



Almacenamiento de datos seguro en AWS (S3)

</> Índice

- ¿Qué es S3?
- Características de S3
- Modelo de consistencia de datos en S3
- S3 – Almacenamiento clave-valor
- Clases de almacenamiento en S3
- Costes en S3
- Lab - S3

</> ¿Qué es AWS Simple Storage Servers (S3)?

- S3 proporciona a los desarrolladores y equipos IT un almacenamiento de datos seguro, durable y altamente escalable
- S3 es sencillo de utilizar. Se pueden gestionar nuestros datos a través de una interfaz a través de cualquier navegador



</> ¿Qué es AWS Simple Storage Servers (S3)?



- S3 es un lugar seguro donde almacenar nuestros archivos
 - S3 es un almacenamiento basado en objetos
- Los datos están repartidos entre múltiples instalaciones y dispositivos



Características de S3 (I)

S3 nos ofrece las siguientes características:

- S3 está basado en objetos – Nos permite subir archivos
- Los archivos pueden ocupar desde 0 Bytes a 5 TB
- No hay límite de almacenamiento
- Los archivos se guardan en Buckets
- El espacio de nombres (namespace) de S3 es universal. Esto es, cada nombre asociado a un Bucket tiene que ser único a nivel global
- Acceso temporal para usuarios o dispositivos cuando sea necesario
- devacademyaws2019.us-east-1.amazonaws.com
- Cuando subimos un archivo a S3, recibimos un código HTTP 200 si el archivo ha sido subido correctamente

</> Modelo de consistencia de datos en S3

- *Read after Write consistency* para **PUTS** de nuevos objetos
- *Eventual Consistency* para **PUTS** que sobrescriban objetos existentes y **DELETES** (estas operaciones pueden tardar un tiempo en propagarse)



</> S3 - Almacenamiento clave-valor

S3 no es más que un simple almacenamiento clave-valor basado en objetos:

- Clave (Es simplemente el nombre del objeto)
- Valor (Es simplemente la secuencia de bytes que conforman los datos del objeto)
- Version ID (Importante para tener un versionado de los ficheros)
- Metadata (Datos referentes a la información de lo que estamos almacenando)
- Subrecursos:
 - Listas de control de acceso (permisos sobre quién puede acceder a nuestro objeto)
 - Torrent



Características de S3 (II)

Más características de S3:

- Amazon garantiza una disponibilidad del 99.9%
- Amazon garantiza una durabilidad de del 99.9999999999%
- Hay distintas clases de almacenamiento
- Nos permite Gestionar el ciclo de vida de los datos
- Versionado
- Encriptado
- Acceso seguro a los datos
- Listas de control y políticas para los Buckets

</> Clases de almacenamiento en S3

- **S3 Standard:** 99.9% de disponibilidad, 99.9999999999% de durabilidad, se almacena de forma redundante en varios dispositivos en distintas instalaciones y está diseñado para aguantar la pérdida concurrente de dos instalaciones (\$\$\$\$)
- **S3 – IA (Infrequently Access):** Se utiliza para datos a los que se necesite acceder con menos frecuencia pero de forma rápida cuando se necesite. Comisión por obtener los datos (\$\$\$)
- **S3 One Zone – IA:** Cuando no se requiere que los datos estén en varias zonas de disponibilidad (\$\$)
- **Glacier:** Sólo se utiliza para archivar datos. Hay varios tipos. La obtención de datos para el tipo Standard tarda entre 3 y 5 horas (\$)

</> Clases de almacenamiento en S3

	S3 Estándar	S3 Estándar – Acceso poco frecuente	S3 Única zona – Acceso poco frecuente	Amazon Glacier
Diseñado para ofrecer durabilidad	99,999999999%	99,999999999%	99,999999999%†	99,999999999%
Diseñado para ofrecer disponibilidad	99,99%	99,9%	99,5%	N/D
SLA de disponibilidad	99,9%	99%	99%	N/D
Zonas de disponibilidad	≥3	≥3	1	≥3
Cargo mínimo de capacidad por objeto	N/D	128 KB*	128 KB*	N/D
Cargo mínimo por duración de almacenamiento	N/D	30 días	30 días	90 días
Tarifa de recuperación	N/D	por GB recuperado	por GB recuperado	por GB recuperado**
Latencia del primer byte	milisegundos	milisegundos	milisegundos	minutos u horas seleccionados***
Tipo de almacenamiento	Objeto	Objeto	Objeto	Objeto
Transiciones del ciclo de vida	Sí	Sí	Sí	Sí

</> Costes en S3

Cuando utilizamos S3 nos van a cobrar por:

- Cantidad de almacenamiento
- Número de peticiones
- Gestión del almacenamiento (Storage Management Pricing)
- Transferencia de datos entre regiones (Data Transfer Pricing)
- Transfer Acceleration

</> ¿Cuál es la mejor manera de aprender sobre S3?

"For the things we have to learn before we can do them, we learn by doing them"

Aristóteles

¿Qué hemos aprendido?

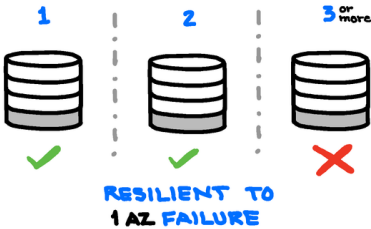
- Los Buckets tienen asignado un namespace universal. No podemos repetir nombres
- Cada vez que subimos un objeto obtenemos un código HTTP 200
- Hay distintas clases: S3, S3 – IA, S3 One Zone – IA, Glacier
- Los datos del Bucket pueden ir encriptados
- Podemos controlar el acceso a los buckets con ACLs o usando Políticas para los Buckets
- Por defecto los objetos que subimos a los Buckets son privados
- Podemos versionar los Buckets
- Podemos crear webs estáticas fácilmente

STORAGE CLASSES

STANDARD
and
INFREQUENT ACCESS

11 NINES DURABILITY

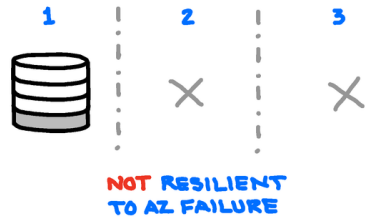
AVAILABILITY ZONES



1-ZONE INFREQUENT ACCESS

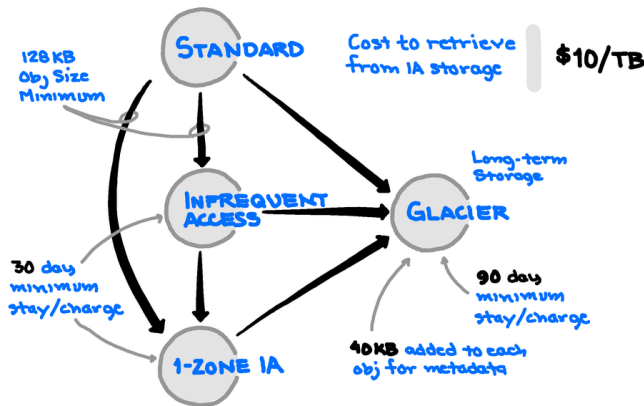
11 NINES DURABILITY

AVAILABILITY ZONES



99.5% / YR AVAILABILITY

STORAGE CLASS TRANSITIONS



SERVER SIDE ENCRYPTION



QUERY IN PLACE

AMAZON S3

OBJECT STORAGE



S3 SELECT



Retrieve subset of data
Use simple SQL

Cost:

\$2/TB scanned
70¢/TB returned } Standard Storage

\$10/TB scanned
70¢/TB returned } IA Storage

OBJECT LOCKING

NEW

PREVENT OBJECT DELETION

GOVERNANCE

CAN'T DELETE TIL DATE

COMPLIANCE

NEVER DELETE (EVER)



jerry@lucidchart



@awsgeek

[all pricing from US West (Oregon)]

REQUESTS (\$USD per 1,000,000 requests)



	GET	POST	PUT	COPY	LIST	SELECT	other
Standard	0.4	5.0	5.0	5.0	5.0	0.4	0.4
Infrequent Access	1.0	10.0	10.0	10.0		1.0	1.0
1-Zone IA	1.0	10.0	10.0	10.0		1.0	1.0

Storage class transitions : 1¢ per 1,000 objects

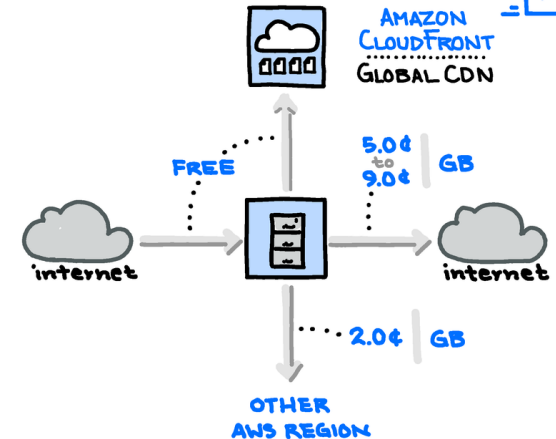
Transition to Glacier : 5¢ per 1,000 objects

STORAGE

	TB	\$
INFREQUENT ACCESS	[0, 50]	2.3¢
	[51, 500]	2.2¢
	[501, ∞]	2.1¢
STD 1-ZONE		1.25¢
		1.0¢
GLACIER		0.4¢

per GB/Mo

DATA TRANSFER





Consejos - S3

- Definir bien los permisos <https://aws.amazon.com/es/blogs/aws/aws-config-update-new-managed-rules-to-secure-s3-buckets/>
- No debemos usar . en los nombres por tema de certificados
- Podemos activar el versionado si queremos tener versionado de nuestros elementos. No tiene todas las funcionalidades de un control de versiones
- Transferencias de S3 a EC2 en la misma región son gratis
- Entender los tipos de S3 y cuál es el más adecuado (IA para logs, Glacier datos antiguos, etc.) Tener en cuenta las integraciones con CloudFront
- Podemos utilizarlo para alojar páginas estáticas
- Podemos habilitar la replicación en distintas regiones
- Si borramos un Bucket, su nombre puede tardar horas en liberarse

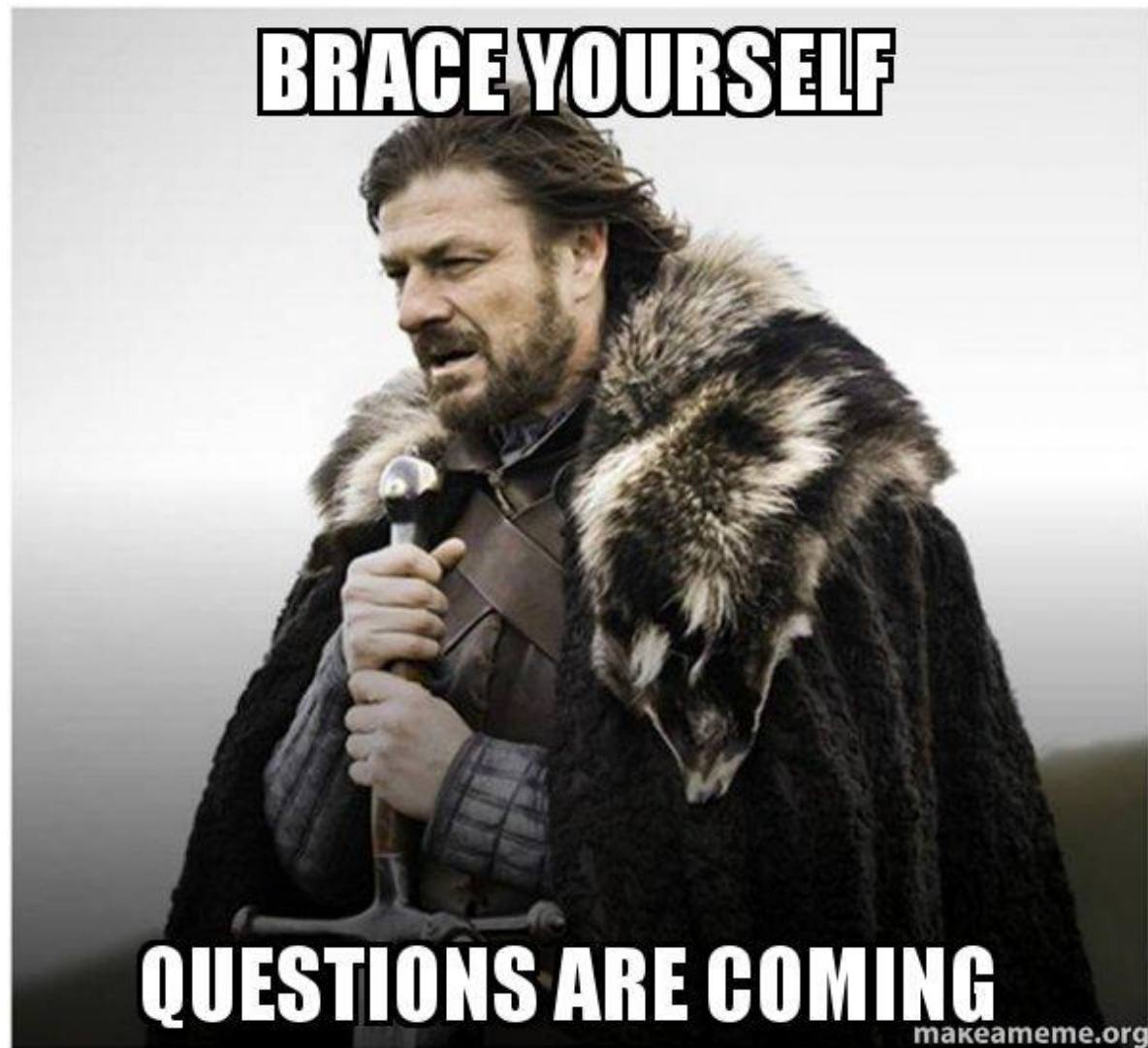
</> Importante - S3

- Hay un límite de 1000 Buckets por cuenta
- Nunca asumir que las operaciones se realizan de forma secuencial
- El SLA no es muy alto (99.9%)
- Cuidado con la consistencia eventual

¿Qué hemos aprendido en este módulo?

- Conocer qué es S3 y para qué utilizarlo
- Conocer las características de S3
- Entender el modelo de consistencia de datos de S3
- Entender S3 como un almacenamiento de clave-valor
- Conocer las diferencias entre las distintas clases de almacenamiento para S3
- Conocer los costes de S3
- Aprender a usar el servicio de S3 de AWS


</> Preguntas





- Unai Arrien
- Email de contacto: *unai.arrien@gmail.com*

info@devacademy.es 

687374918 

@DevAcademyES 