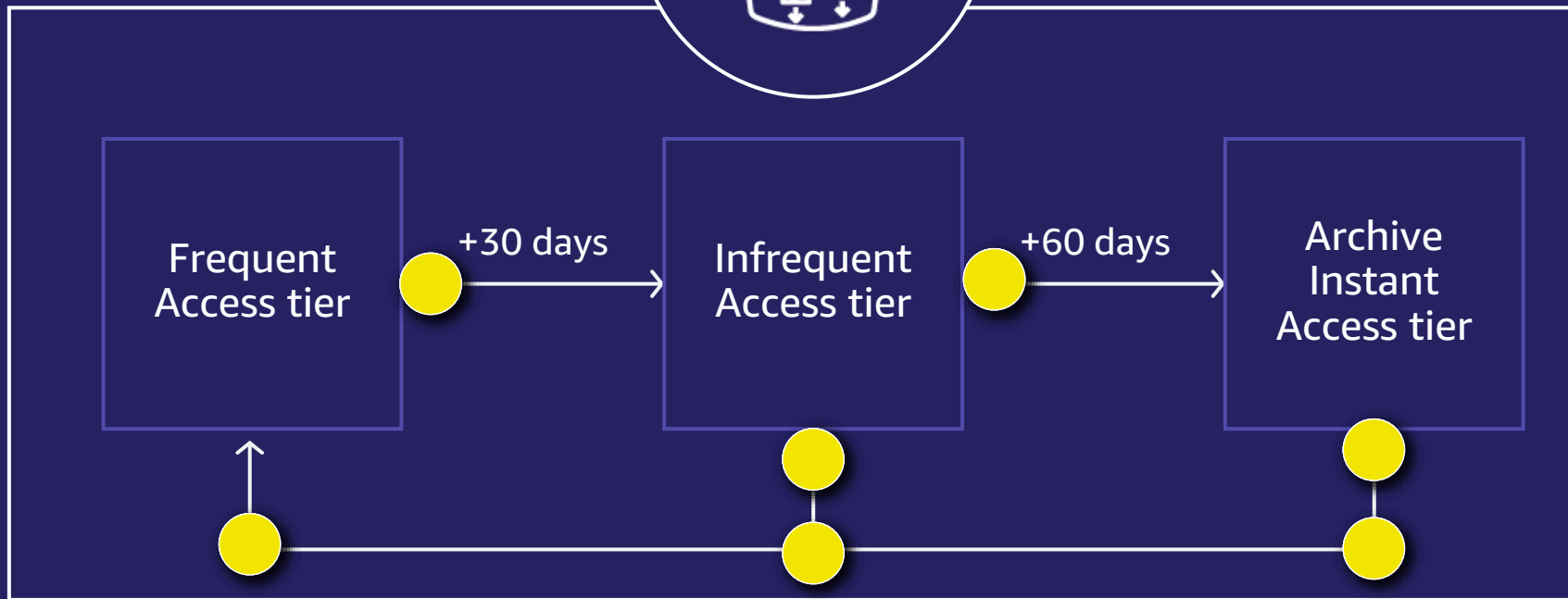
Designed for  
99.999999999%  
durability

**Lowest cost storage with milliseconds retrieval for rarely accessed data**

# Archive Instant Access tier in S3 Intelligent-Tiering



Automatically delivers up to **68%** in storage cost savings





# Patterns of cost optimization



1

Data with **known** or **predictable** access patterns

2

Data with **unknown** or **changing** access patterns

# Patterns of cost optimization



1

Data with **known** or **predictable** access patterns

2

Data with unknown or changing access patterns

# Patterns of cost optimization



1

Data with known or predictable access patterns

2

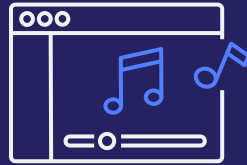
Data with **unknown** or **changing** access patterns

# Data with **predictable** access patterns



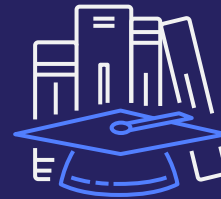
---

Medical records



---

Media streaming



---

Learning resources



---

Backups

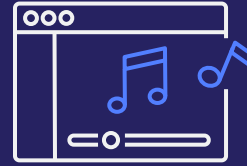
# Data with **unknown** access patterns



Data Lakes



Data Analytics



Content Sharing



Long Tail Data

ACROSS EVERY INDUSTRY

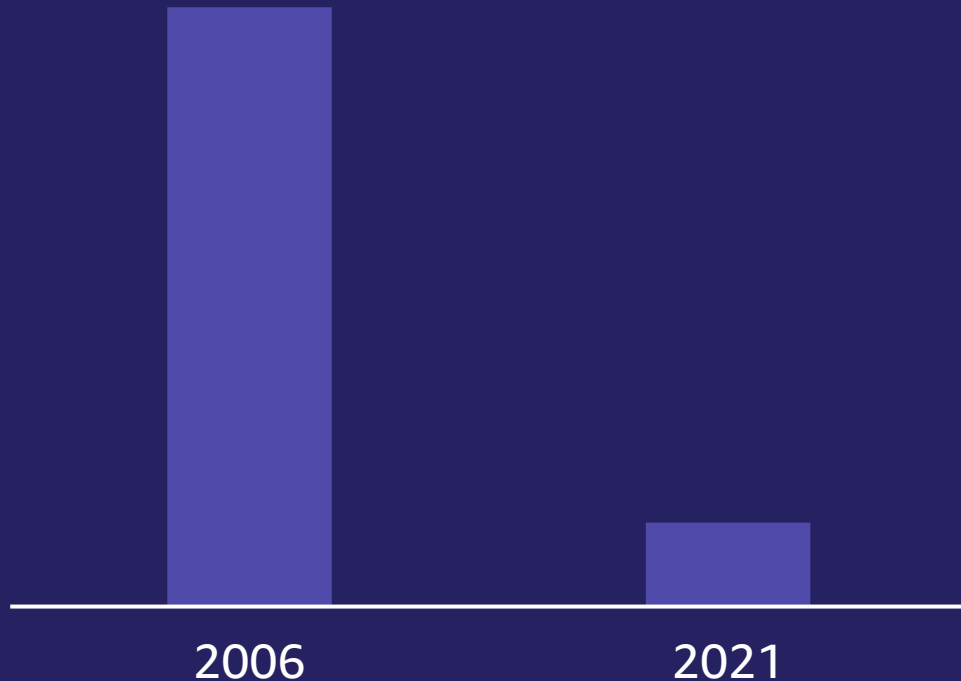
# Using Amazon S3 storage classes



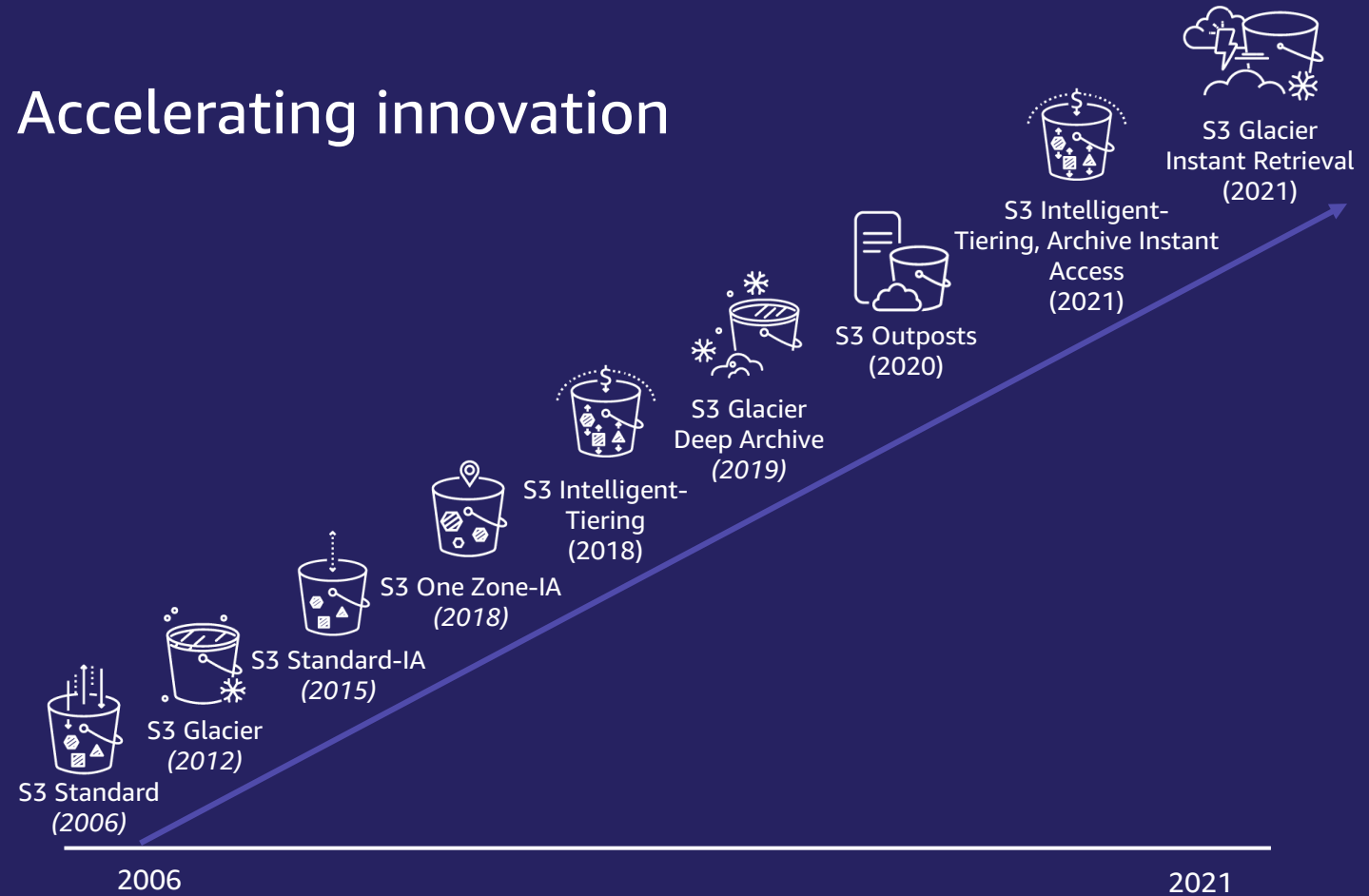
# Amazon S3 storage classes

OPTIMIZE YOUR STORAGE COST BY UTILIZING ALL AMAZON S3 STORAGE CLASSES

## Decreasing storage prices



## Accelerating innovation



# Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

## AWS Region $\geq$ 3 Availability Zones

### Data with changing access patterns

- Milliseconds access
- No retrieval fees

### Frequently accessed data

- Milliseconds access

### Infrequently accessed data

- Milliseconds access
- Minimum object size

### Rarely accessed data

- Milliseconds access
- Minimum object size

### Archive data

- Retrieval options from minutes to hours
- Minimum object size
- Free bulk retrievals

### Long term archive data

- Retrieval in hours
- Minimum object size

## AWS AZ

### Re-creatable, less accessed data

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

## AWS Outposts

### On-premises data

- Milliseconds access
- Encrypted with SSE-S3



# Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

## AWS Region $\geq$ 3 Availability Zones

| Frequently accessed data  | Infrequently accessed data   | Rarely accessed data   | Archive data   | Long term archive data  |
|---|--|--|--|---|
| <ul style="list-style-type: none"> <li>• Milliseconds access</li> </ul> | <ul style="list-style-type: none"> <li>• Milliseconds access</li> <li>• Minimum object size</li> </ul> | <ul style="list-style-type: none"> <li>• Milliseconds access</li> <li>• Minimum object size</li> </ul> | <ul style="list-style-type: none"> <li>• Retrieval options from minutes to hours</li> <li>• Minimum object size</li> <li>• Free bulk retrievals</li> </ul> | <ul style="list-style-type: none"> <li>• Retrieval in hours</li> <li>• Minimum object size</li> </ul> |

## AWS AZ

- Re-creatable, less accessed data
- Milliseconds access
  - Retrieval charge per GB
  - Minimum object size

## AWS Outposts

- On-premises data
- Milliseconds access
  - Encrypted with SSE-S3

### Data with changing access patterns

- Milliseconds access
- No retrieval fees



# Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

## AWS Region $\geq$ 3 Availability Zones

Data with changing access patterns

- Milliseconds access
- No retrieval fees

Frequently accessed data

- Milliseconds access

Infrequently accessed data

- Milliseconds access
- Minimum object size

Rarely accessed data

- Milliseconds access
- Minimum object size

Archive data

- Retrieval options from minutes to hours
- Minimum object size
- Free bulk retrievals

Long term archive data

- Retrieval in hours
- Minimum object size

## AWS AZ

Re-creatable, less accessed data

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

## AWS Outposts

On-premises data

- Milliseconds access
- Encrypted with SSE-S3

# Your choice of Amazon S3 storage classes



S3 Intelligent-Tiering



S3 Standard



S3 Standard-IA



S3 Glacier Instant Retrieval



S3 Glacier Flexible Retrieval (formerly S3 Glacier)



S3 Glacier Deep Archive



S3 One Zone-IA



S3 Outposts

## AWS Region $\geq$ 3 Availability Zones

### Data with changing access patterns

- Milliseconds access
- No retrieval fees

### Frequently accessed data

- Milliseconds access

### Infrequently accessed data

- Milliseconds access
- Minimum object size

### Rarely accessed data

- Milliseconds access
- Minimum object size

### Archive data

- Retrieval options from minutes to hours
- Minimum object size
- Free bulk retrievals

### Long term archive data

- Retrieval in hours
- Minimum object size

## AWS AZ

### Re-creatable, less accessed data

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

## AWS Outposts

### On-premises data

- Milliseconds access
- Encrypted with SSE-S3

# Economics and primary factors to consider



Frequency of access



Duration of storage

# Upload data directly to an S3 storage class

- Begin saving at ingest by choosing the right S3 storage class
- Specify **GLACIER\_IR** in the PUT API request header

```
PUT /my-image.jpg HTTP/1.1
Host: myBucket.s3.<Region>.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: GLACIER_IR
```

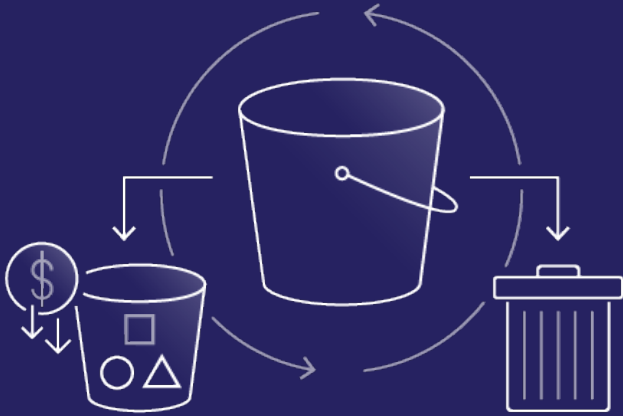
# Lifecycle data with predictable access patterns



# Relationship between object size and total request costs



# S3 Lifecycle filters and actions



- 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



# Fine tune your Lifecycle policies

Lifecycle rules take action based on object age:

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



S3 Standard



S3 Glacier  
Instant Retrieval

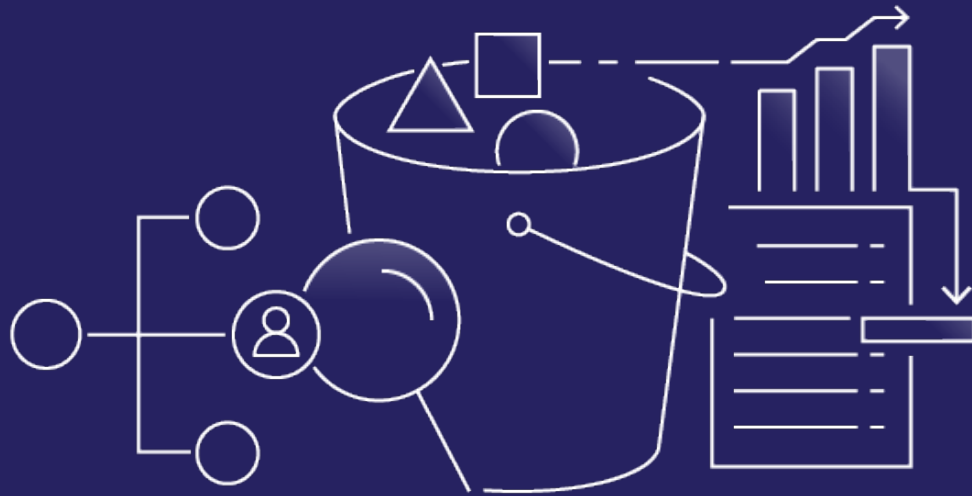


S3 Glacier Deep  
Archive



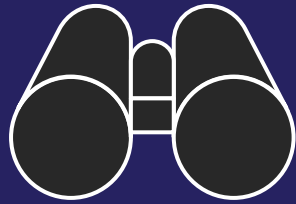
# Choosing the right storage class with S3 Storage Lens

# Amazon S3 Storage Lens overview

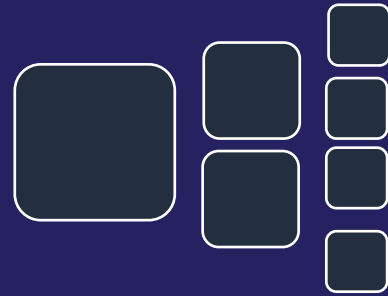


- Interactive dashboard experience in the S3 console – free to all customers
- Organization-wide visibility
- Drill-down by Region, storage class, bucket, and prefix
- Granular usage & activity metrics
- Call outs for cost efficiency & data protection best practices
- Publish metrics to CloudWatch

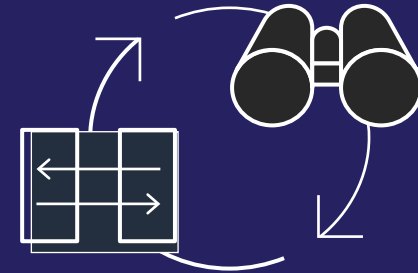
# Why we built S3 Storage Lens



Organization wide  
visibility into usage and  
activity metrics



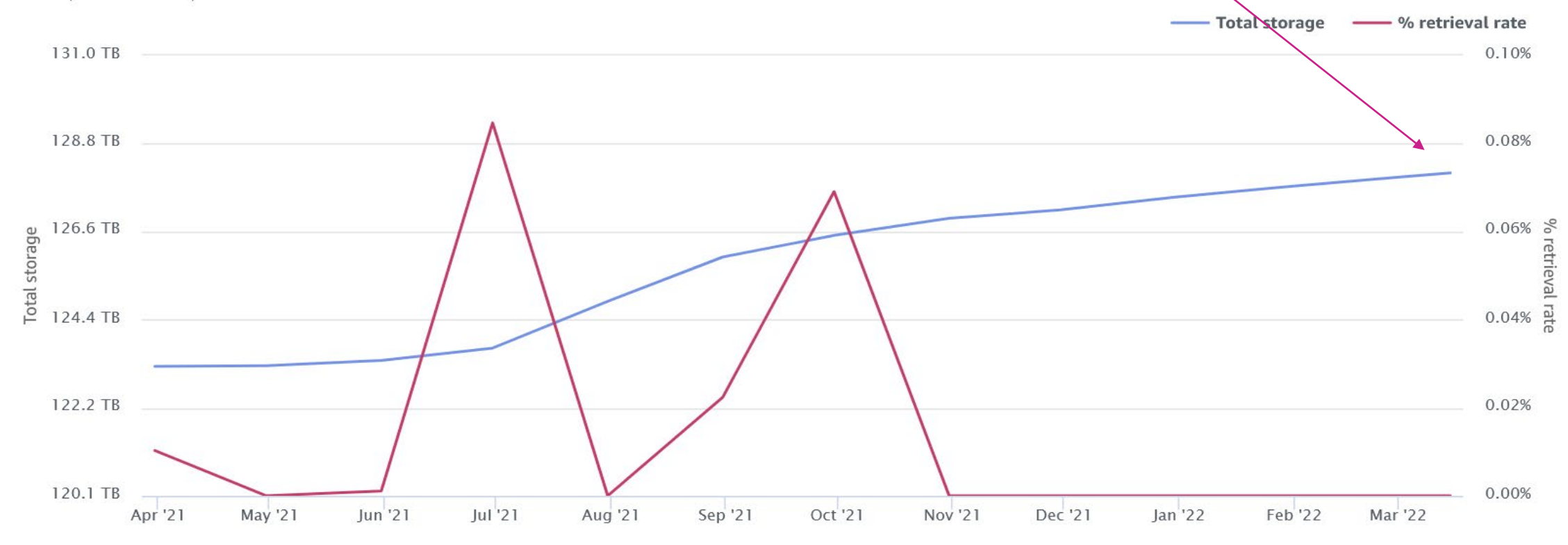
Ability to drill down to  
gain detail understanding  
of usage by buckets and  
prefixes



Apply insights to  
optimize storage cost

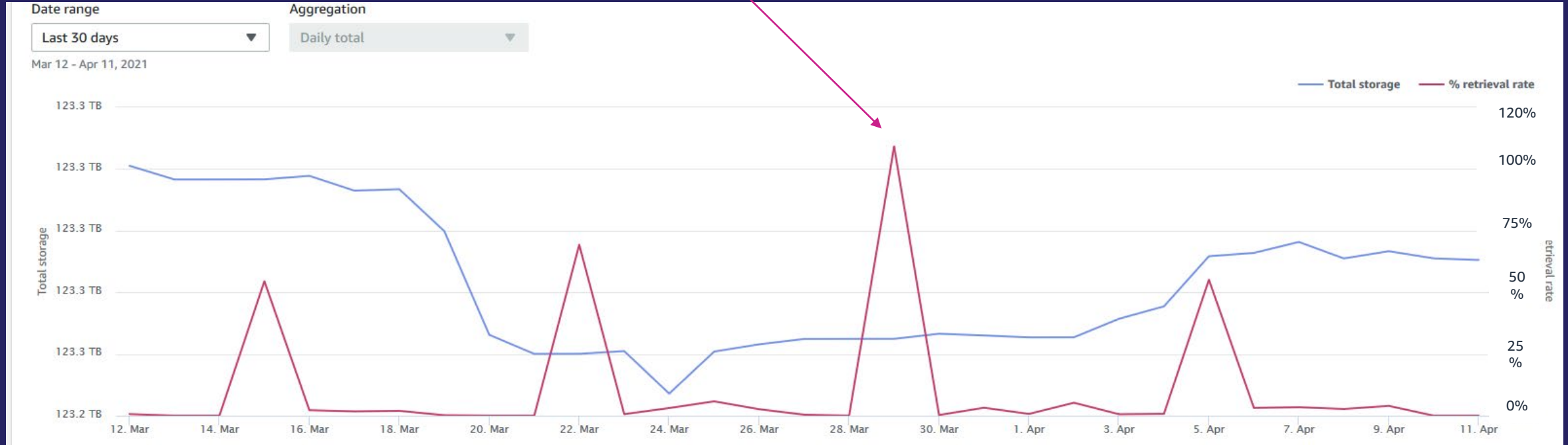
# Data with **predictable** access patterns usage characteristics

Workloads with predictable patterns often have **low retrieval ratios** for long periods of time after becoming rarely accessed

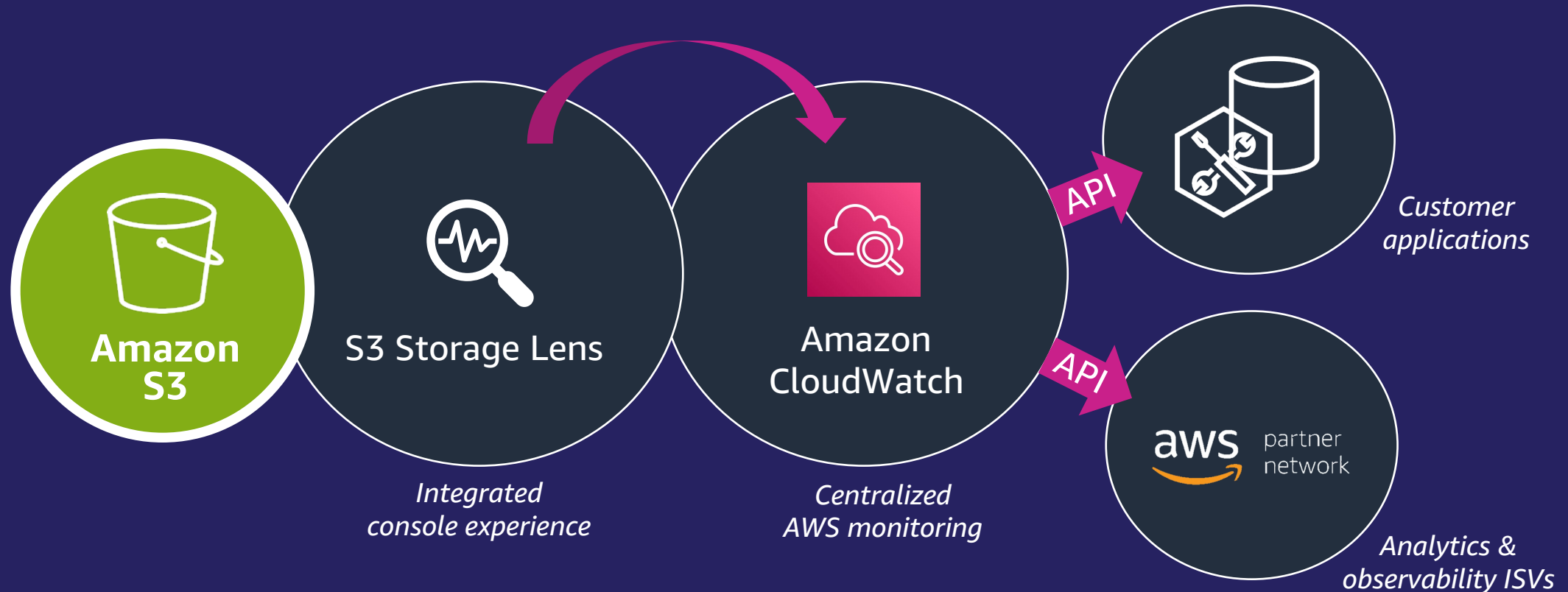


# Data with **unknown or changing** access patterns usage characteristics

Workloads with changing access patterns can look cold for a time but **retrievals unexpectedly spike** above 100% retrieval ratios



# Amazon CloudWatch publishing for S3 Storage Lens



# Why we built Amazon CloudWatch publishing for S3 Storage Lens



Customize dashboard expressions to get new metrics and chart styles



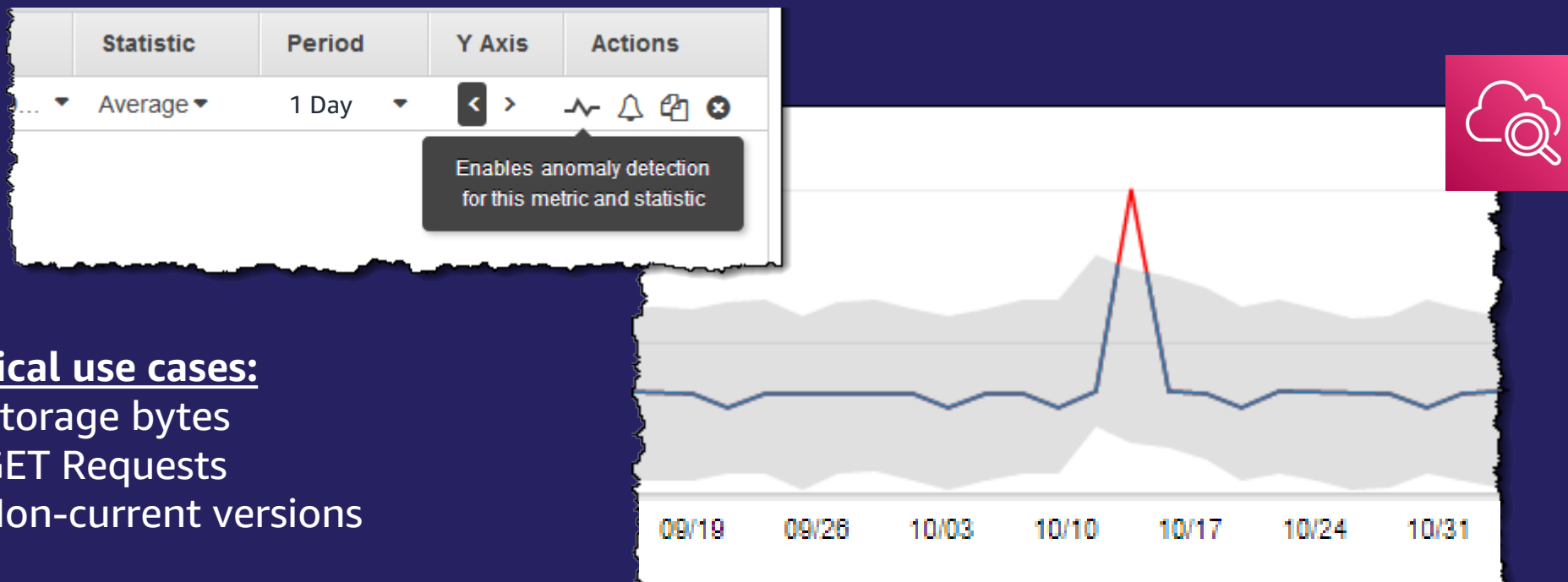
Configure alarms on S3 Storage Lens metrics to trigger SNS notifications



Build custom monitoring applications using S3 Storage Lens APIs to identify anomalies



# CloudWatch Anomaly detection using Storage Lens metrics



## Typical use cases:

- Storage bytes
- GET Requests
- Non-current versions

# S3 Intelligent-Tiering



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

**\$250,000,000**

---

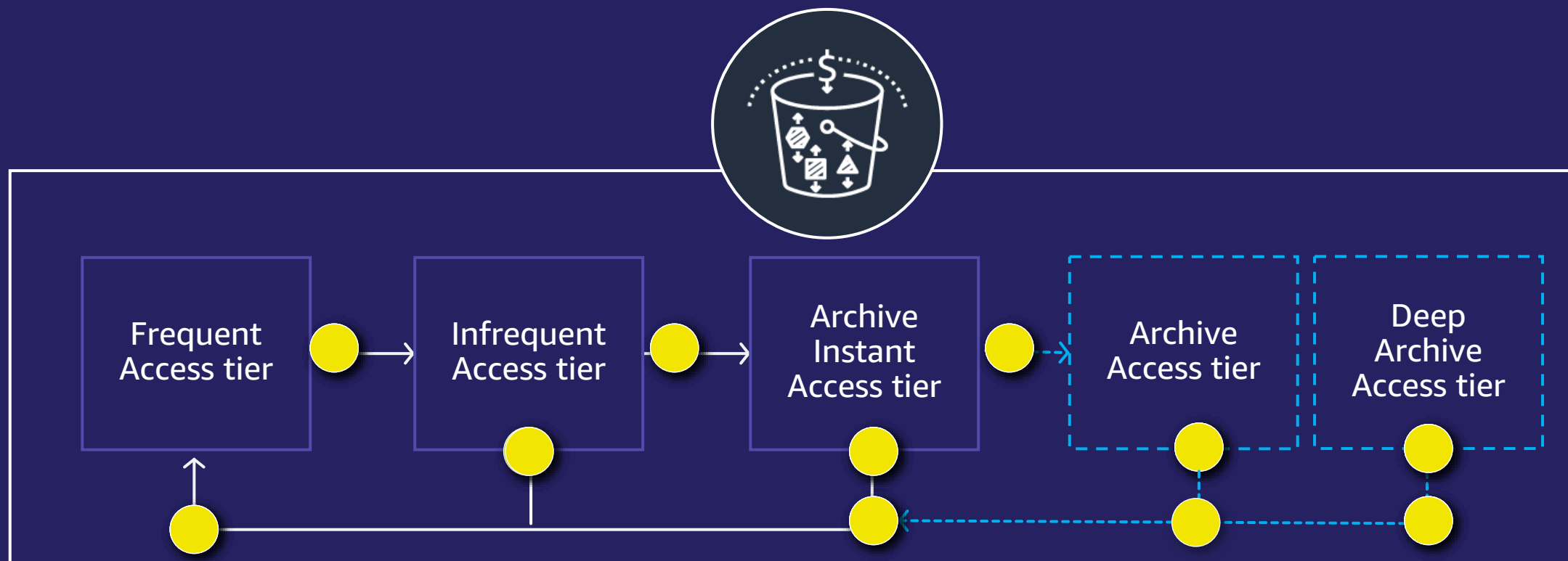


# What is S3 Intelligent-Tiering?



- 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 99.9999999999% (11 9s) of durability

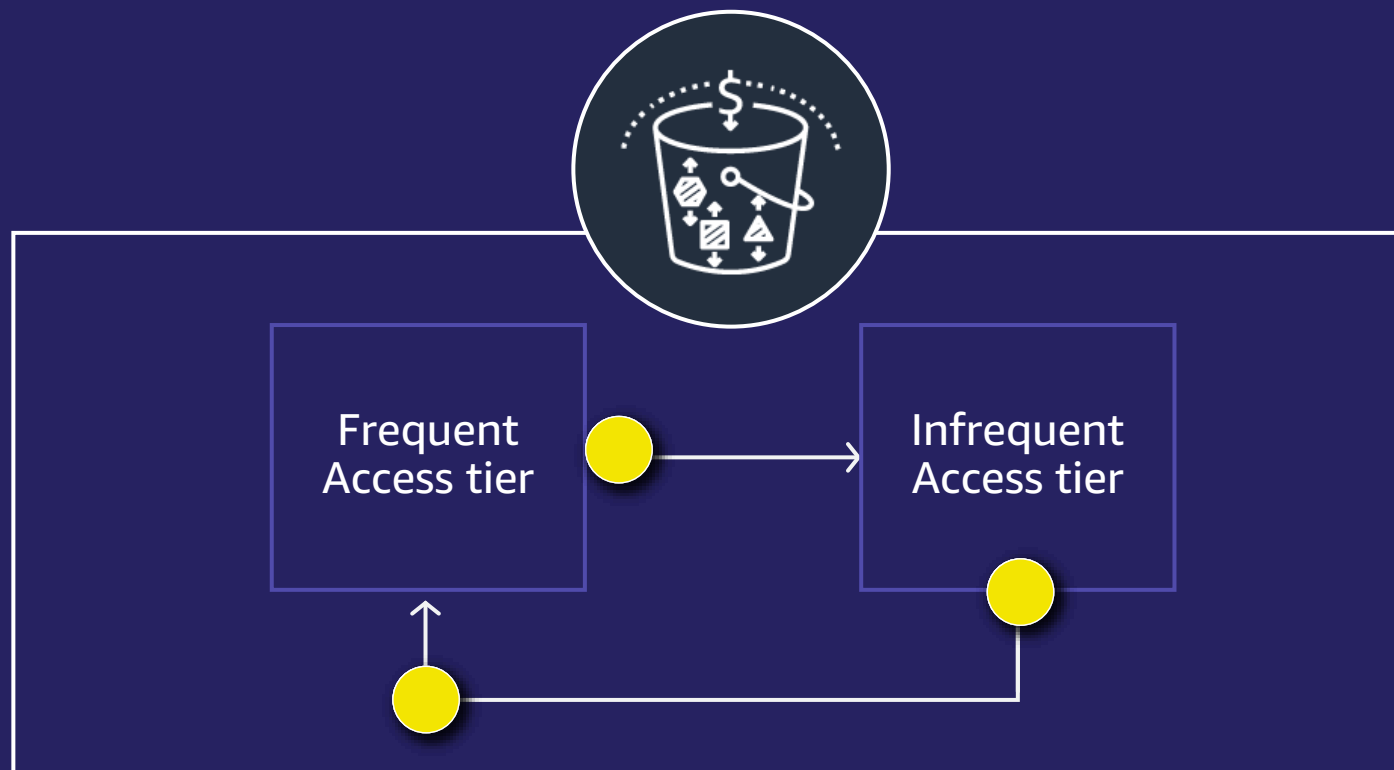
# Use S3 Intelligent-Tiering for data with **unknown or changing access patterns**



Milliseconds access (automatic)

Minutes to hours (optional)

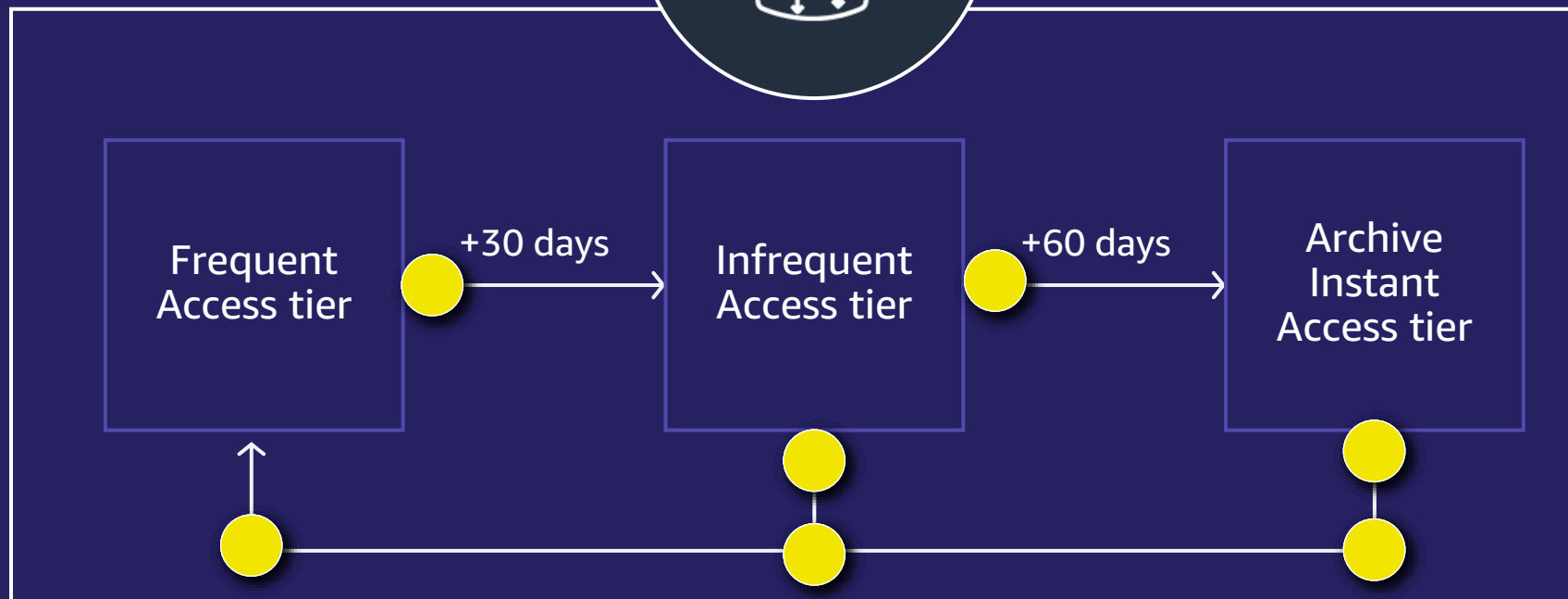
# Before re:Invent 2021: S3 Intelligent-Tiering automatically optimized cost in two access tiers



# Since re:Invent 2021: S3 Intelligent-Tiering automatically optimizes cost in three access tiers



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



# 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



# 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



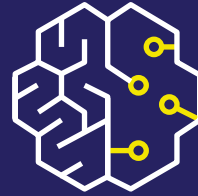
Designed for 99.999999999% durability

# S3 Intelligent-Tiering pricing enhancements



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

# 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
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
```

# Define your archive strategy



# Content is growing rapidly, your storage cost don't have to



Media



User data



Files



Connected devices

# S3 Glacier is the **best place** to archive your data



S3 Glacier Instant Retrieval

Fastest access to archive storage



S3 Glacier Flexible Retrieval

Flexible retrieval options



S3 Glacier Deep Archive

Lowest storage cost in the cloud



S3 Intelligent-Tiering

Automatic archiving based on the last access of your objects



# S3 Glacier is the best place to archive



**S3 Glacier  
Instant Retrieval**



**S3 Glacier  
Flexible Retrieval**



**S3 Glacier  
Deep Archive**

|                                | <b>S3 Glacier<br/>Instant Retrieval</b> | <b>S3 Glacier<br/>Flexible Retrieval</b>                               | <b>S3 Glacier<br/>Deep Archive</b>                 |
|--------------------------------|---|--|--|
| <b>Storage cost</b>            | \$0.004 per GB-month                    | \$0.0036 per GB-month  | \$0.00099 per GB-month                             |
| <b>Data retrieval</b>          | Milliseconds retrieval                  | Expedited: 1-5 minutes<br>Standard: 3-5 hours<br>FREE Bulk: 5-12 hours | Standard: Within 12 hours<br>Bulk: Within 48 hours |
| <b>Minimum object duration</b> | 90 days                                 | 90 days  | 180 days   |



What if you have **unknown**  
or **changing** access patterns?

---

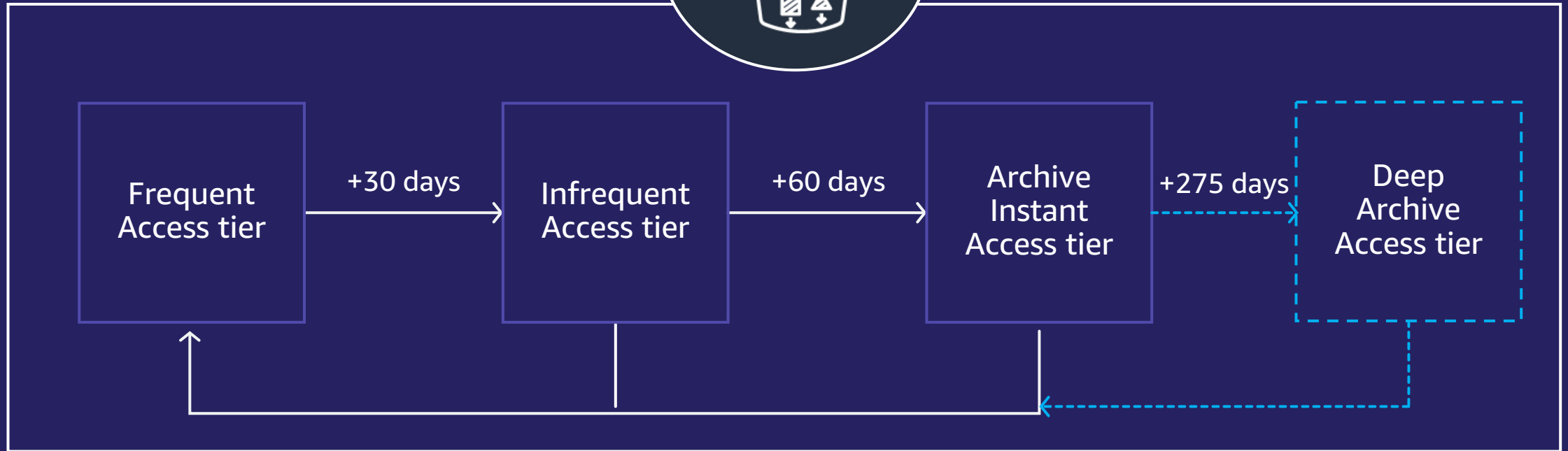




# Get the **lowest storage cost in the cloud automatically**



Up to **95%** storage cost savings using  
**Deep Archive Access**




**Milliseconds access (automatic)**

**Hours (optional)**



# Archive data when you are ready

**Archive Access tier**  
When enabled, Intelligent-Tiering will automatically move objects that haven't been accessed for a minimum of 90 days to the Archive Access tier.


 **Retrieval time compatibility**  
To access objects that have moved to the Intelligent-Tiering Archive Access tier, you must move them back to the Frequent Access tier which can take 3 - 5 hours. An expedited option is available. Ensure that this retrieval time is compatible with your application.

**Days until transition to the Archive Access tier**  
The number of consecutive days without access before tiering down to the Archive Access tier.

90

Whole number greater than or equal to 90 and up to 730 days. When both are selected, the Deep Archive Access tier value must be larger than or equal to the Archive Access tier value.

**Deep Archive Access tier**  
When enabled, Intelligent-Tiering will automatically move objects that haven't been accessed for a minimum of 180 days to the Deep Archive Access tier.

 **Retrieval time compatibility**  
To access objects that have moved to the Intelligent-Tiering Deep Archive Access tier, you must restore them back to the Frequent Access tier which can take up to 12 hours. Ensure that this retrieval time is compatible with your application.

**Days until transition to the Deep Archive Access tier**  
The number of consecutive days without access before tiering down to the Deep Archive Access tier can be extended for up to 2 years.

180

Whole number greater than or equal to 180 and up to 730 days. When both are selected, the Deep Archive Access tier value must be larger than or equal to the Archive Access tier value.

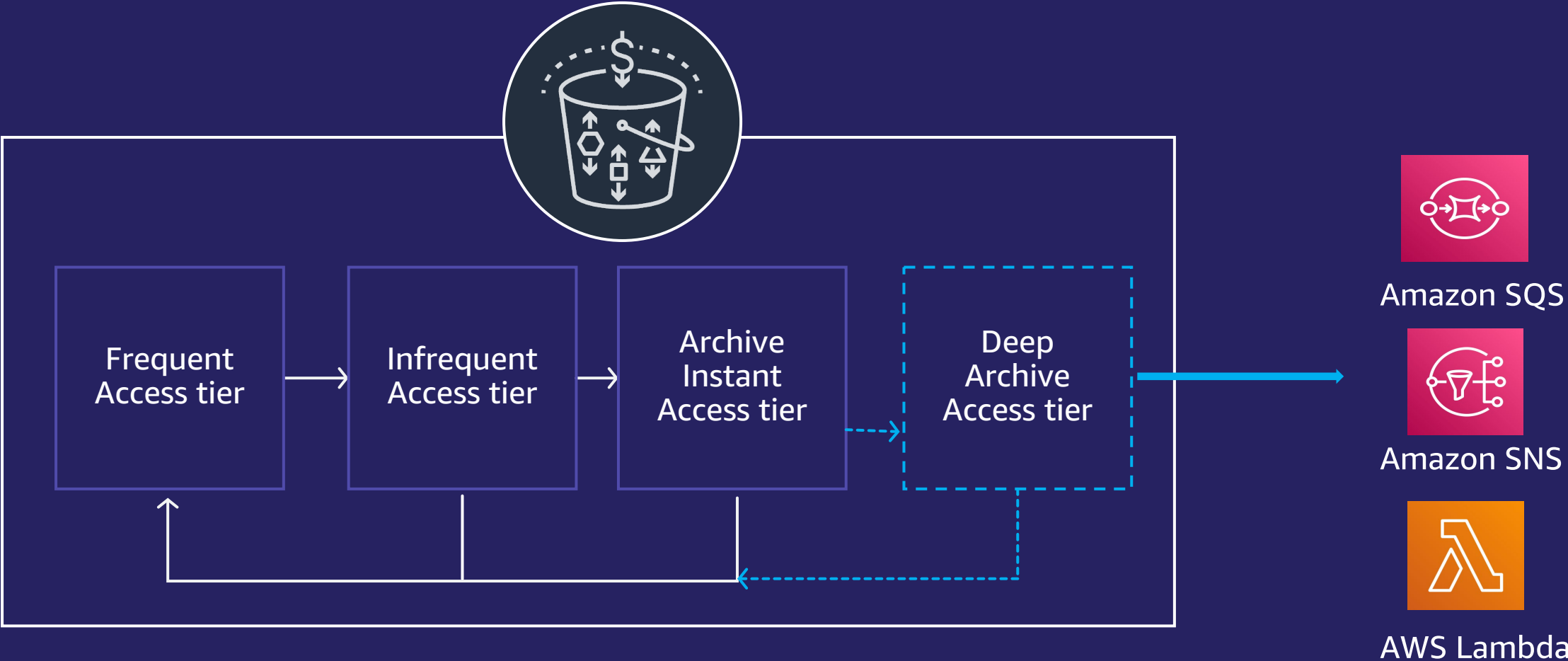
Configure the Archive Access or Deep Archive Access tiers, or both

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

# Enable event notifications to track objects moved to async archive



# Key takeaways

- 1 Use S3 Storage Lens to analyze your storage
- 2 Use our building blocks like S3 Lifecycle for data with **known** or **predictable** access patterns
- 3 Use S3 Intelligent-Tiering by default for data with **unknown** or **changing** access patterns
- 4 Put in a place an archiving storage strategy using Amazon S3 Glacier storage classes



**Thank you!**