



Evolving Cloud Storage Landscape

Rob Czarnecki
Amazon S3 Glacier

Agenda

AWS storage

S3 storage classes

S3 archive storage

AWS Partner Network integrations

Economics

Questions

AWS Storage

Adoption patterns

BBC

STEMCELL
TECHNOLOGIES

LIONSGATE



loanlogics

Celgene

Kellogg's

FINRA

photobox

Sysco

aws storage

Broad portfolio of storage services

Block storage



General purpose SSD
Provisioned IOPS SSD
Throughput-optimized HDD
Cold HDD
Elastic volumes

Backup



AWS Backup

Data transfer



AWS DataSync
AWS Transfer for SFTP

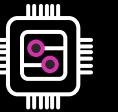


File storage



Amazon EFS Standard
Amazon EFS Infrequent Access (Amazon EFS IA)

Object storage



Amazon S3 Standard
Amazon S3 Standard-IA
Amazon S3 One Zone-IA
Amazon S3 Glacier
Amazon S3 Intelligent-Tiering
Amazon S3 Glacier Deep Archive

Diversity of use cases in AWS Storage

Customers choose the adoption pattern that matters for their business



Backup and restore

- Non-disruptive
- Easy place to start
- Integrated with all major vendors



Archive and compliance

- Media workflows
- Tape replacement
- Public Sector, Financial Services, Healthcare/ Life Sciences



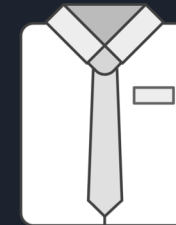
Home directories

- Simple to move
- Not sensitive to latency
- Significant cost savings



Data Lakes

- Variety of analytics tools
- Built for streaming data
- Data visualization



Critical Applications

- Integrated with major vendors
- Fully managed infrastructure
- Lift-and-shift migrations



Hybrid Storage

- Take advantage of the economies of cloud storage, while applications running on-premises
- Storage Gateway for hybrid cloud workloads

AWS pricing principles



No upfront
investment



Pay-as-you-go
approach



Pay less by using
more



Pay less as AWS
grows

S3 storage classes

Compliance records

Geospatial or lunar imagery

Internet of Things (IoT) sensor data

Medical imagery and records

Analytics

Media master files

Data lakes

Customer call-center records

Digital record preservation

Mobile sync and storage

Home-recording video

Origin storage for CDN

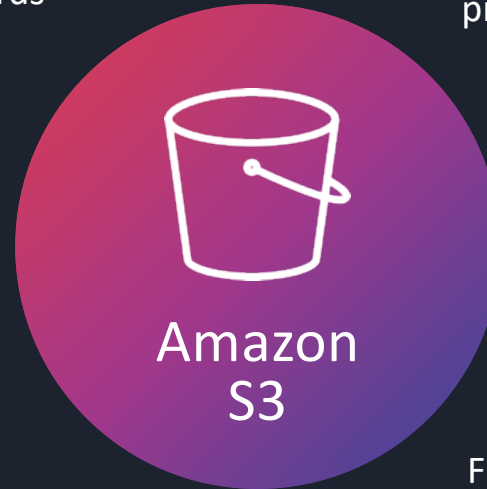
Seismic and reservoir simulation data

Durable backups

Pharmaceutical study data

DNA sequences

Surveillance video/closed-circuit television



Amazon S3

ML training data

Financial transaction records

Website hosting

Media assets

Meteorological and environmental research

User-generated content

Oil and gas topography

Autonomous vehicle data

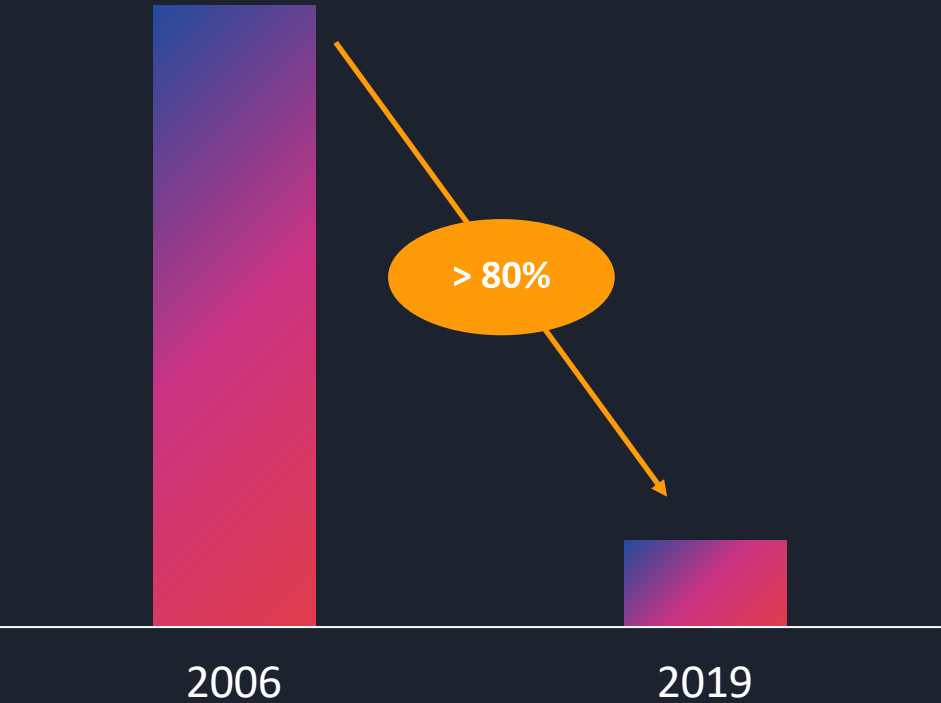
Log files

Mapping data

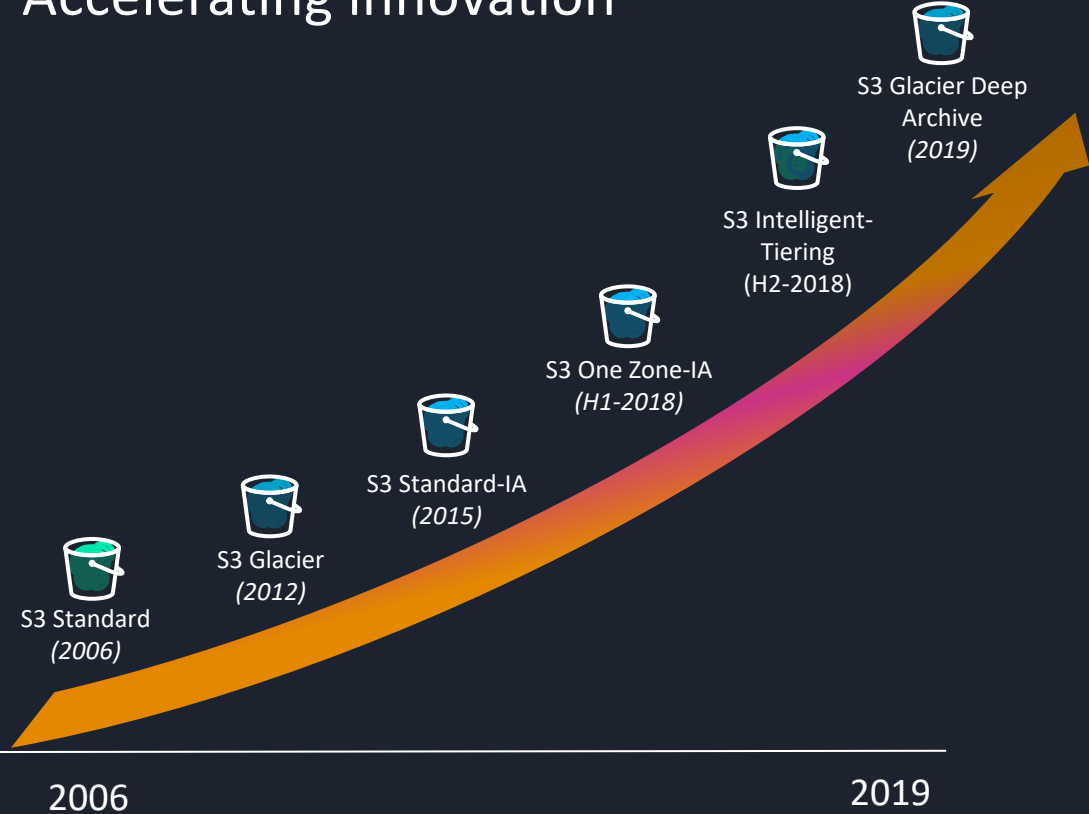


Decreasing prices and more storage options

Decreasing storage prices



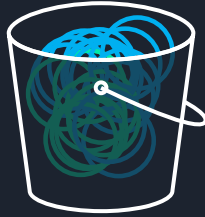
Accelerating innovation



Your choice of Amazon S3 storage classes



S3 Standard



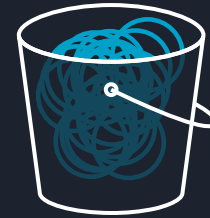
S3 Intelligent-Tiering



S3 Standard-IA



S3 One Zone-IA



S3 Glacier



S3 Glacier Deep Archive



Frequent

Access frequency

Infrequent

- Active, frequently accessed data
- Milliseconds access
- ≥ 3 AZ
- \$0.0210/GB

- Data with changing access patterns
- Milliseconds access
- ≥ 3 AZ
- \$0.0210 to \$0.0125/GB
- Monitoring fee per Obj.
- Min storage duration

- Infrequently accessed data
- Milliseconds access
- ≥ 3 AZ
- \$0.0125/GB
- Retrieval fee per GB
- Min storage duration
- Min object size

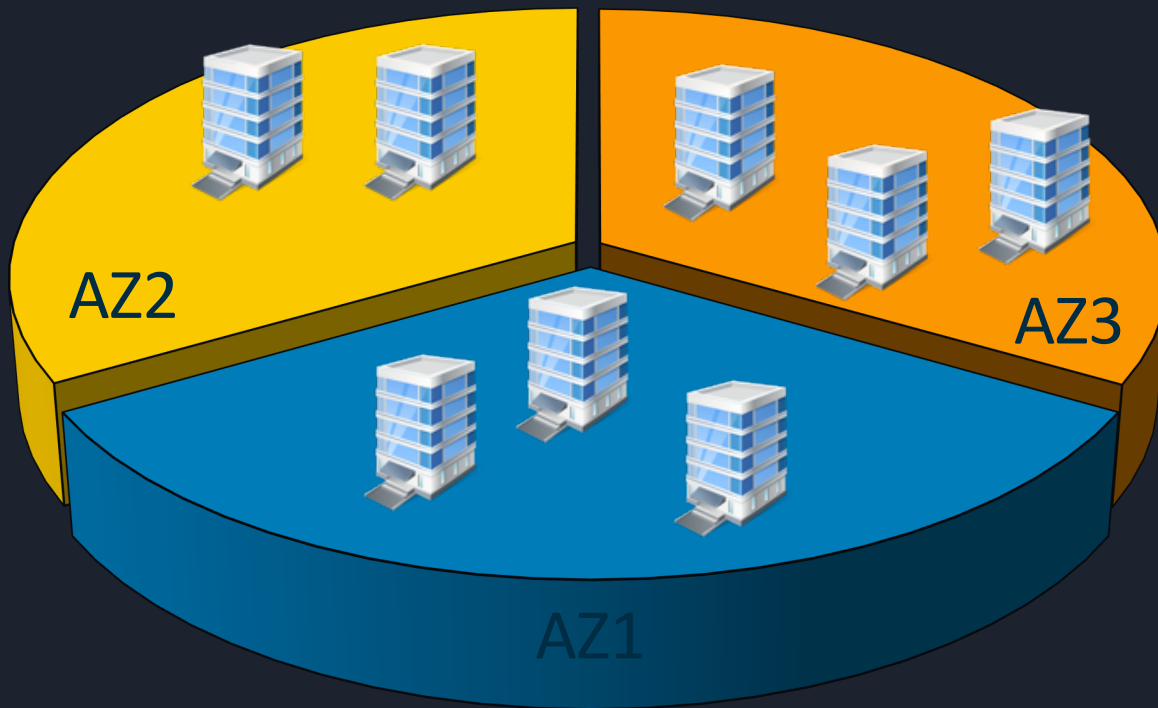
- Re-creatable, less accessed data
- Milliseconds access
- 1 AZ
- \$0.0100/GB
- Retrieval fee per GB
- Min storage duration
- Min object size

- Archive data
- Select minutes or hours
- ≥ 3 AZ
- \$0.0040/GB
- Retrieval fee per GB
- Min storage duration
- Min object size

- Long term Archive data
- Select hours
- ≥ 3 AZ
- \$0.00099/GB
- Retrieval fee per GB
- Min storage duration
- Min object size

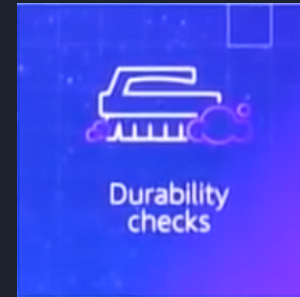
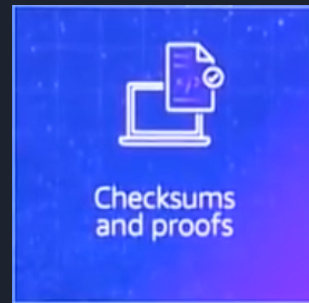
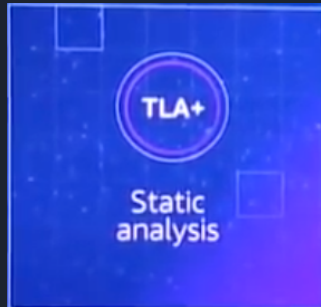
S3 achieves 99.999999999% durability via geographic protection for a single 'copy'

A Single AWS "Region"



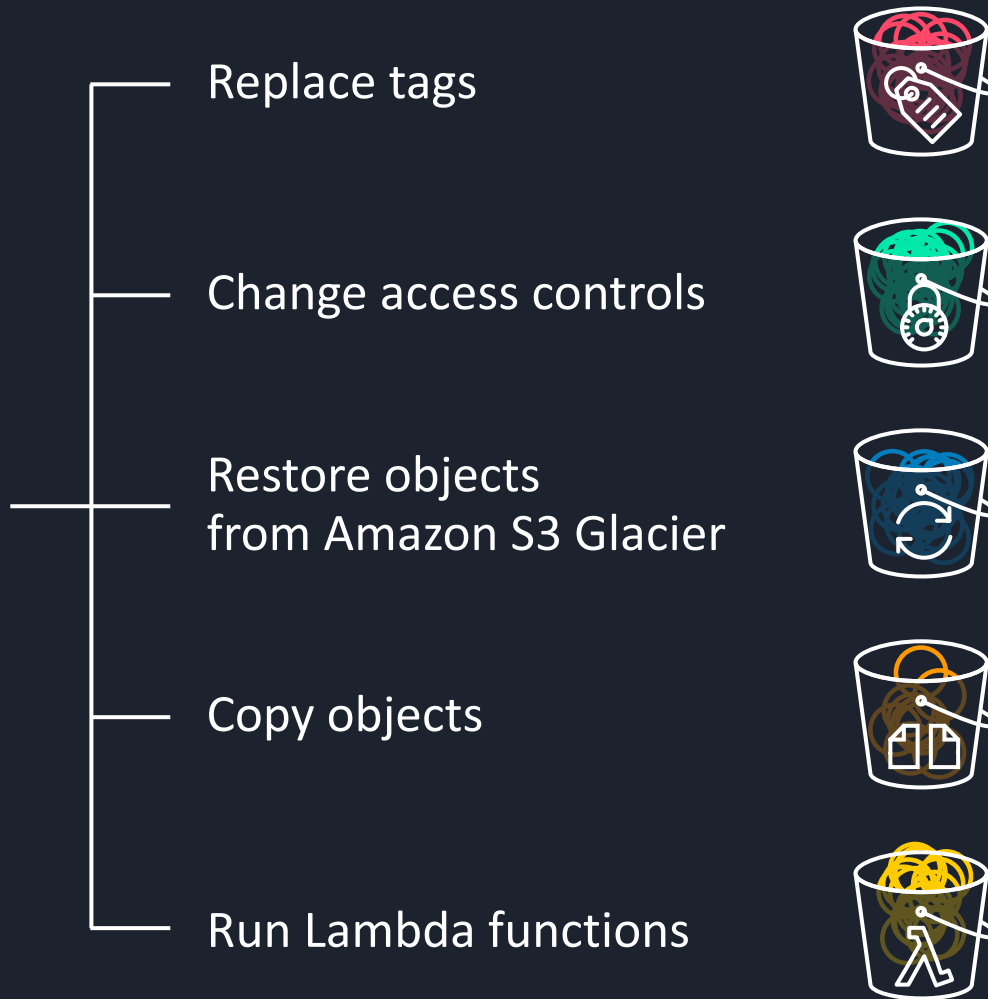
- ▀ Availability Zones physically separated, independent power, network, flood-plane
- ▀ Objects stored across AZ's
- ▀ Automated validation in all storage classes with automated self-healing

Culture of Durability



Amazon S3 Batch Operations

Take large-scale actions on Amazon S3 objects



Manage billions of objects at a time

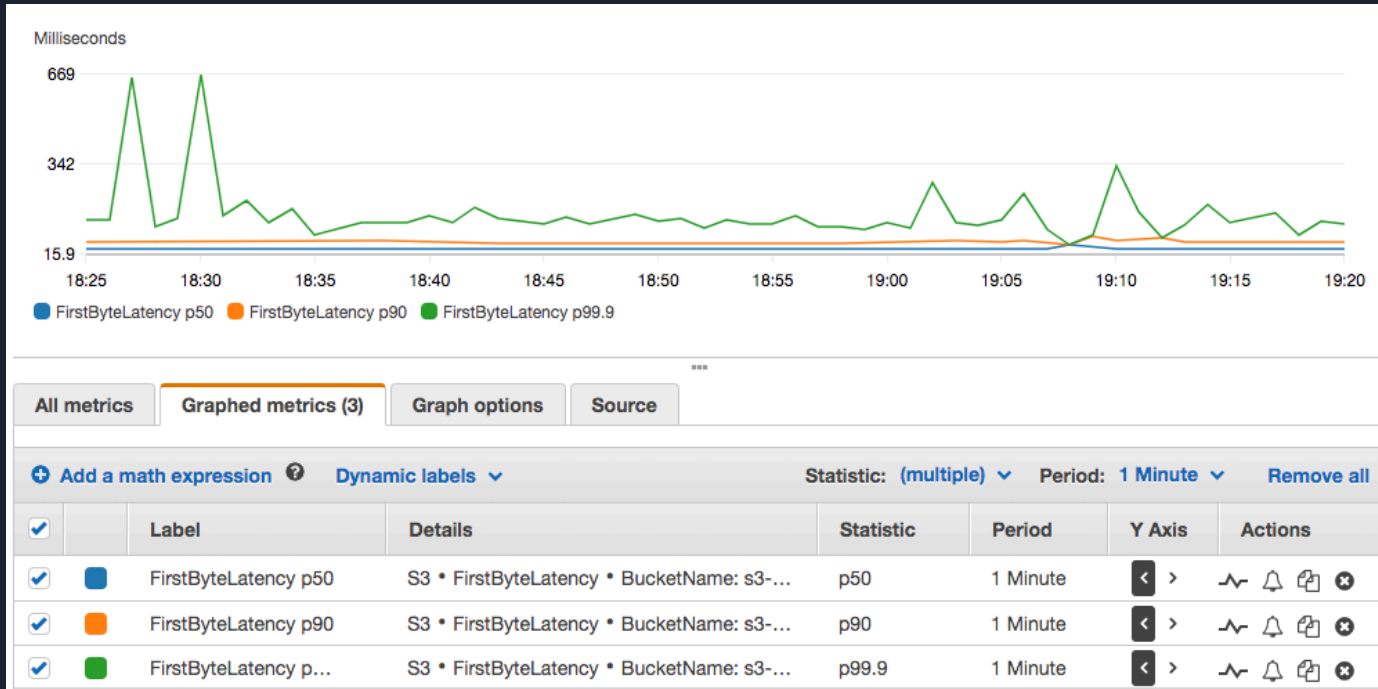
S3 Batch Operations allows you to replace tags, change ACLs, restore and objects, and run Lambda functions

Manages retries, tracks progress, sends completion notifications, generates reports, and delivers events

New!

Amazon S3 CloudWatch Percentiles Metrics

S3 request metrics on any percentile (e.g. p90, p99, p99.9, p100)



Use percentiles to understand the distribution of S3 request metrics

Visualize and alarm on any percentile to identify outliers or unusual application behavior

Avoid false alarms and save costs spent in monitoring and tracking requests

S3 archive storage

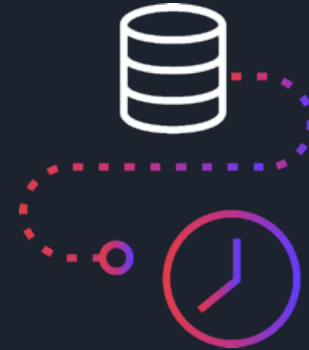
S3 Glacier and S3 Glacier Deep Archive?



No infrastructure
to manage



Designed for
99.9999999999%
durability



Recover data in
minutes to hours



Priced from
\$0.00099 per GB-
month

Archival storage through Amazon S3 API

Write to Amazon S3 Glacier directly via the Amazon S3 API

Direct PUT, Replicate or Lifecycle via
Amazon S3 API



Amazon S3 Glacier

Amazon S3 Glacier Deep
Archive

Features

- Lifecycle management
- Direct PUT to Amazon S3 Glacier
- Amazon S3 Object Lock (WORM storage)
- Amazon S3 Glacier Restore Speed Upgrade
- Amazon S3 Glacier Restore Notifications
- CRR/SRR direct to Amazon S3 Glacier

Multiple data transfer services



AWS Snowmobile



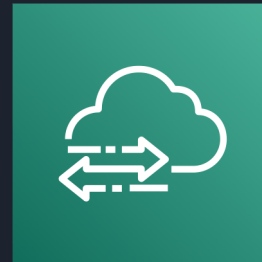
AWS Snowball



AWS Direct Connect



AWS Transfer for SFTP



AWS DataSync

Accessing data in S3 Glacier Deep Archive

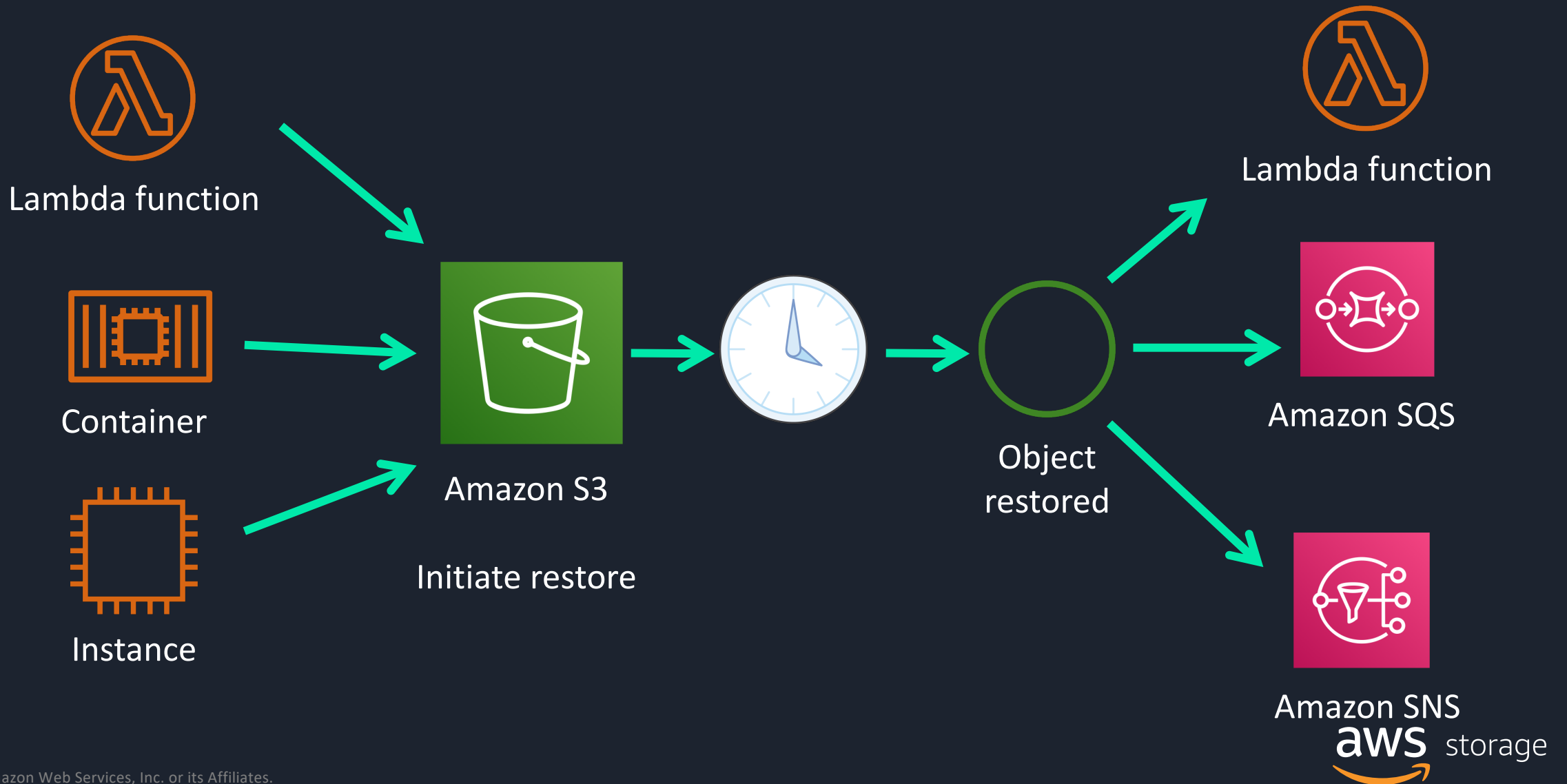
Restore before GET

- Issue a `RESTORE` request for the object (select speed, duration of restore)
- S3 restores a copy of the object into; notification issued
- Issue a `GET` request for the object (millisecond access)

Use restore notification to drive workflows through SNS, SQS, or Lambda functions

- ```
$ aws s3api restore-object --bucket mybucket --key dir1/example.obj --restore-request '{"Days":25,"GlacierJobParameters":{"Tier":"Standard"}}'
```

# Glacier restore – Architecture pattern



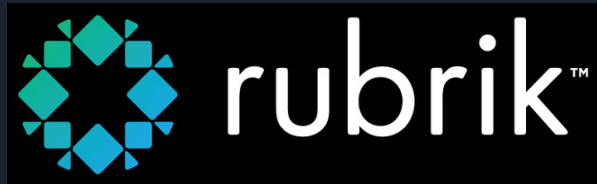
# Multiple restore speeds

| Retrieval tier | S3 Glacier | S3 Glacier Deep Archive |
|----------------|------------|-------------------------|
| Expedited      | 1–5 mins*  | not available           |
| Standard       | 3–5 hours  | within 12 hours         |
| Bulk           | 5–12 hours | within 48 hours         |



# AWS Partner Network integrations

# Archive partner integrations



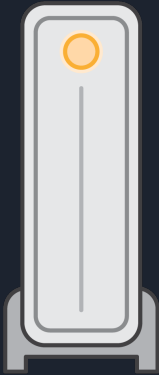
COMMVAULT®





# Economics

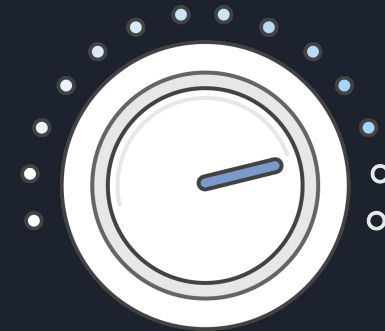
# Example workflow



250TB stored



50,000 objects  
(avg. size 25GB)



0.5% accessed  
monthly

# Long-term retention with standard retrieval costs

## S3 Glacier Deep Archive

- PUT requests
  - \$2.50 (50,000 @ \$0.05/1k)
- Storage
  - \$3,041.28 (250TB @ \$0.00099/GB-mo)
- Standard retrieval requests
  - \$0.30 (250 @ \$0.05/1k per mo)
- Standard retrieval bytes
  - \$307.20 (1.25TB @ \$0.01/GB per mo)
- Total cost (12 months)
  - \$3,351.28

## S3 Glacier

- PUT requests
  - \$2.50 (50,000 @ \$0.05/1k)
- Storage
  - \$12,288.00 (250TB @ \$0.004/GB-mo)
- Standard retrieval requests
  - \$0.15 (250 @ \$0.10/1k per mo)
- Standard retrieval bytes
  - \$153.60 (1.25TB @ \$0.02/GB per mo)
- Total cost (12 months)
  - \$12,444.25

# Long-term retention with bulk retrieval costs

## S3 Glacier Deep Archive

- PUT requests
  - \$2.50 (50,000 @ \$0.05/1k)
- Storage
  - \$3,041.28 (250TB @ \$0.00099/GB-mo)

- Bulk retrieval requests
  - \$0.075 (250 @ \$0.025/1k per mo)
- Bulk retrieval bytes
  - \$38.40 (1.25TB @ \$0.0025/GB per mo)
- Total cost (12 months)
  - \$3,082.26

## S3 Glacier

- PUT requests
  - \$2.50 (50,000 @ \$0.05/1k)
- Storage
  - \$12,288.00 (250TB @ \$0.004/GB-mo)

- Bulk retrieval requests
  - \$0.075 (250 @ \$0.025/1k per mo)
- Bulk retrieval bytes
  - \$38.40 (1.25TB @ \$0.0025/GB per mo)
- Total cost (12 months)
  - \$12,328.96

# Questions