A CLOUD GURU

S3 Object Lock & Glacier Vault Lock

Ryan Kroonenburg
AWS COMMUNITY HERO & ALEXA CHAMPION
FOUNDER OF A CLOUD GURU

What Is S3 Object Lock?

A CLOUD GURU

You can use S3 Object Lock to store objects using a **write once, read many (WORM)** model. It can help you prevent objects from being deleted or modified for a fixed amount of time or indefinitely.

You can use **S3 Object Lock** to meet regulatory requirements that require WORM storage, or add an extra layer of protection against object changes and deletion.

S3 OBJECT LOCK & GLACIER VAULT LOCK

S3 Object Lock Modes: Governance Mode

A CLOUD GURU

Governance Mode

In governance mode, **users can't overwrite or delete an object version or alter its lock settings** unless they have special permissions. With governance mode, you protect objects against being deleted by most users, but you can still grant some users **permission to alter the retention settings** or delete the object if necessary.

S3 OBJECT LOCK & GLACIER VAULT LOCK

S3 Object Lock Modes: Compliance Mode

A CLOUD GURU

Compliance Mode

In compliance mode, **a protected object version can't be overwritten or deleted by any user**, including the root user in your AWS account. When an object is locked in compliance mode, its retention mode can't be changed and its retention period can't be shortened. Compliance mode ensures an object version **can't be overwritten or deleted** for the duration of the retention period.

S3 OBJECT LOCK & GLACIER VAULT LOCK

Retention Periods

Retention Periods

A retention period **protects an object version for a fixed amount of time**. When you place a retention period on an object version, Amazon S3 stores a timestamp in the object version's metadata to indicate when the retention period expires. After the retention period expires, the object version can be **overwritten or deleted** unless you also placed a legal hold on the object version.

S3 OBJECT LOCK & GLACIER VAULT LOCK

Legal Holds

A CLOUD GURU

Legal Holds

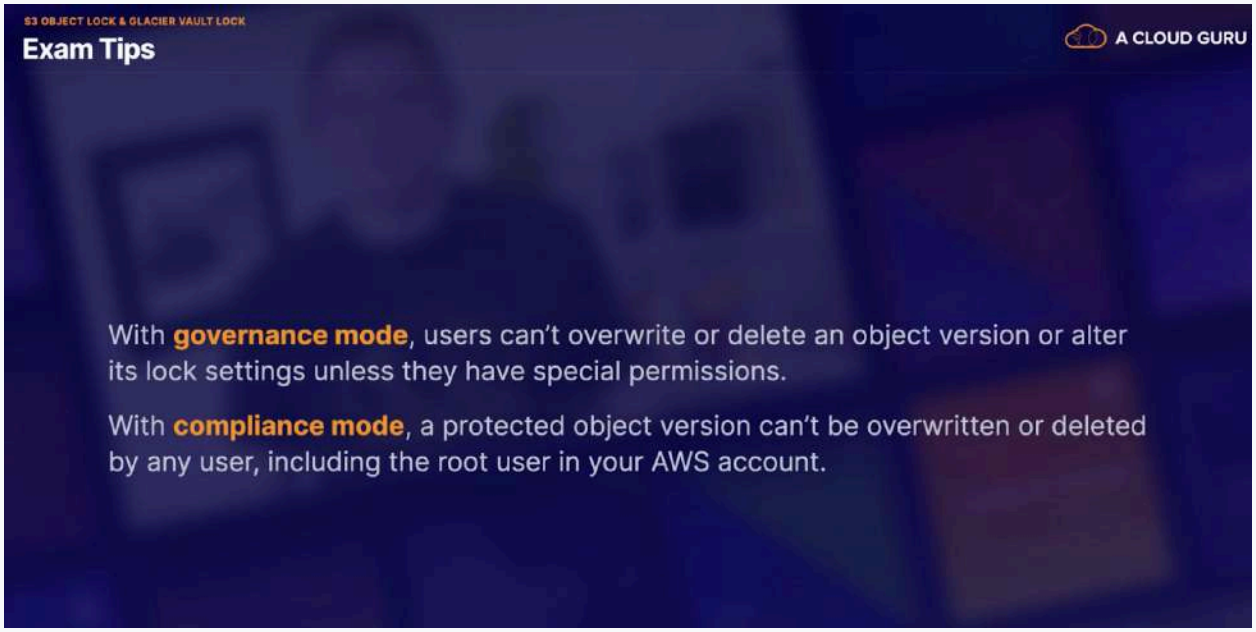
S3 Object Lock also enables you to place a legal hold on an object version. Like a retention period, **a legal hold prevents an object version from being overwritten or deleted**. However, a legal hold doesn't have an associated retention period and remains in effect until removed. Legal holds can be freely placed and removed by any user who has the `s3:PutObjectLegalHold` permission.

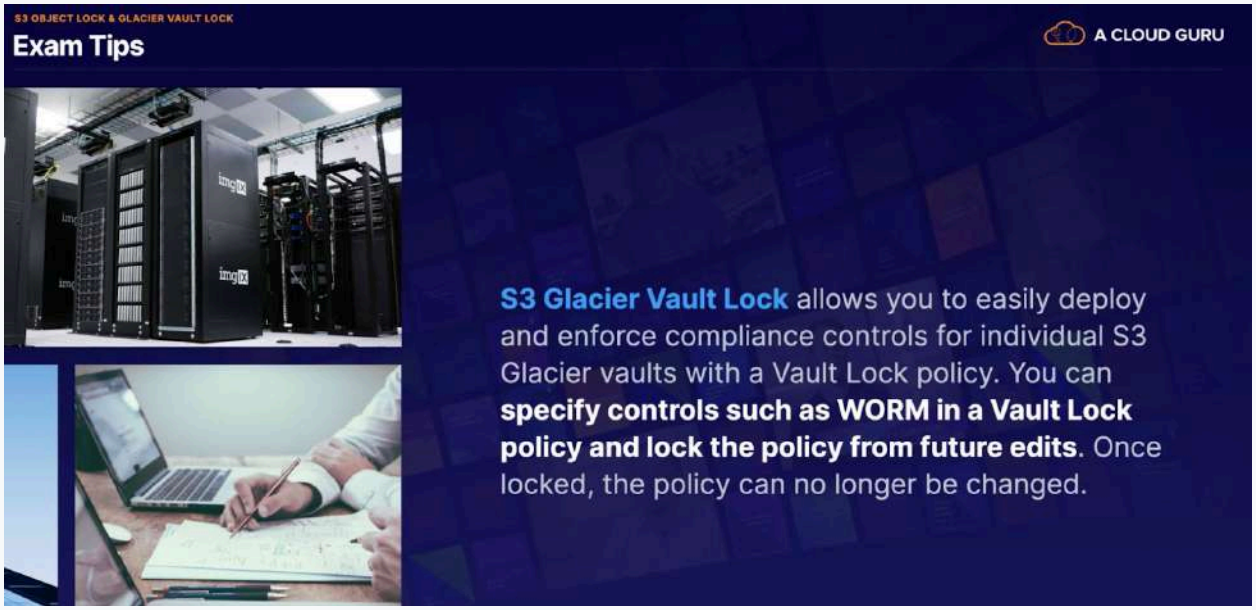
Glacier Vault Lock

S3 Glacier Vault Lock allows you to **easily deploy and enforce compliance controls for individual S3 Glacier vaults with a Vault Lock policy**. You can specify controls, such as WORM, in a Vault Lock policy and lock the policy from future edits. Once locked, the policy can no longer be changed.

Exam Tips

- Use **S3 Object Lock** to store objects using a write once, read many (WORM) model.
- Object locks can be on individual objects or **applied across the bucket** as a whole.
- Object locks come in two modes: **governance mode** and **compliance mode**.





S3 Performance [SAA-C02]

A CLOUD GURU

S3 Performance

Ryan Kroonenburg
AWS COMMUNITY HERO & ALEXA CHAMPION
FOUNDER OF A CLOUD GURU

S3 PERFORMANCE

What Are S3 Prefixes?

A CLOUD GURU

S3 Prefix

- ✓ *mybucketname/folder1/subfolder1/myfile.jpg* > **/folder1/subfolder1**
- ✓ *mybucketname/folder2/subfolder1/myfile.jpg* > **/folder2/subfolder1**
- ✓ *mybucketname/folder3/myfile.jpg* > **/folder3**
- ✓ *mybucketname/folder4/subfolder4/myfile.jpg* > **/folder4/subfolder4**

S3 PERFORMANCE

S3 Performance

A CLOUD GURU

S3 Performance

S3 has extremely low latency. You can get the first byte out of S3 within **100-200 milliseconds**

You can also achieve a high number of requests: **3,500 PUT/COPY/POST/DELETE** and **5,500 GET/HEAD** requests per second per prefix.


S3 PERFORMANCE
S3 Performance
A CLOUD GURU
S3 Performance
1
You can get better performance by spreading your reads across **different prefixes**. For example, if you are using **two prefixes**, you can achieve **11,000 requests per second**.
2
If we used all **four prefixes** in the last example, you would achieve **22,000 requests per second**.


S3 PERFORMANCE
KMS Request Rates
A CLOUD GURU
S3 LIMITATIONS WHEN USING KMS
• If you are using **SSE-KMS** to encrypt your objects in S3, you must keep in mind the **KMS limits**.
• When you **upload** a file, you will call **`GenerateDataKey`** in the KMS API.
• When you **download** a file, you will call **`Decrypt`** in the KMS API.

S3 PERFORMANCE

KMS Request Rates

A CLOUD GURU

S3 Limitations When Using KMS

- Uploading/downloading will count toward the **KMS quota**.
- Region-specific, however, it's either **5,500**, **10,000**, or **30,000** requests per second.
- Currently, you **cannot** request a quota increase for KMS.

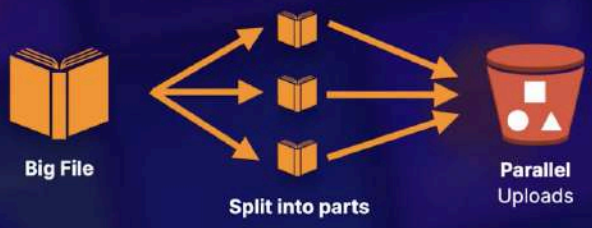
S3 PERFORMANCE

S3 Performance: Uploads

A CLOUD GURU

Multipart Uploads

- Recommended for files **over 100 MB**
- Required for files **over 5 GB**
- Parallelize uploads (increases **efficiency**)



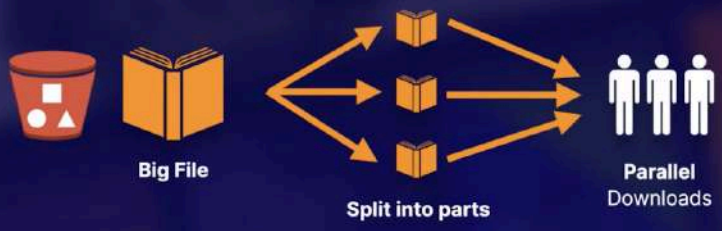
S3 PERFORMANCE

S3 Performance: Downloads

A CLOUD GURU

S3 Byte-Range Fetches

- Parallelize **downloads** by specifying byte ranges.
- If there's a failure in the download, it's only for a specific byte range.



S3 PERFORMANCE

S3 Performance: Downloads

A CLOUD GURU

S3 Byte-Range Fetches

Can be used to **speed up** downloads

Can be used to just download **partial amounts of the file** (e.g., header information)

S3 PERFORMANCE

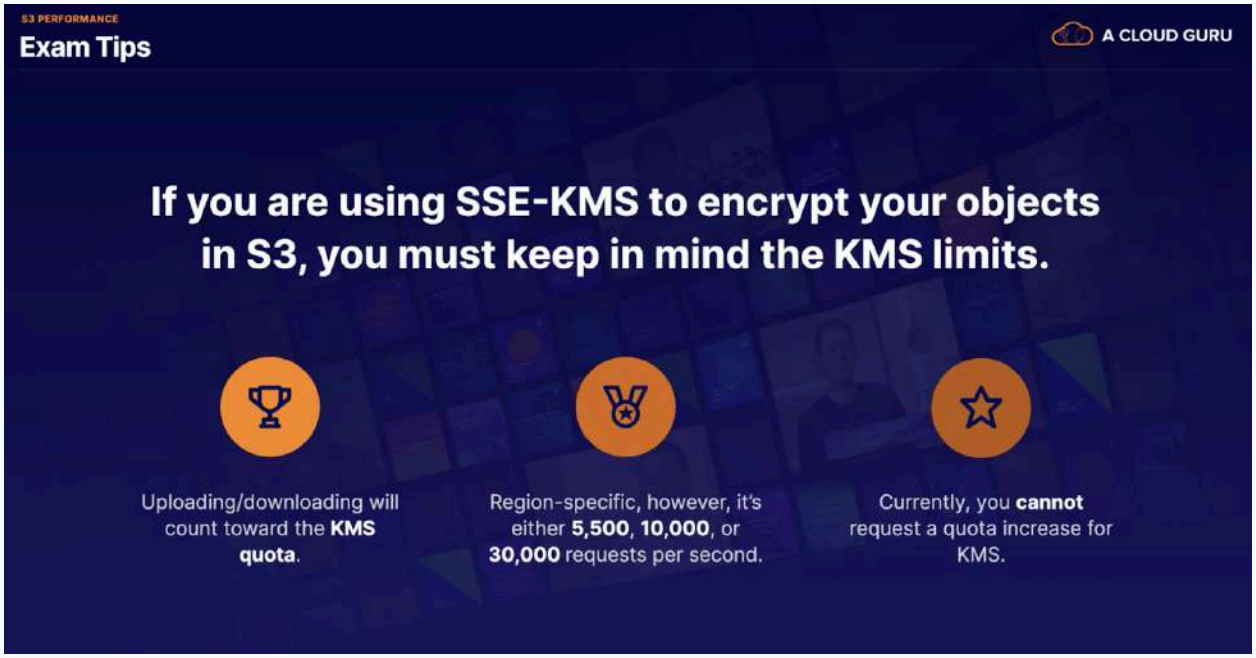
Exam Tips

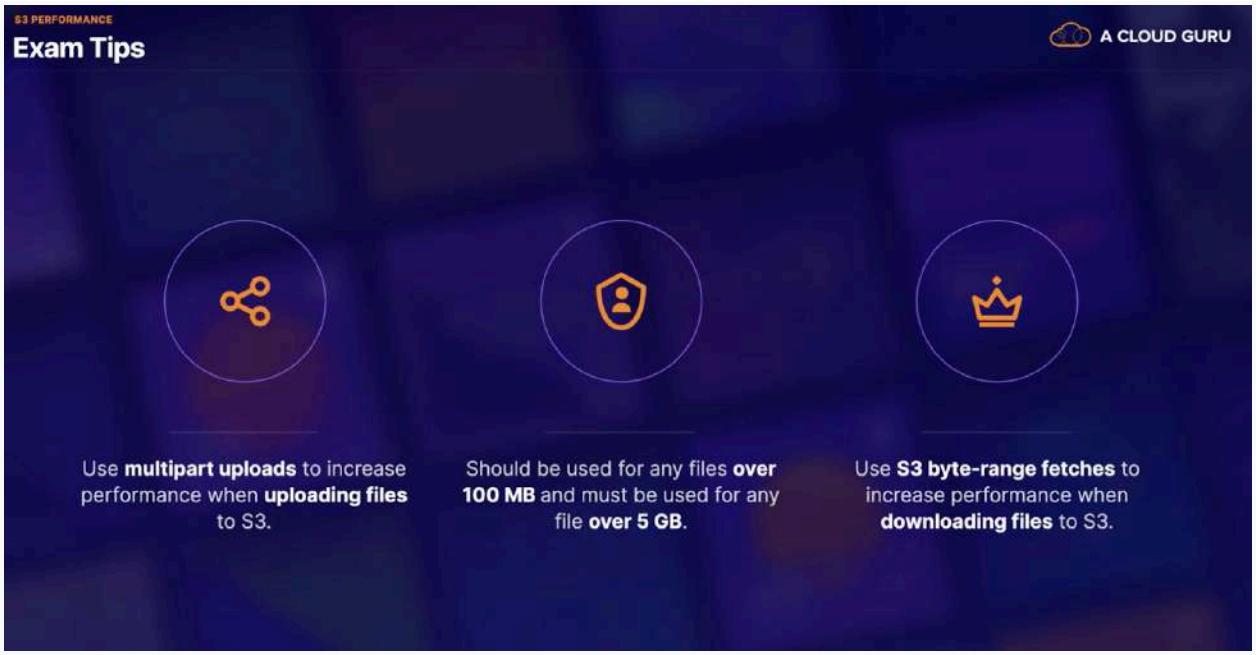
A CLOUD GURU

mybucketname/folder1/subfolder1/myfile.jpg > **/folder1/subfolder1**

You can also achieve a high number of requests: **3,500 PUT/COPY/POST/DELETE** and **5,500 GET/HEAD** requests per second per prefix.

You can get better performance by spreading your reads across **different prefixes**. For example, if you are using **two prefixes**, you can achieve **11,000 requests per second**.





S3 Select & Glacier Select [SAA-C02]

A CLOUD GURU

S3 Select & Glacier Select

Ryan Kroonenburg
AWS COMMUNITY HERO & ALEXA CHAMPION
FOUNDER OF A CLOUD GURU

What Is S3 Select?

S3 Select enables applications to retrieve only a subset of data from an object by using simple SQL expressions. By using S3 Select to retrieve only the data needed by your application, you can **achieve drastic performance increases** — in many cases, you can get as much as a 400% improvement.

S3 SELECT & GLACIER SELECT

S3 Select Example

A CLOUD GURU

Let's assume all your data is stored in S3 in zip files that contain CSV files. Without S3 Select, you would need to **download**, **decompress**, and **process the entire CSV** to get the data you needed.

S3 SELECT & GLACIER SELECT

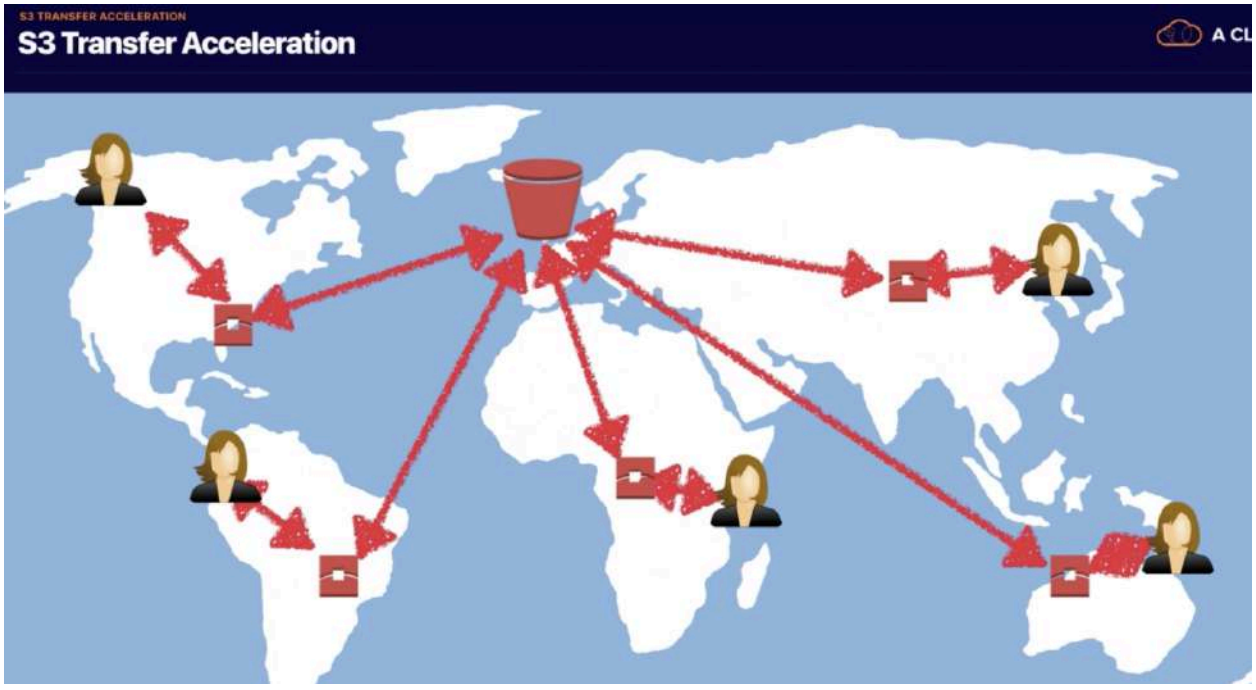
S3 Select Example

A CLOUD GURU

With S3 Select, you can use a **simple SQL expression** to return only the data from the store you're interested in instead of retrieving the entire object. This means you're dealing with an order of magnitude less data, which **improves the performance of your underlying applications**.

AWS Organizations [SAA-C02]

S3 Transfer Acceleration



CloudFront Overview

CLOUDFRONT

What is CloudFront?

A CLOUD GURU

A content delivery network (CDN) is a system of distributed servers (network) that deliver webpages and other web content to a user based on the geographic locations of the user, the origin of the webpage, and a content delivery server.

CLOUDFRONT

CloudFront - Key Terminology

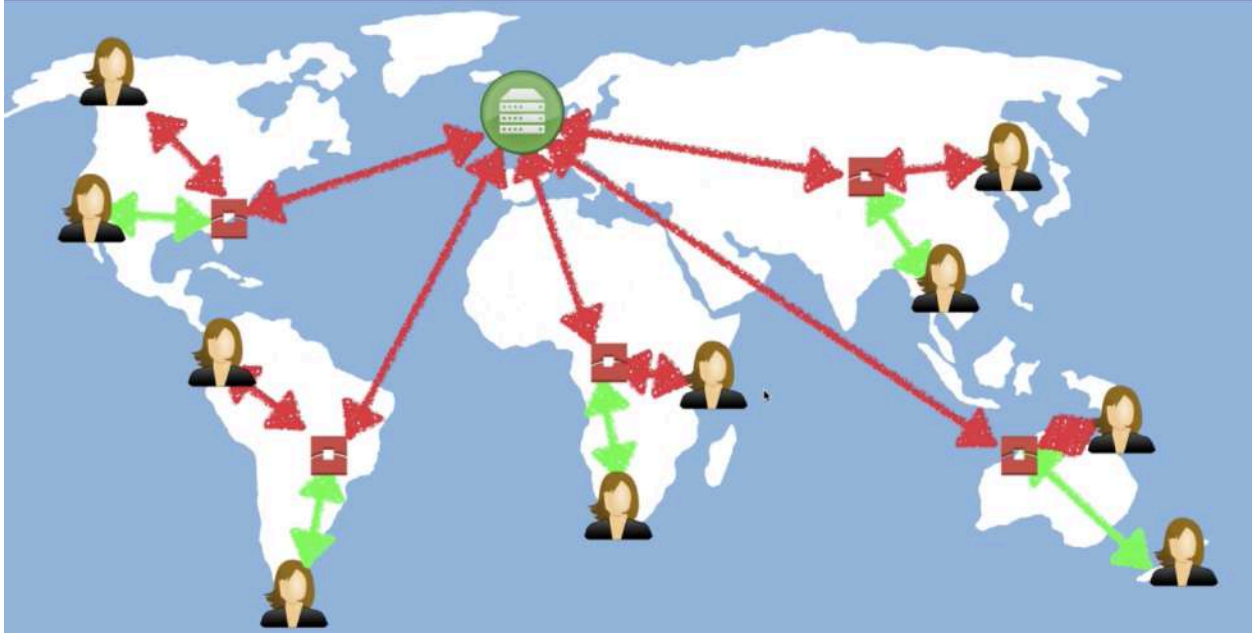
A CLOUD G

- **Edge Location** - This is the location where content will be cached. This is separate to an AWS Region/AZ.
- **Origin** - This is the origin of all the files that the CDN will distribute. This can be an S3 Bucket, an EC2 Instance, an Elastic Load Balancer, or Route53.
- **Distribution** - This is the name given the CDN which consists of a collection of Edge Locations.

CLOUDFRONT

CloudFront - How It Works

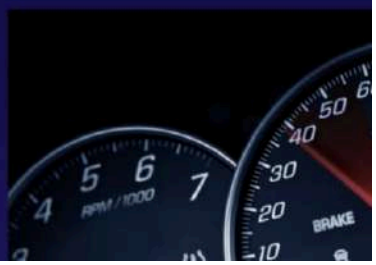
A CLOUD G



CLOUDFRONT

What is CloudFront?

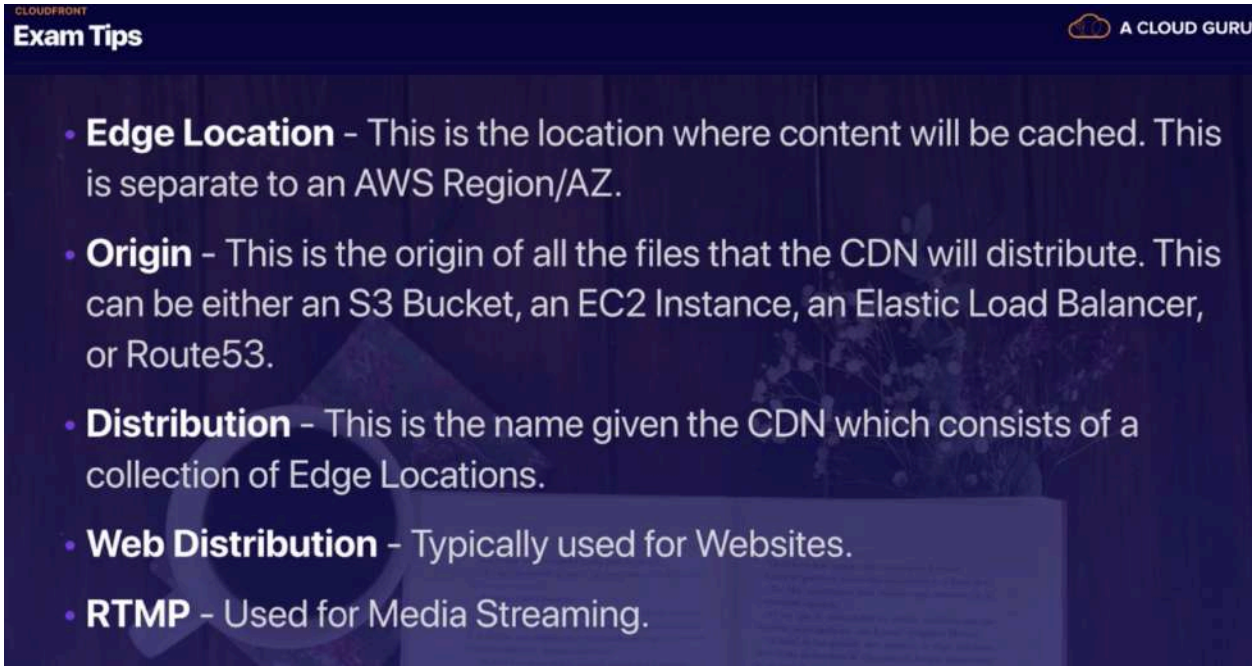
Amazon CloudFront can be used to deliver your entire website, including dynamic, static, streaming, and interactive content using a global network of edge locations. Requests for your content are automatically routed to the nearest edge location, so content is delivered with the best possible performance.

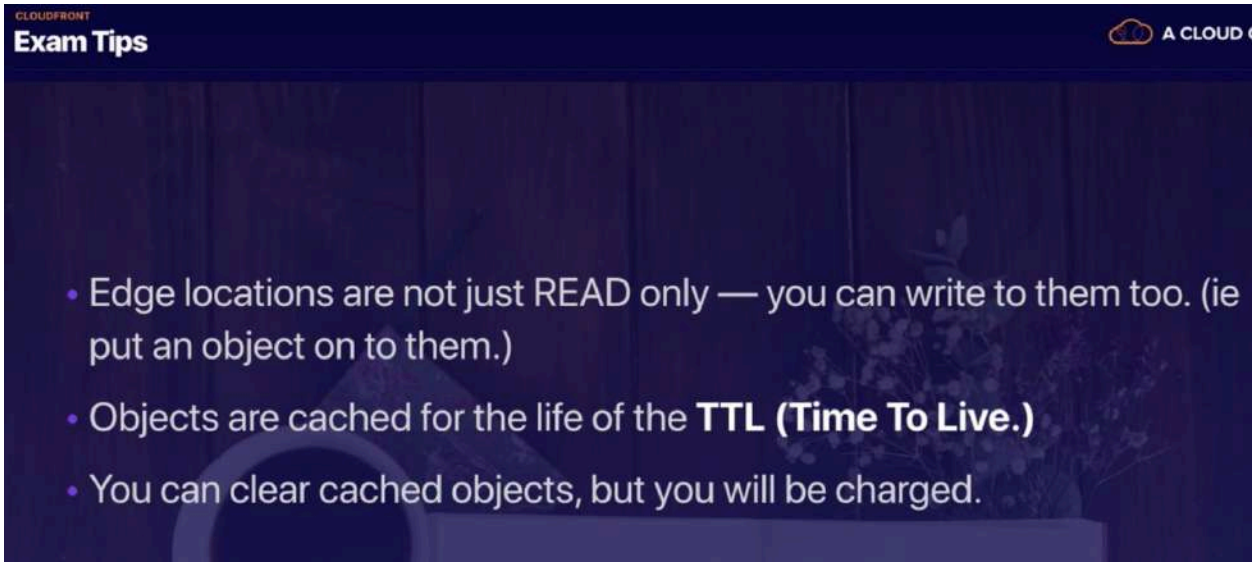


CLOUDFRONT

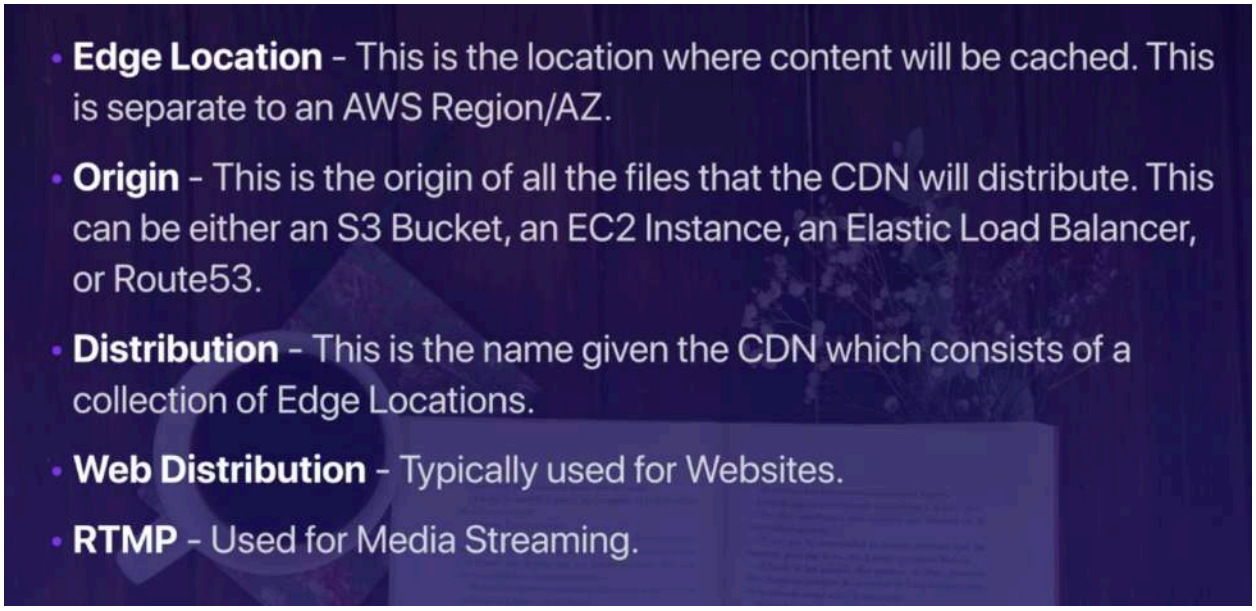
CloudFront - Key Terminology

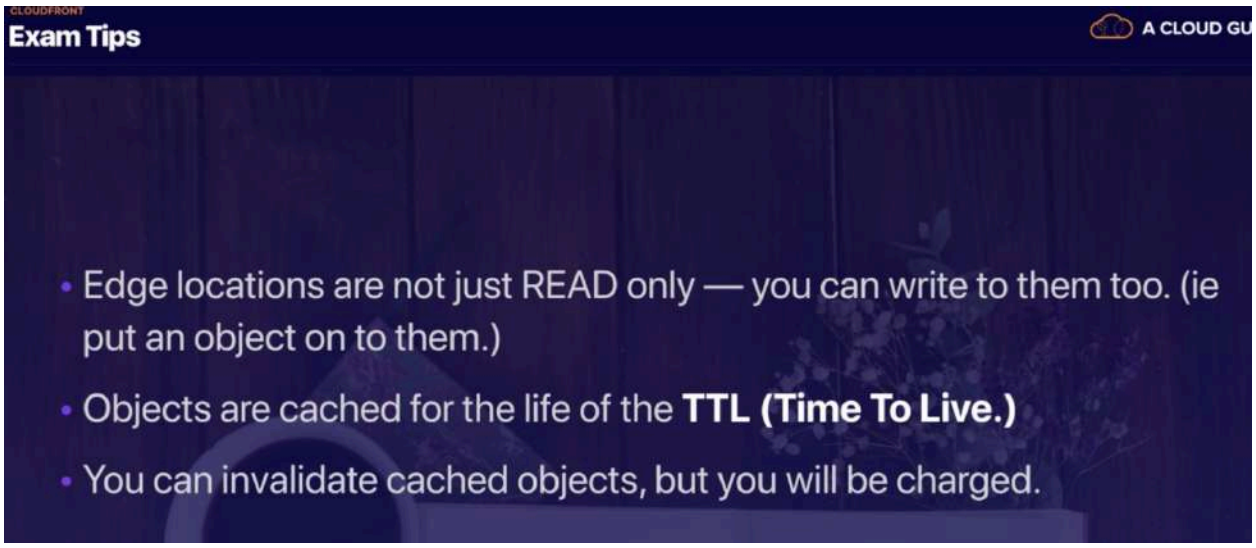
- Web Distribution - Typically used for Websites.
- RTMP - Used for Media Streaming.





CloudFront Lab


• Edge Location - This is the location where content will be cached. This is separate to an AWS Region/AZ.
• Origin - This is the origin of all the files that the CDN will distribute. This can be either an S3 Bucket, an EC2 Instance, an Elastic Load Balancer, or Route53.
• Distribution - This is the name given the CDN which consists of a collection of Edge Locations.
• Web Distribution - Typically used for Websites.
• RTMP - Used for Media Streaming.


CLOUDFRONT
Exam Tips
A CLOUD GU
• Edge locations are not just READ only — you can write to them too. (ie put an object on to them.)
• Objects are cached for the life of the TTL (Time To Live.)
• You can invalidate cached objects, but you will be charged.

Snowball

SNOWBALL

What Is Snowball

A CLOUD GURU

Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of AWS. Using Snowball addresses common challenges with large-scale data transfers including high network costs, long transfer times, and security concerns. Transferring data with Snowball is simple, fast, secure, and can be as little as one-fifth the cost of high-speed Internet.

SNOWBALL

Snowball

A CLOUD GUR

Snowball comes in either a 50TB or 80TB size. Snowball uses multiple layers of security designed to protect your data including tamper-resistant enclosures, 256-bit encryption, and an industry-standard Trusted Platform Module (TPM) designed to ensure both security and full chain-of-custody of your data. Once the data transfer job has been processed and verified, AWS performs a software erasure of the Snowball appliance.

SNOWBALL

What Is Snowball Edge?

A CLOUD GURU

AWS Snowball Edge is a 100TB data transfer device with on-board storage and compute capabilities. You can use Snowball Edge to move large amounts of data into and out of AWS, as a temporary storage tier for large local datasets, or to support local workloads in remote or offline locations.

SNOWBALL

Snowball Edge

A CLOUD GURU

Snowball Edge connects to your existing applications and infrastructure using standard storage interfaces, streamlining the data transfer process and minimizing setup and integration. Snowball Edge can cluster together to form a local storage tier and process your data on-premises, helping ensure your applications continue to run even when they are not able to access the cloud.

SNOWBALL

What Is Snowmobile?

A CLOUD GURU

AWS Snowmobile is an Exabyte-scale data transfer service used to move extremely large amounts of data to AWS. You can transfer up to 100PB per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck. Snowmobile makes it easy to move massive volumes of data to the cloud, including video libraries, image repositories, or even a complete data center migration. Transferring data with Snowmobile is secure, fast and cost effective.