

The Amazon S3 Storage Classes and Cost Optimization

Matt Sidley, Sr. Product Manager, Amazon S3

January, 2019

Agenda

Amazon S3 and AWS Storage overview

Pillars of cost optimization

Data placement

Introducing S3 Intelligent-Tiering



Cost optimization patterns

Putting it all together

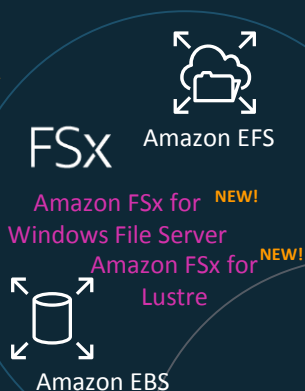
More choice for more applications

Block storage



- General Purpose SSD
- Provisioned IOPS SSD
- Throughput-Optimized HDD
- Cold HDD
- Elastic Volumes

Re-host



Re-architect

File storage



EFS Standard

EFS Infrequent Access

COMING SOON!

Object storage



S3 Standard

S3 Standard-IA

S3 One Zone-IA

S3 Glacier

S3 Intelligent-Tiering

NEW!

S3 Glacier Deep Archive

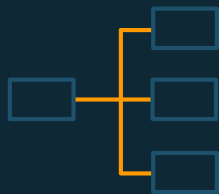
COMING SOON!

Pillars of cost optimization

Pillars of Cost Optimization



Application
Requirements



Data
Organization

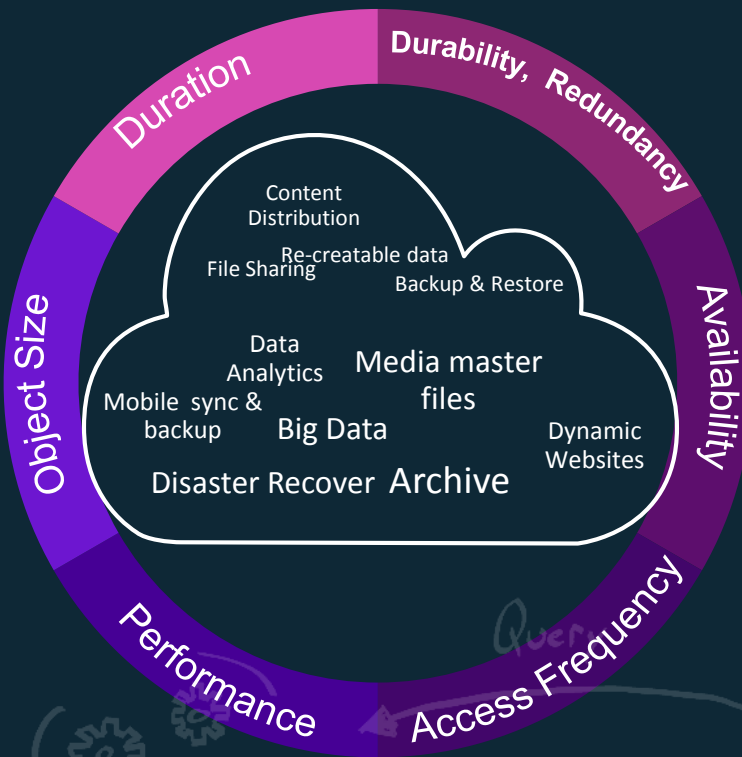


Right
Sizing

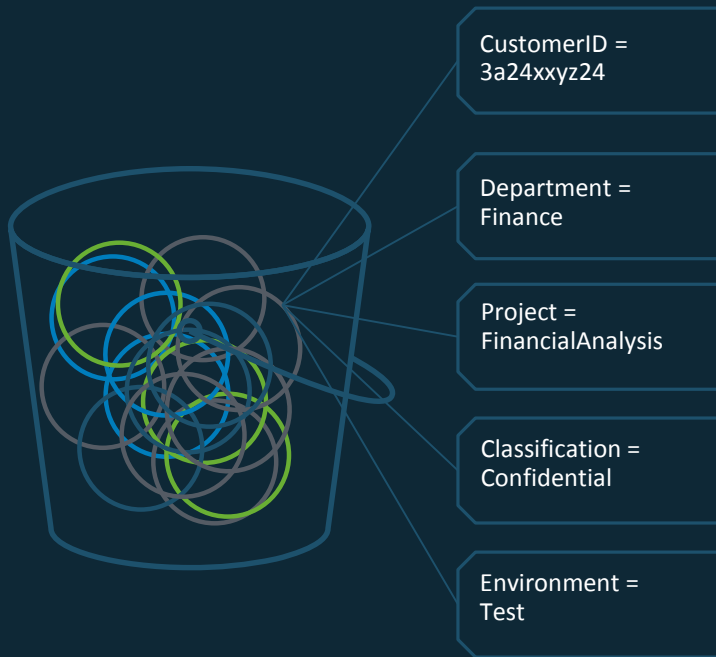


Monitor,
Optimize, Repeat

Define application requirements



Organize data with Object Tags and Prefixes



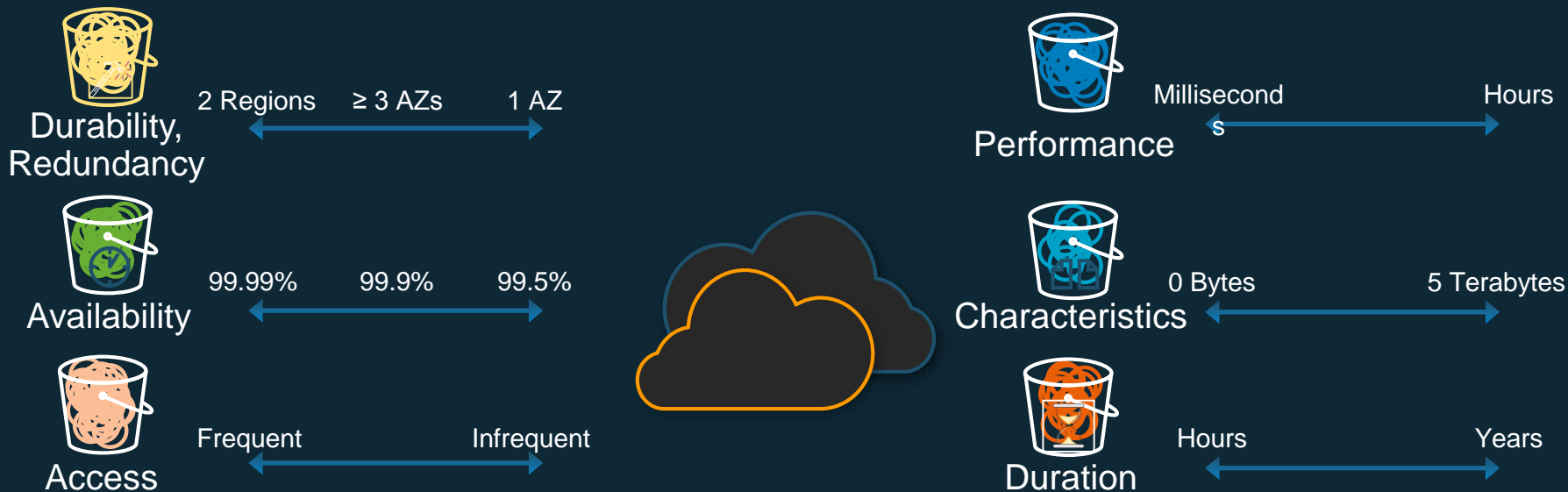
Control access, analyze usage, manage lifecycle policies, and replicate objects

Up to 10 mutable metadata tags (key value pair) per object

Completely customizable (Dept., Project, Environment, etc.)

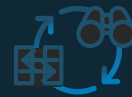
Tag objects when created, later, or both

Choose the storage class that fits best



Reduce storage cost > 80% by choosing the storage class option that best fits your use case

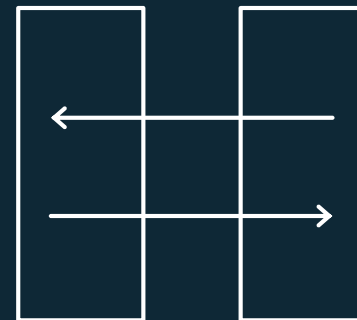
Monitor, analyze, and optimize ...



Monitor with S3 Inventory, Amazon CloudWatch, S3 Server Access Logging



Understand access patterns with **S3 Storage Class Analysis**



Tier and expire storage with **S3 Lifecycle** policies

... or just let **S3 Intelligent-Tiering** do the work and you save on storage costs automatically

S3 Storage class options

AWS pricing principles



No upfront
investment



Pay-as-you-go
approach



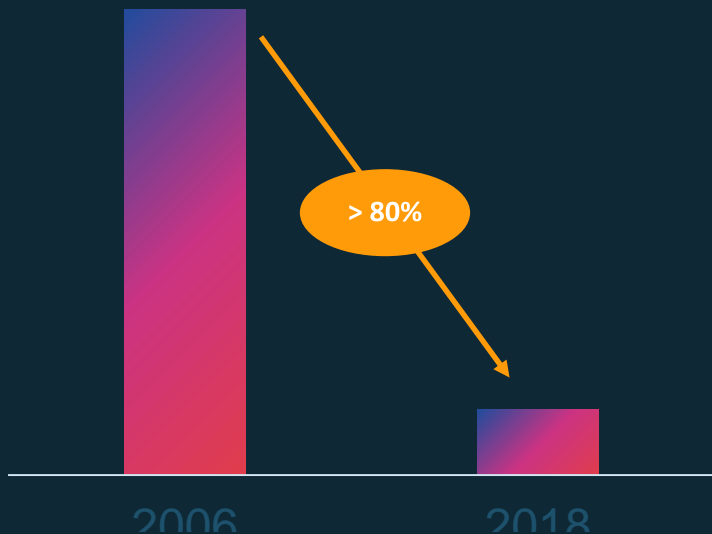
Pay less by
using more



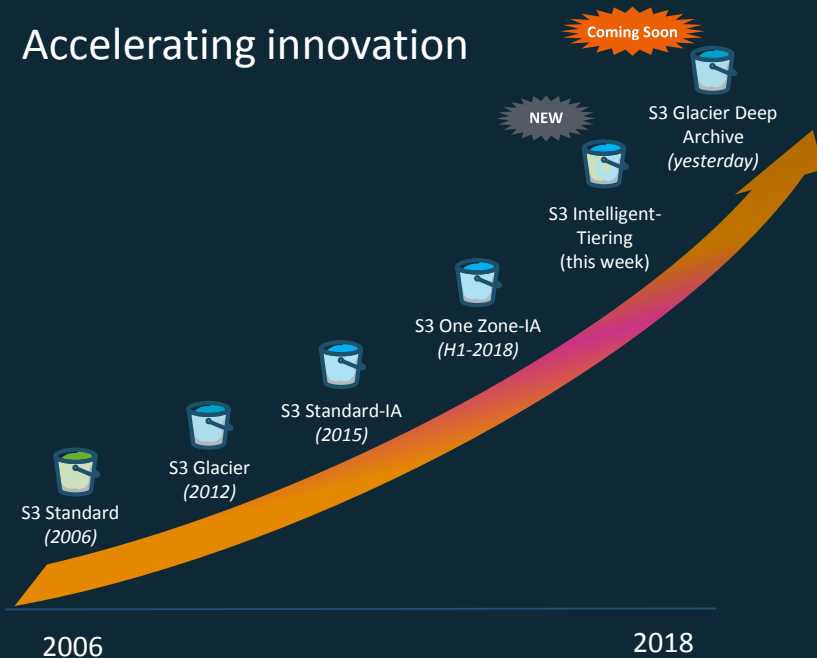
Pay less as AWS
grows

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

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

Data placement

Customers save millions of dollars annually with Storage
Class Analysis and
Lifecycle Management



S3 Storage Class Analysis helps optimizing cost



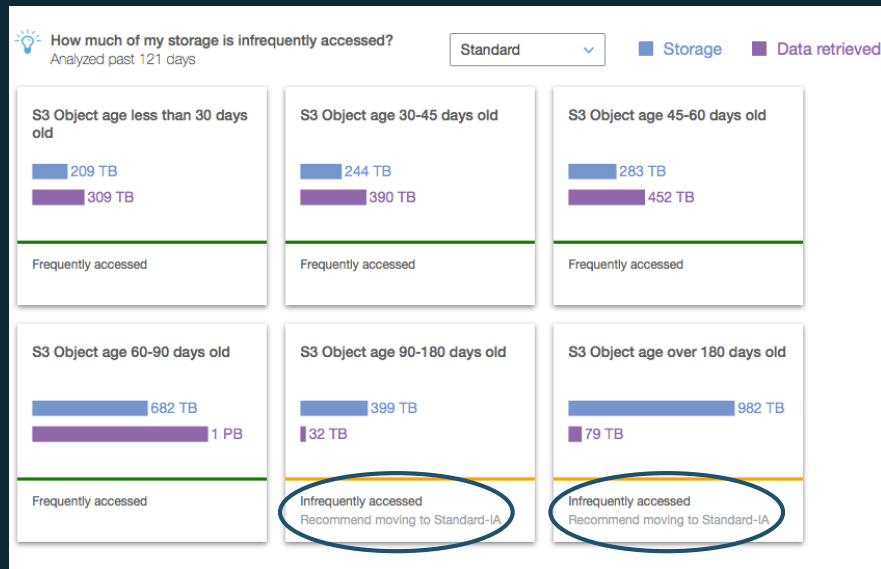
Identifies **storage age groups** that are less frequently accessed

Recommends storage age groups to **lifecycle from Standard to S-IA**

Analyzes **retrievals against storage**

Great for **predictable workloads**
(object age indicates access frequency)

Fine tune analysis by bucket, prefix, or object tag

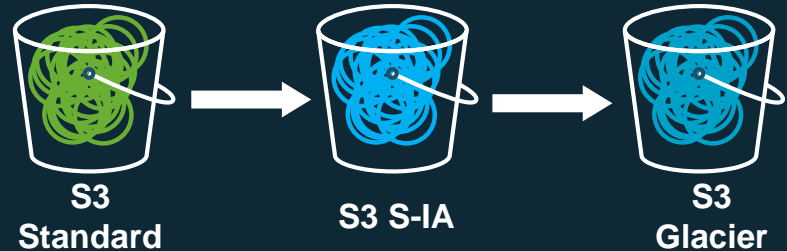


Set S3 Lifecycle Policy to tier and expire storage

S3 Lifecycle Policy to tier to lower cost storage classes and expire storage

S3 Storage Class Analysis results help set up a S3 Lifecycle Policy

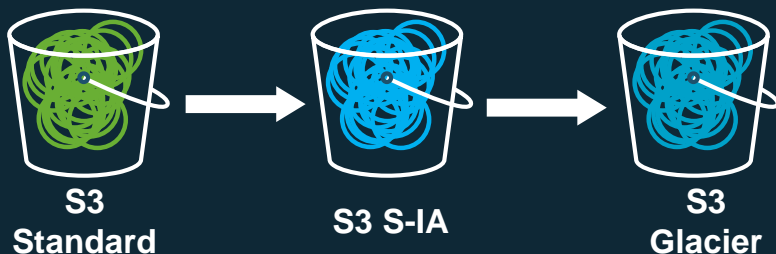
Policies are based on age of object and set by bucket, prefix, or object tag



Lifecycle Management Example Policies

Lifecycle rules take action based on object age:

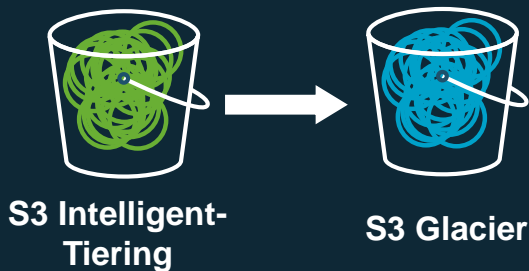
- Move all objects older than 60 days to S3 S-IA, move all objects older than 180 days to S3 Glacier



Lifecycle Management Example Policies

Lifecycle rules take action based on object age:

- Move all objects older than 180 days to S3 Glacier



Lifecycle Management Example Policies

Lifecycle rules take action based on object age:

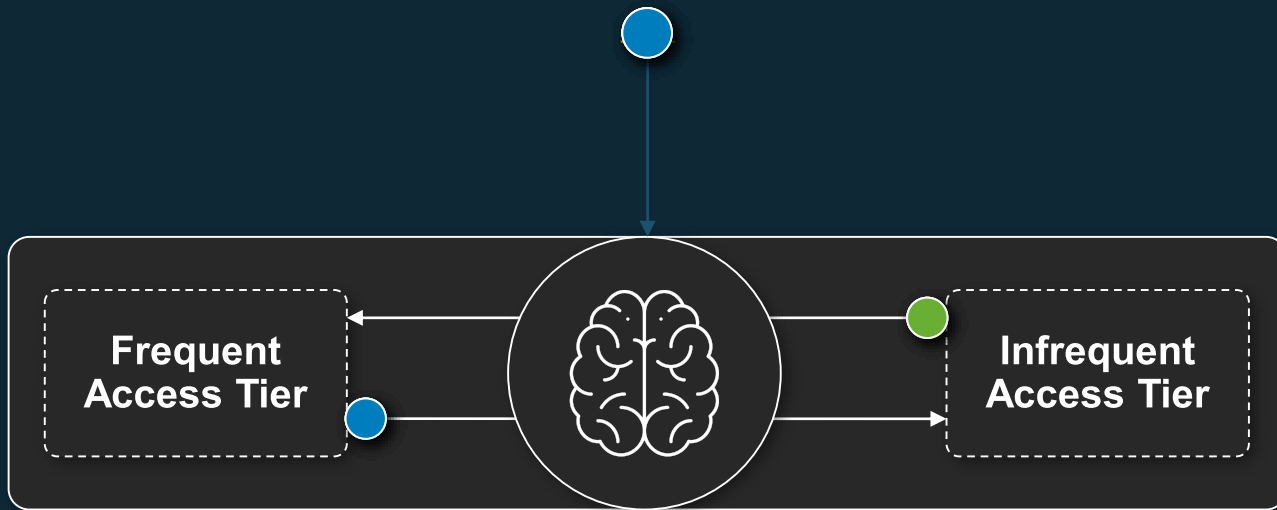
- Move all objects older than 180 days to S3 Glacier, move all objects older than 365 days to S3 Glacier Deep Archive



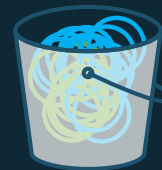
Introducing S3 Intelligent-Tiering



S3 Intelligent-Tiering automates cost savings



The story behind S3 Intelligent-Tiering



Heavy Lifting
Fragmented applications,
constraints on resources
and experience

Unmatched experience
>1M S3 customers,
Trillions of objects,
Millions of requests per second

Amazon Machine Learning
predict future access patterns,
inform storage of objects in most cost-
effective way

S3 Intelligent-Tiering

New cloud storage class that automates cost savings for customers

S3 Intelligent-Tiering storage class



Automatically optimizes storage costs for data with changing access patterns

Moves objects between two storage tiers:

- Frequent access tier optimized for frequent use of data
- Lower cost infrequent access tier optimized for less accessed data

Monitors access patterns and auto-tiers on granular object level

No performance impact, no operational overhead

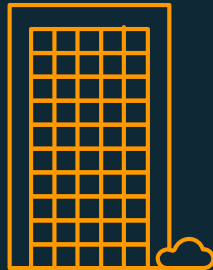
Milliseconds access, ≥ 3 AZs, Monitoring fee per Object, Minimum storage duration

Ideal use cases for S3 Intelligent-Tiering



Big Data, Data Lakes

Storage with changing access patterns used by multiple applications



Enterprises

Storage accessed by fragmented applications from various organizations



Startups

Constraint on resources and experience to optimize storage themselves

Dynamic cost optimization with no performance impact and no operational overhead

S3 Intelligent-Tiering Cost Savings Example

Assumptions:

- 10 PB of Data
- US-East-1
- Minimum Object Size of 128KB

Figure 1. Cost Savings for S3 Intelligent-Tiering vs. S3 Standard

		Average Object Size in S3 Intelligent-Tiering					
		128 KB	512 KB	1 MB	10 MB	100 MB	1 GB
Percentage of Data in S3 Intelligent-Tiering IA Tier	10%	-93%	-20%	-8%	3%	4%	4%
	20%	-89%	-16%	-4%	7%	8%	8%
	30%	-85%	-12%	0%	11%	12%	12%
	40%	-81%	-8%	4%	15%	16%	16%
	50%	-77%	-4%	8%	19%	20%	20%
	60%	-73%	0%	12%	23%	24%	24%
	70%	-69%	4%	16%	27%	28%	28%
	80%	-65%	8%	20%	31%	32%	32%
	90%	-61%	12%	24%	35%	36%	37%
	100%	-57%	16%	28%	39%	41%	41%

Cost optimization patterns

The power of S3 storage classes



S3 Glacier Deep Archive

Industry-leading performance, scalability, availability, and durability

Unmatched security, compliance, and audit capabilities

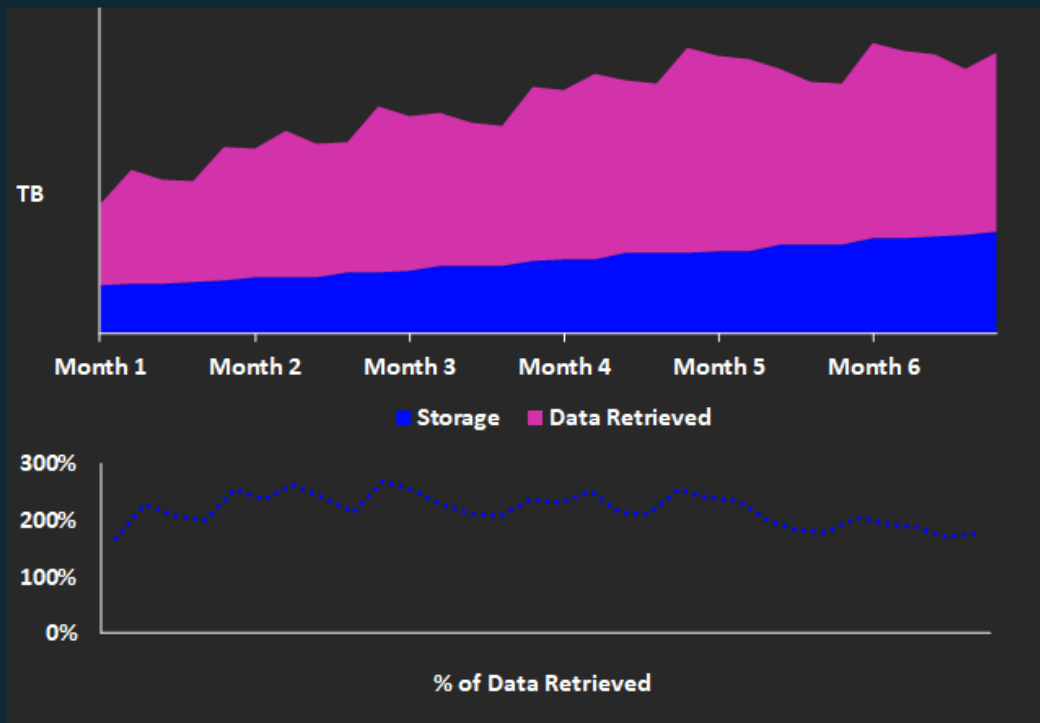
Cost optimization on a granular object-level

Storage classes that fit your workload

Set of analytics services such as S3 Select, Athena, Redshift, and EMR

Optimize your storage cost by utilizing all S3 storage classes

Workload pattern 1 – frequently accessed data



Workload characteristics:

- Frequently accessed storage (>100% of storage retrieved)
- Sometimes small objects (avg. object size ~KB)
- Storage duration sometimes short

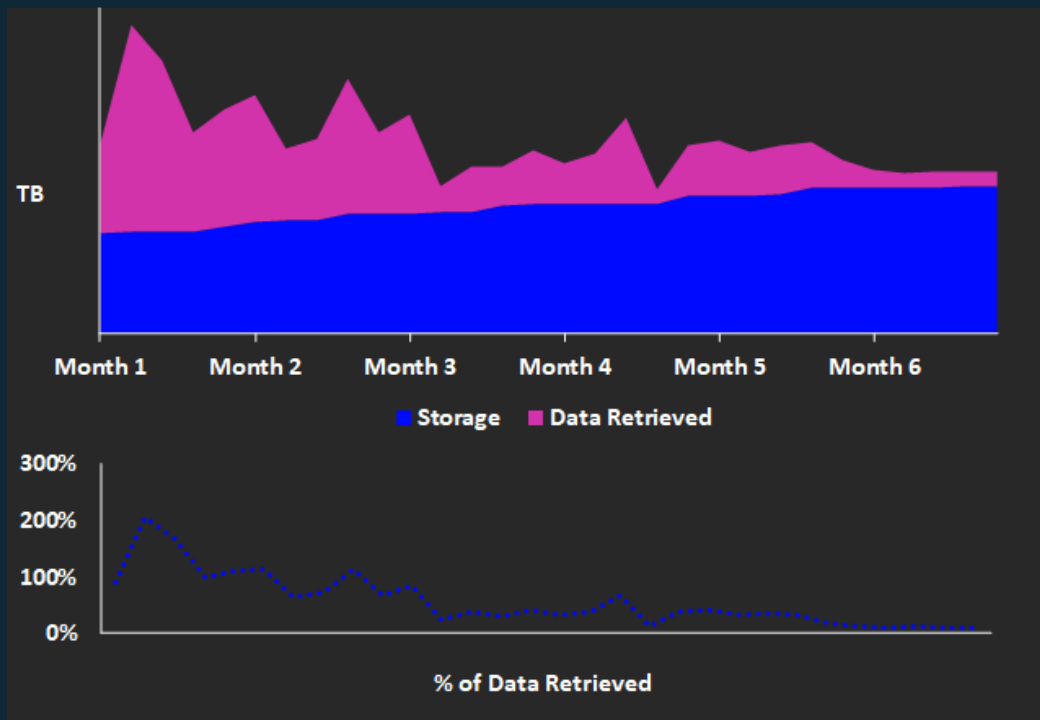
Common use cases:

- Big data analytics, dynamic website hosting, IoT sensor data, DNA sequencing, financial simulations, origin storage for CDN

Storage classes:

- S3 Standard, maybe S3 Intelligent-Tiering

Workload pattern 2 – infrequently accessed data



Workload characteristics:

- Over time infrequently accessed storage (<100% of storage retrieved after 90 days)
- Large objects (avg. object size ~MB)
- Storage duration long term

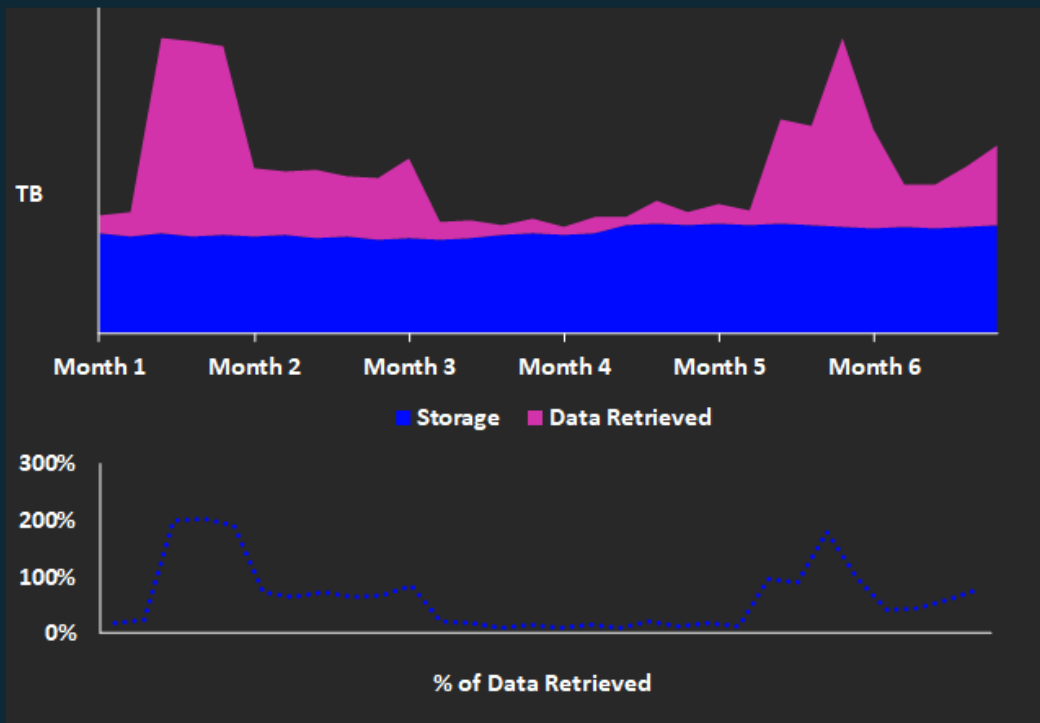
Common use cases:

- Mobile sync & backup, data logs, media assets for gaming, customer generated content, data stored for disaster recovery

S3 storage classes:

- Lifecycle from S3 Standard to S3 Standard-IA or S3 One Zone-IA for re-creatable data
- Use S3 Intelligent-Tiering for automated tiering
- Use S3 Glacier for archive

Workload pattern 3 – data with changing access



Workload characteristics:

- Data with changing or unpredictable access patterns
- Mix of object sizes (avg. object size ~MB)
- Storage duration long term

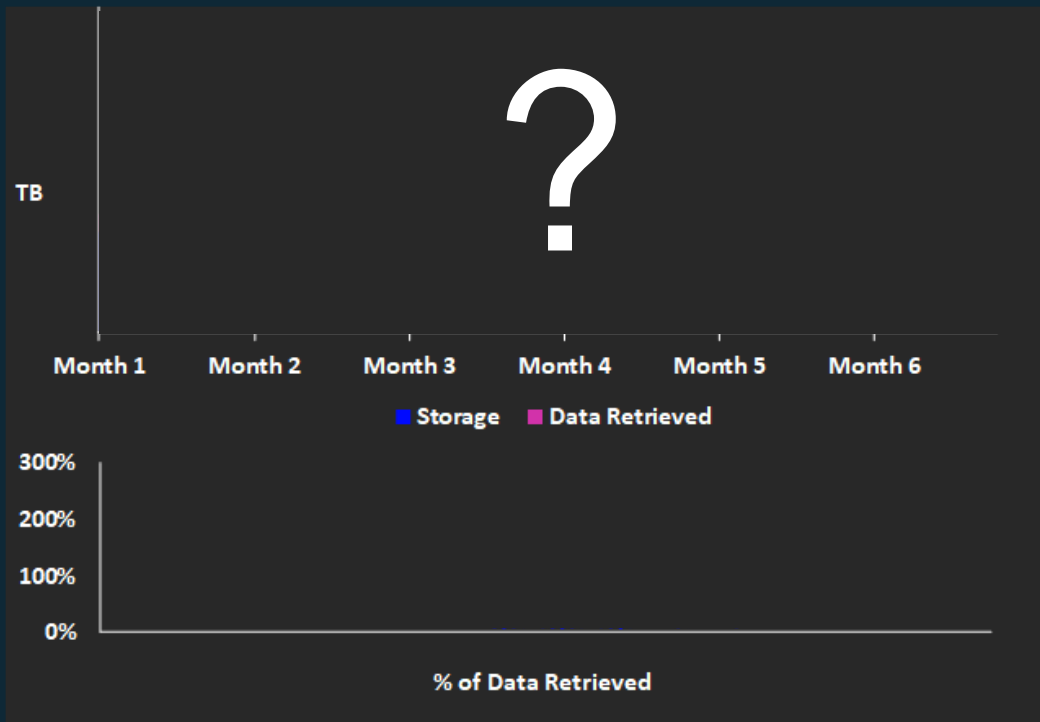
Common use cases:

- Machine Learning training data, Satellite and Geospatial imagery, Financial Transaction Records, Autonomous vehicle data, data lakes

Storage classes:

- S3 INT

Workload pattern 4 – unknown access patterns



Workload characteristics:

- Unknown workload
- You only know that objects are large (~MB) and storage duration is long (~months)

→ **S3 INT**

Workload characteristics:

- Unknown workload
- Unknown object size and short lived objects (<months)

→ **Start with S3 Standard and after some time lifecycle large objects into S3 INT**

Putting it all together

Putting it all together



Understand your application requirements



Use tags and prefixes to organize your data



Optimize across all storage classes



Cost optimize on an object level (or tag, prefix, bucket)



S3 Intelligent-Tiering for automated cost savings



AWS Cloud enables you to be more innovative, agile, and cost effective

Thank you!

Matt Sidley, Sr. Product Manager, Amazon S3