S3 Lifecycle Rules

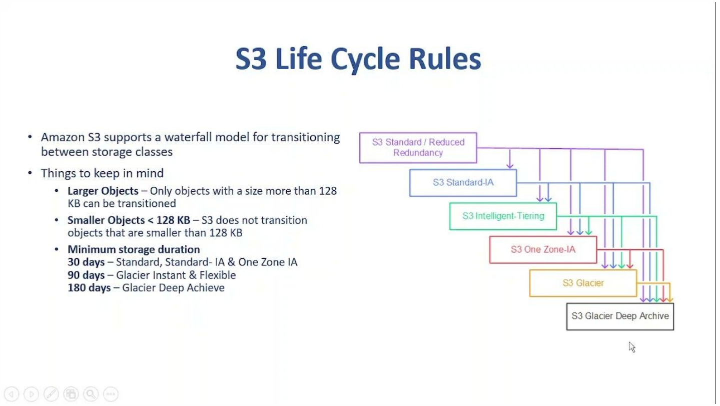
Amazon S3 Lifecycle Rules allow users to automatically transition objects between storage classes or delete them after a defined period, optimizing storage costs and data management. Rules can be applied to entire buckets, specific prefixes (folders), or objects with specific tags.

Key features:

- Transition Actions: Move objects to a lower-cost storage class (e.g., from S3 Standard to S3 IA, then to Glacier).
- Expiration Actions: Automatically delete objects or delete incomplete multipart uploads after a set time.
- Filter Options: Apply rules based on object prefix or tags.
- Granularity: Lifecycle rules support both current and previous versions of objects (if versioning is enabled).

Use cases:

- Archiving old log files to S3 Glacier after 30 days.
- Deleting temporary files after 7 days to reduce costs.
- Managing large datasets with tiered storage policies.



S3 Life Cycle Rules

- Amazon S3 supports a waterfall model for transitioning between storage classes
- Things to keep in mind
  - **Larger Objects** – Only objects with a size more than 128 KB can be transitioned
  - **Smaller Objects < 128 KB** – S3 does not transition objects that are smaller than 128 KB
  - **Minimum storage duration**
    30 days – Standard, Standard- IA & One Zone IA
    90 days – Glacier Instant & Flexible
    180 days – Glacier Deep Achieve

S3 Analytics

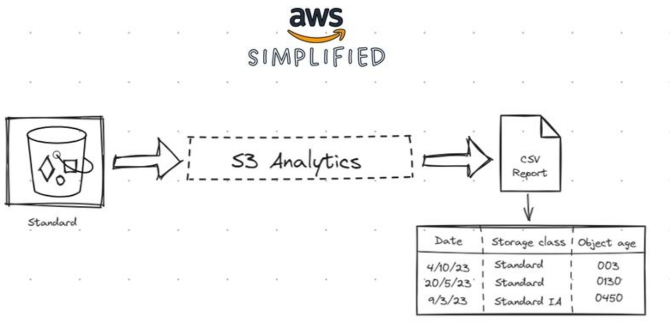
Amazon S3 Storage Class Analysis helps you determine when to transition data to a more cost-effective storage class based on access patterns. It analyzes object usage over time and provides reports that show which data is rarely accessed.

Key points:

- Purpose: Identifies infrequently accessed data that could be moved to lower-cost tiers (like S3 IA or S3 Glacier).
- Scope: Can be configured for the whole bucket, specific prefixes, or tagged objects.
- Reports: CSV files stored in a specified S3 bucket. They include daily analysis of object access patterns.
- Automation: Based on analytics, you can create lifecycle policies to automatically transition data.

Use cases:

- Understanding long-term access trends before applying lifecycle rules.
- Optimizing storage costs for large datasets by identifying cold data.



Date	Storage class	Object age
4/10/23	Standard	003
20/5/23	Standard	0130
9/3/23	Standard IA	0450

Requester Pays

Amazon S3 Requester Pays feature shifts the cost of data transfer and requests from the bucket owner to the user accessing the data. This is especially useful for publicly shared datasets where the bucket owner doesn't want to bear the download costs.

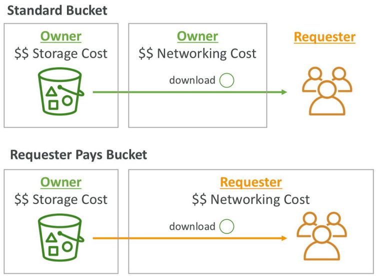
Key points:

- Enabled per bucket: The bucket owner must activate the feature.
- Requesters must include x-amz-request-payer: requester in their requests.
- Applies to: Data retrieval, GET, LIST, and other request types (requester is billed).

Use Case:

- Ideal for open data initiatives, research datasets, or public repositories.

Important: Requesters must have an AWS account and will be billed under their account for any actions that incur charges.

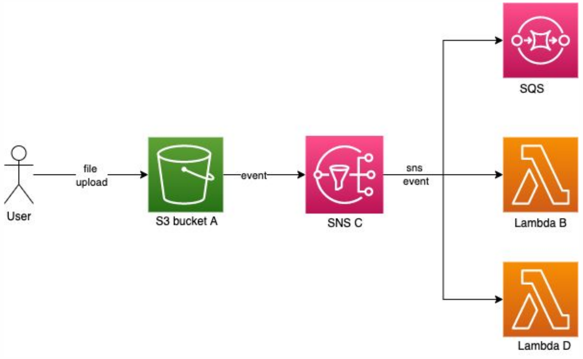


S3 Event Notifications

Amazon S3 Event Notifications enable automatic triggering of workflows when specific events occur in a bucket, such as object creation, deletion, or restoration. These notifications can be sent to Amazon SNS, SQS, or trigger an AWS Lambda function.

Supported Destinations:

- SNS topic (for fan-out messaging)
- SQS queue (for decoupling processes)
- Lambda function (for serverless processing)



Events:

- s3:ObjectCreated:* – when a new object is uploaded
- s3:ObjectRemoved:* – when an object is deleted
- s3:ObjectRestore:* – for restoration from Glacier

Configuration Scope: Can apply to a specific prefix (folder) or suffix (e.g., .jpg).

Use Case: Automate processing pipelines (e.g., image resizing, virus scanning, analytics) upon object uploads.

Limitations:

- Only one event notification configuration per bucket per destination.
- Not all object operations are supported (e.g., transitions via lifecycle rules don’t trigger events).

Baseline Performance

Amazon S3 is built to deliver high durability, availability, and baseline performance at scale — automatically and without manual tuning.

Automatic Scaling:

- S3 automatically scales to support high request rates without additional configuration.

Request Rate Per Prefix:

- Your application can achieve at least 3,500 PUT/COPY/POST/DELETE requests per second, and 5,500 GET/HEAD requests per second per prefix in a bucket.

Latency:

- Typical latency for GET and HEAD requests is around 100–200 milliseconds.

S3 Performance – Multipart Upload & Transfer Acceleration

Multipart Upload is a feature that enables you to upload large objects in parts, which can be done in parallel to improve performance.

When to Use: For files larger than 100 MB (recommended), required for objects >5 GB.

Benefits:
Parallel upload of parts → faster uploads.
Recovery from failure: if a part fails, only that part needs to be re-uploaded.
Efficient retries and network usage.

S3 Transfer Acceleration uses Amazon CloudFront's globally distributed edge locations to accelerate uploads to S3.

When to Use: For long-distance or high-latency uploads to S3.

Benefits:
Automatically routes data through the fastest path.
Works with a special endpoint: bucketname.s3-accelerate.amazonaws.com.
No need to change the bucket itself — only the endpoint used.

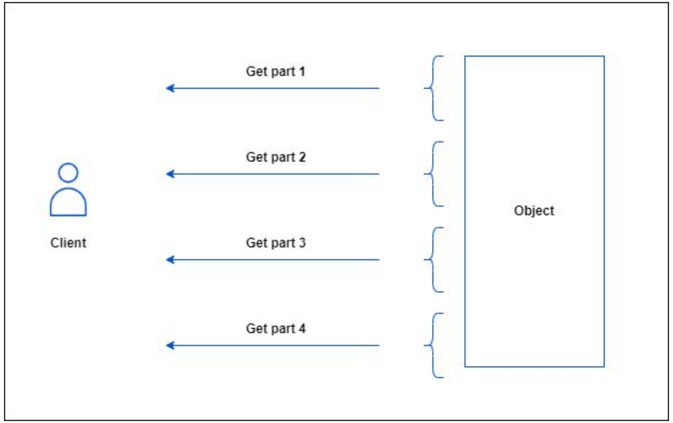
Byte-Range Fetches

Byte-Range Fetches in Amazon S3 allow applications to retrieve only a specific portion (range of bytes) of an object rather than downloading the entire file.

Partial Downloads: Ideal for large files like videos, logs, and archives.

Use Case: Media streaming, resume interrupted downloads, parallel processing of different object parts.

Performance Benefit: Enables concurrent byte-range requests to improve throughput and reduce latency.



S3 Batch Operations

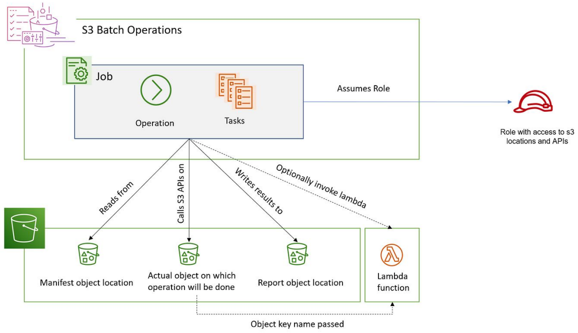
Amazon S3 Batch Operations allow you to perform large-scale operations on millions or billions of S3 objects with a single job. It's ideal for automating repetitive tasks across many files.

Supported Actions:

- Copying objects
- Setting object tags or ACLs
- Initiating object restores from Glacier
- Running AWS Lambda functions on each object
- Deleting objects

Job Definition:

- You define a manifest (list of object keys, usually stored in CSV format in S3).
- Select the operation to perform and configure permissions and logging.
- Optional integration with AWS Lambda for custom logic on each object.



Amazon S3 Storage Lens

Amazon S3 Storage Lens is an analytics tool that provides visibility into object storage usage and activity across an entire organization or account. It helps you monitor, optimize, and apply best practices to your S3 storage.

- Aggregated Metrics: Collects usage metrics such as object counts, storage size, number of requests, and more—at the account, bucket, prefix, or tag level.
- Dashboards: Comes with a default dashboard in the AWS Console to visualize trends, identify cost-saving opportunities, and detect anomalies.
- Scopes: Can analyze all buckets or a subset based on accounts, regions, or tags.
- Data Retention: Metrics are stored for 15 months for historical trend analysis.

Free vs. Paid (Advanced) Tier:

- Free Tier:
  - Includes 28 metrics (e.g., object counts, total storage, request counts).
  - Basic dashboard and filtering capabilities.

- Advanced (Paid) Tier:
  - Adds over 35 additional metrics (e.g., data retrievals, incomplete multipart uploads, lifecycle rule effectiveness).
  - Enables advanced filtering, prefix-level granularity, CloudWatch publishing, and exporting metrics to S3.

S3 Exam Questions

A university hosts public research data in an S3 bucket. Due to rising download costs from global access, the IT team wants to ensure that any user downloading data is responsible for paying the data transfer charges.

What should the team do to implement this requirement?

A. Enable CloudFront with signed URLs to enforce download quotas

B. Enable S3 Requester Pays on the bucket

C. Set up a Lambda@Edge function to charge requesters

D. Configure object-level encryption with cost-tracking tags

A video hosting platform uploads very large media files (100+ GB) to S3. Developers report intermittent failures and long upload times. The solution must improve reliability and performance while handling potential network interruptions.

Which action should the Solutions Architect take?

A. Use Transfer Acceleration and standard PUT requests

B. Enable S3 Object Lock to prevent partial uploads

C. Use S3 Multipart Upload with retry logic on failed parts

D. Increase the bucket-level request throughput quota via a support ticket

S3 Encryption

S3 supports encryption at rest and in transit.

Encryption at Rest:

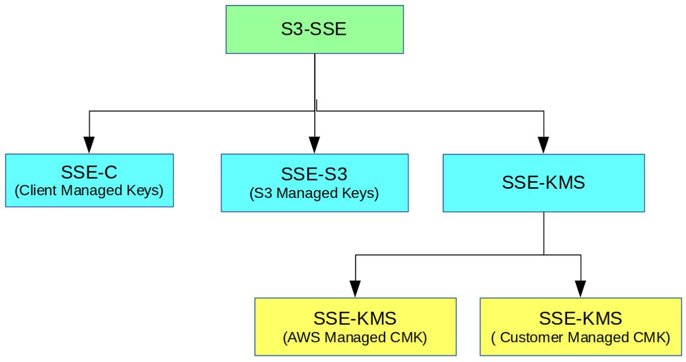
- SSE-S3 (AES-256): S3 encrypts each object with a unique key, and the key itself is encrypted with a master key.
- SSE-KMS: AWS Key Management Service (KMS) allows control over key rotation and provides detailed audit logs.
- SSE-C: Customer provides the encryption key, which must be included with each request.
- Client-side encryption: Data is encrypted before upload.

You can enforce encryption by:

- Using a bucket policy to deny unencrypted uploads.
- Enabling default encryption on the bucket.

Encryption in Transit:

- Achieved via SSL/TLS (HTTPS), encrypting data between clients and S3.



S3 CORS

CORS allows your S3 bucket to handle requests coming from a different origin (e.g., your frontend hosted on another domain).

Must configure a CORS rule on the S3 bucket.

Rules include:

- AllowedOrigins: Domains allowed to access the bucket.
- AllowedMethods: HTTP methods (GET, PUT, etc.) permitted.
- AllowedHeaders: Headers that can be sent with requests.
- ExposeHeaders: Headers exposed in the response.
- MaxAgeSeconds: Time the browser caches the preflight response.

This is required for frontend applications hosted on other domains to interact with S3 via JavaScript (e.g., uploading images or reading data).

MFA Delete

MFA Delete adds an extra layer of security by requiring Multi-Factor Authentication (MFA) to permanently delete a versioned object or change the versioning state of a bucket.

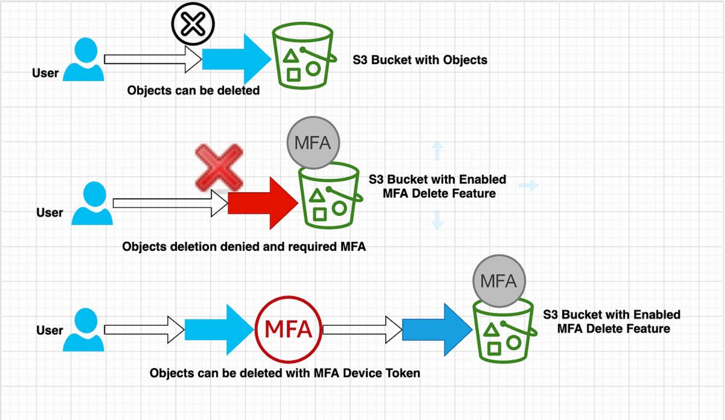
Important Notes:

- Only the root user can enable or disable MFA Delete.
- To use MFA Delete, Versioning must be enabled on the bucket
- MFA Delete is only available via the AWS CLI or API, not the Management Console.

When enabled, any action to permanently delete an object or suspend versioning requires:

- Your account credentials.
- A valid MFA code from a virtual or hardware MFA device.

It is used to prevent accidental or malicious deletions.


User
Objects can be deleted
S3 Bucket with Objects
MFA
User
S3 Bucket with Enabled MFA Delete Feature
Objects deletion denied and required MFA
MFA
User
MFA
S3 Bucket with Enabled MFA Delete Feature
Objects can be deleted with MFA Device Token

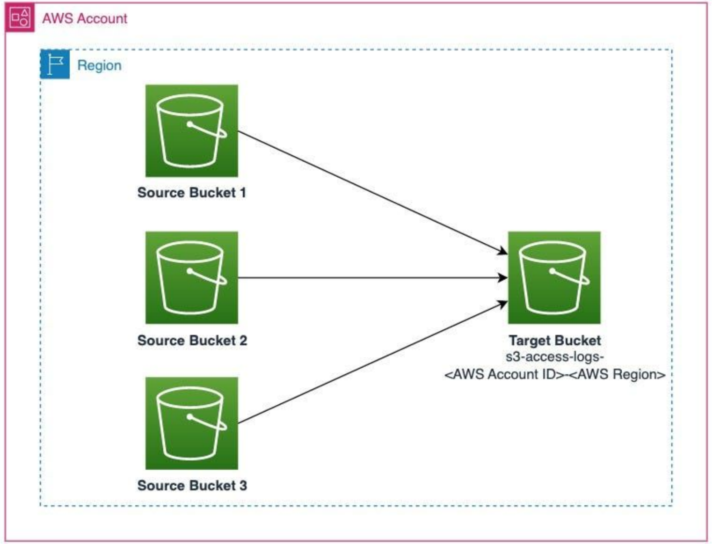
S3 Access Logs

Access logs in S3 provide detailed records for requests made to a bucket, including:

- Request type (GET, PUT, DELETE, etc.)
- Requester’s identity
- Time of request
- Response status
- Error codes (if any)

Use Cases:

- Security auditing (who accessed what and when)
- Billing analysis (which requests generated cost)
- Usage patterns (optimize data access and structure)



The target logging bucket must be in the same AWS region

You specify a target bucket where logs will be delivered.

Logs are stored as text files with request data in a predefined format.

Pre-Signed URL

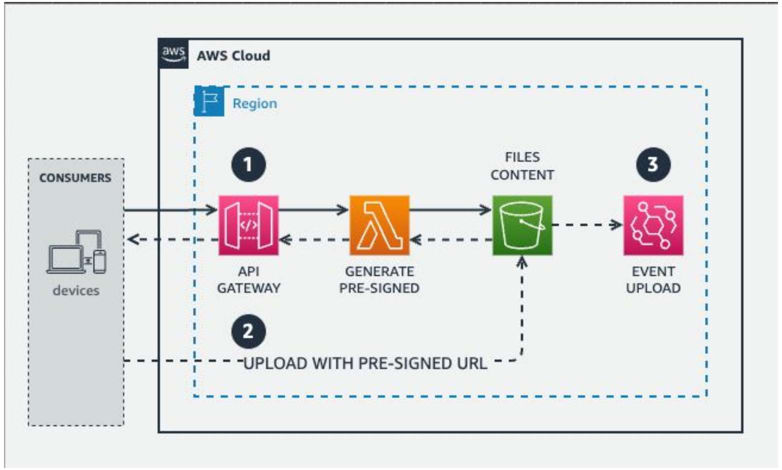
A pre-signed URL grants temporary access to a private S3 object without making it public.

How It Works:

1. A pre-signed URL is generated using the object's key, an expiration time, and the credentials of the user generating it.
2. Anyone with the URL can upload, download, or access the object (depending on the method used).
3. The permissions of the URL are based on the IAM permissions of the user who created it.

Use Cases:

- Letting users download files without making the bucket public.
- Allowing limited-time uploads (e.g., profile picture uploads).
- Secure, time-bound access to sensitive files.



Glacier Vault Lock

Glacier Vault Lock allows you to enforce compliance controls on Glacier vaults using a Vault Lock policy.

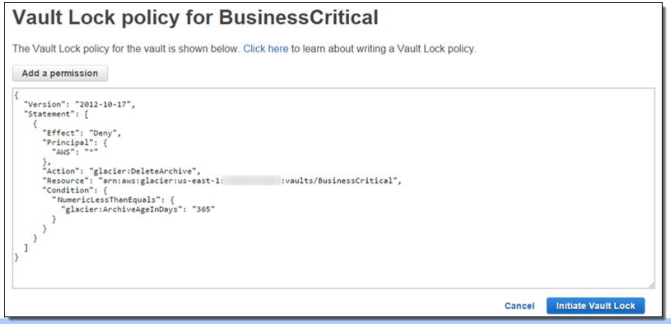
Key Features:

- Once locked, the policy becomes immutable (cannot be changed).
- Ideal for compliance and governance use cases (e.g., financial regulations).
- You initiate a Vault Lock by creating a lock policy and then locking it permanently after a testing phase.

Steps:

1. Initiate lock with a policy (in testing mode).
2. Test if the policy works as expected.
3. Finalize the lock — the policy becomes immutable.

Important: Vault Lock is different from S3 bucket policies — it is meant for write-once-read-many (WORM) scenarios.



Vault Lock policy for BusinessCritical

The Vault Lock policy for the vault is shown below. Click here to learn about writing a Vault Lock policy.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Deny",
      "Principal": {
        "AWS": "*"
      },
      "Action": "glacier:DeleteArchive",
      "Resource": "arn:aws:glacier:us-east-1:[redacted]:vaults/BusinessCritical",
      "Condition": {
        "NumericLessThanEquals": {
          "glacier:ArchiveAgeInDays": "365"
        }
      }
    }
  ]
}
```

S3 Object Lock

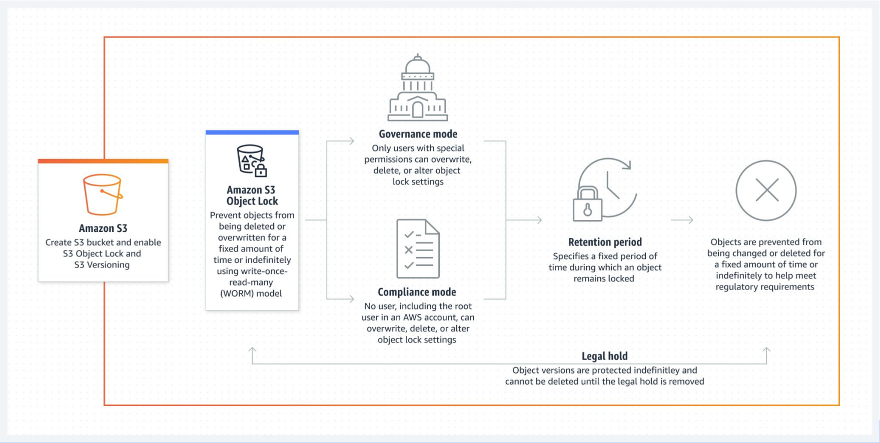
S3 Object Lock lets you store objects using a WORM (Write Once, Read Many) model. It prevents object deletion for a defined retention period or indefinitely.

Retention Modes:
Governance Mode: Users can't overwrite or delete the object unless they have special permissions (bypass governance).
Compliance Mode: No one can delete or modify the object, even root users.

Retention Settings:
Retention Period: Prevents deletion until the specified date.
Legal Hold: Like a legal pause — prevents deletion regardless of retention mode/duration.

Use Cases:
Regulatory compliance (e.g., financial, legal documents)
Protecting data from tampering or accidental deletion

Object Lock must be enabled when the bucket is created.


Amazon S3
Create S3 bucket and enable S3 Object Lock and S3 Versioning
Amazon S3 Object Lock
Prevent objects from being deleted or overwritten for a fixed amount of time or indefinitely using write-once-read-many (WORM) model
Governance mode
Only users with special permissions can overwrite, delete, or alter object lock settings
Compliance mode
No user, including the root user in an AWS account, can overwrite, delete, or alter object lock settings
Retention period
Specifies a fixed period of time during which an object remains locked
Objects are prevented from being changed or deleted for a fixed amount of time or indefinitely to help meet regulatory requirements
Legal hold
Object versions are protected indefinitley and cannot be deleted until the legal hold is removed

S3 Access Points

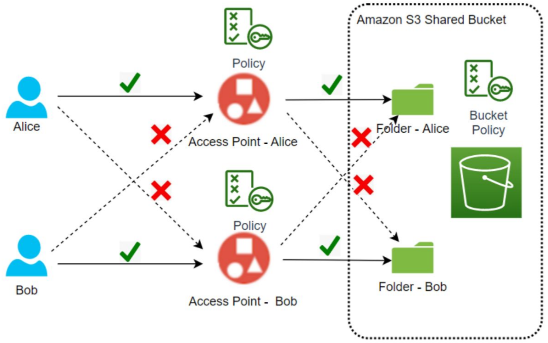
S3 Access Points simplify and manage access to shared datasets in S3.

Key Features:

- Each access point has a unique DNS name and its own access policy.
- Designed for multi-tenant environments or applications needing different access rules.
- You can restrict access to a specific VPC.
- Access points can have IAM or resource policies that define what actions are allowed.

Use Case:

- Instead of managing complex bucket policies, you can create multiple access points with simpler, purpose-specific policies.



Access Points – VPC Origin

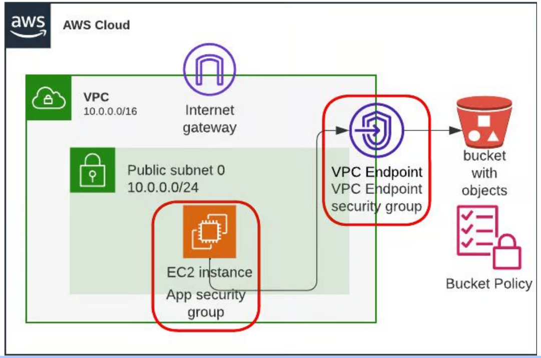
You can configure S3 Access Points to only accept requests that come from a specific Amazon VPC.

Key Points:

- Improves security by restricting data access to private network environments.
- Useful for organizations that require data access from within their VPC only, and not over the internet.
- This is configured as part of the access point settings.
- You can define VPC origin access control via resource policies attached to the access point.

Use Case:

- When you want applications within a VPC to interact with S3 without exposing the data publicly, Access Points with VPC restrictions provide tight control.



S3 Object Lambda

S3 Object Lambda allows you to transform data on-the-fly as it is retrieved from S3 by using a Lambda function.

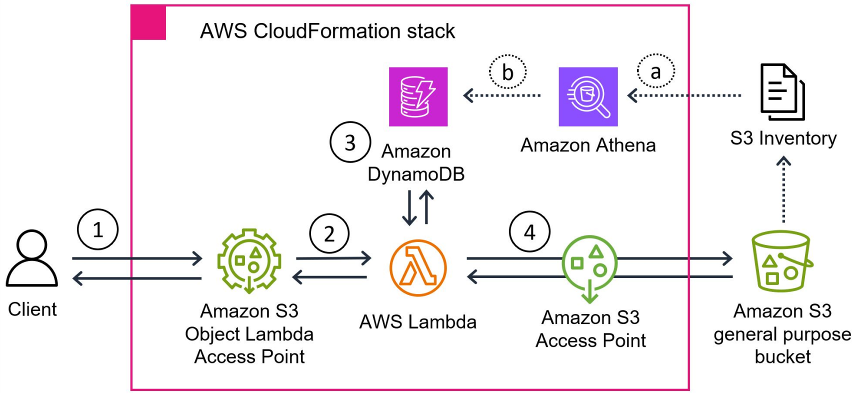
- You can modify, filter, or mask data before it's returned to the client.
- Lambda is invoked automatically when an object is requested through the Object Lambda Access Point.
- You keep the data in S3 unchanged while delivering customized responses to different users.

Examples:

- Redact sensitive information in a CSV file.
- Convert image formats dynamically.
- Apply real-time filtering to logs or data rows.

Architecture:

1. User sends GET request via Object Lambda Access Point.
2. Lambda function is triggered.
3. Lambda reads the original object and modifies it.
4. Modified result is returned to the user.



More S3 Exam Questions :)

You are tasked with securing a highly sensitive S3 bucket. The bucket must log all access, prevent accidental or malicious deletions of versioned objects, and allow only specific VPC access for internal tools.
Which combination of features meets these requirements?

A. Enable server-side encryption (SSE-S3), enable logging, apply a bucket policy for IP restriction

B. Enable S3 versioning, enable access logs, enable MFA Delete, use S3 Access Points restricted to the VPC

C. Use CloudTrail for object-level logging, apply Glacier Vault Lock, use signed URLs

D. Use S3 Transfer Acceleration, enable CORS rules, enable SSE-KMS encryption

Your web application allows users to upload profile pictures directly to Amazon S3. You want to ensure that only authenticated users can upload their own images, and that the upload link expires after 5 minutes.
What is the most secure and scalable way to implement this?

A. Make the bucket public and restrict uploads using bucket policies

B. Use S3 Object Lambda to transform the image before it is uploaded

C. Enable CORS on the bucket and let users upload via JavaScript directly to a public bucket

D. Generate a pre-signed PUT URL on the backend with a 5-minute expiration and return it to the frontend

Amazon CloudFront – Overview

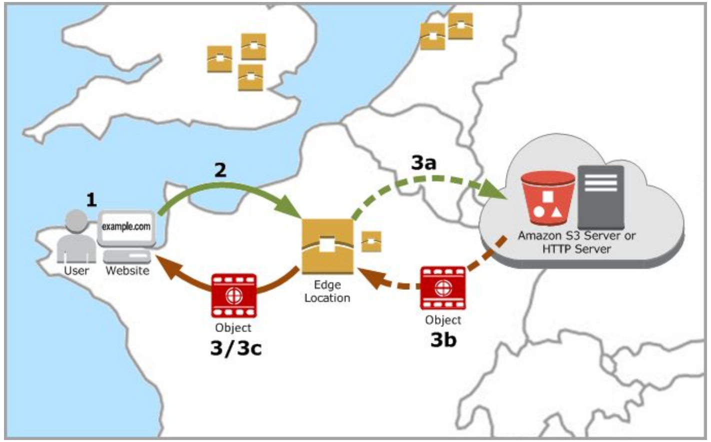
Amazon CloudFront is a Content Delivery Network (CDN) that securely delivers data, videos, applications, and APIs with low latency and high transfer speeds.

Edge Locations: Caches content at global locations closer to users.

- Integration with AWS: Works with S3, EC2, ALB, Lambda@Edge, and more.
- HTTPS Support: SSL/TLS encryption for secure delivery.
- DDoS Protection: Integrated with AWS Shield Standard.
- Caching: Reduces load on origin servers and improves performance.

Use Cases:

- Speed up websites and APIs.
- Deliver media content globally.
- Secure dynamic and static content.
- Enforce geolocation-based access controls.

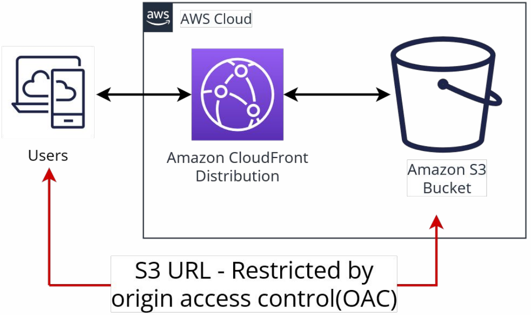


CloudFront Origins

A CloudFront origin is the source location from which CloudFront fetches content when it's not already cached at an edge location.

Supported Origin Types:

- Amazon S3 – for static websites and file delivery.
- Elastic Load Balancer (ALB/CLB) – for dynamic web applications.
- Amazon EC2 – directly connect to compute instances.
- Custom Origin – any HTTP/HTTPS server (e.g., on-premises or non-AWS).


AWS Cloud
Users
Amazon CloudFront Distribution
Amazon S3 Bucket
S3 URL - Restricted by origin access control(OAC)

Advanced Features:

- Origin Groups: Define primary and failover origins.
- Origin Access Control (OAC): Replaces legacy OAI to restrict direct S3 access using signed requests.

CloudFront Geo Restrictions

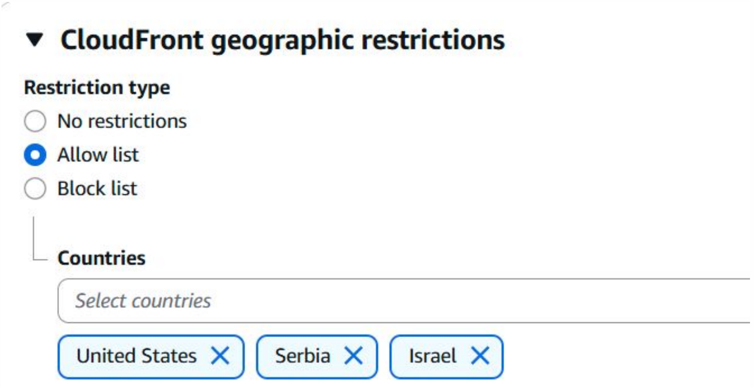
CloudFront Geo Restriction (also called geoblocking) allows you to control access to your content based on the geographic location of the viewer.

Two Main Modes:

1. Allowlist: Only specified countries can access the content.
2. Denylist: Block specific countries from accessing the content.

How It Works:

- CloudFront determines the viewer's country using the IP address.
- Geo restriction rules are applied at the CloudFront edge location.
- Commonly used for licensing compliance or content control.



▼ CloudFront geographic restrictions

Restriction type

○ No restrictions

● Allow list

○ Block list

Countries

Select countries

United States ✕ Serbia ✕ Israel ✕

CloudFront Price Classes

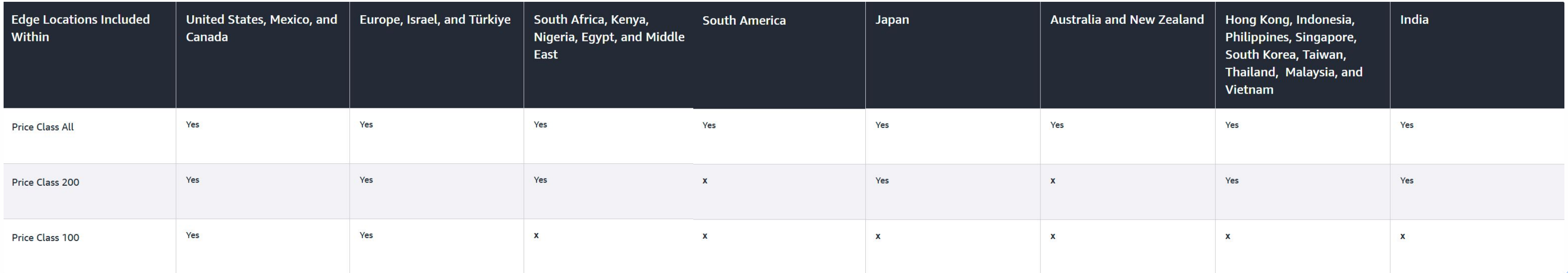
CloudFront Price Classes help control costs by limiting which edge locations serve your content.

Three Price Classes:

1. Price Class 100: Only the most cost-effective edge locations (mainly U.S., Canada, Europe).
2. Price Class 200: Includes additional edge locations in Asia and South America.
3. Price Class All (default): All available edge locations worldwide.

Key Points:

- Selecting a lower price class reduces cost but may increase latency.
- You can change price class anytime without redeploying your distribution.



Edge Locations Included Within	United States, Mexico, and Canada	Europe, Israel, and Türkiye	South Africa, Kenya, Nigeria, Egypt, and Middle East	South America	Japan	Australia and New Zealand	Hong Kong, Indonesia, Philippines, Singapore, South Korea, Taiwan, Thailand, Malaysia, and Vietnam	India
Price Class All	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price Class 200	Yes	Yes	Yes	x	Yes	x	Yes	Yes
Price Class 100	Yes	Yes	x	x	x	x	x	x

CloudFront Cache Invalidations

Cache Invalidation in CloudFront allows you to remove outdated content from edge locations before it expires naturally.

How It Works:

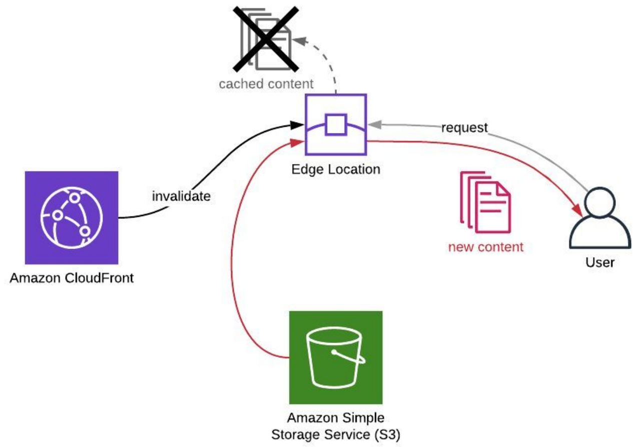
1. You submit an invalidation request specifying object paths (e.g., /index.html, /images/*).
2. Propagation usually takes a few minutes.
3. The next user request will fetch fresh content from the origin.

Cost Note:

- The first 1,000 invalidation paths per month are free.
- Additional requests are charged.

Use Cases:

- Force update of static files after a website deploy.
- Clear sensitive or outdated cached content.



AWS Global Accelerator

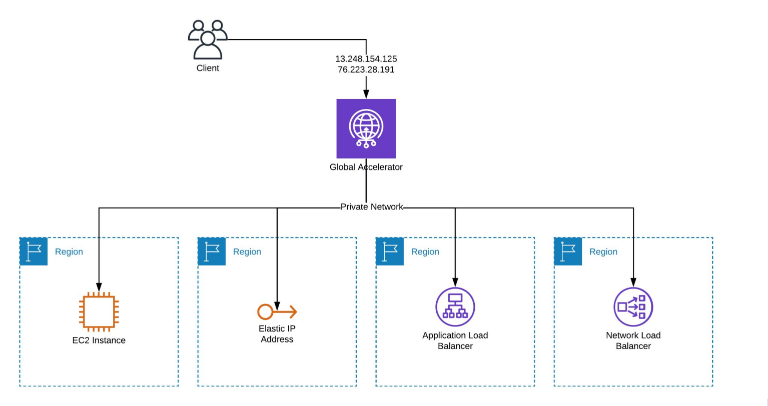
AWS Global Accelerator is a global networking service that improves the availability and performance of your applications by using the AWS global network.

Key Features:

- Static IP Addresses: Two global static IPs provided as fixed entry points.
- Improved Latency: Routes users through the optimal AWS edge location to your endpoint.
- Health Checks: Automatically routes traffic to healthy endpoints.
- Supported Endpoints: Application Load Balancer (ALB), Network Load Balancer (NLB), EC2, and Elastic IPs.

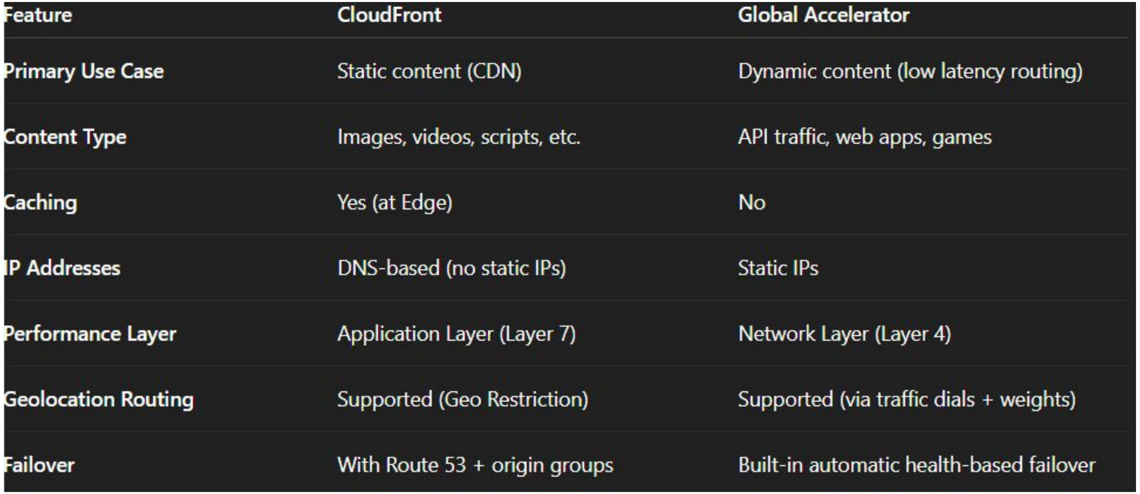
Use Cases:

- Global applications requiring consistent performance.
- Applications with users in multiple continents.
- Multi-region failover and traffic routing.



Global Accelerator vs CloudFront

Although both AWS Global Accelerator and Amazon CloudFront improve performance and availability for end users, they serve different use cases and operate at different layers.



Feature	CloudFront	Global Accelerator
Primary Use Case	Static content (CDN)	Dynamic content (low latency routing)
Content Type	Images, videos, scripts, etc.	API traffic, web apps, games
Caching	Yes (at Edge)	No
IP Addresses	DNS-based (no static IPs)	Static IPs
Performance Layer	Application Layer (Layer 7)	Network Layer (Layer 4)
Geolocation Routing	Supported (Geo Restriction)	Supported (via traffic dials + weights)
Failover	With Route 53 + origin groups	Built-in automatic health-based failover

In Short:

- Use CloudFront for fast delivery of static content via caching.
- Use Global Accelerator for performance and availability of dynamic content or APIs.

Cloud Front + Global Accelerator Exam Questions

A company is using Amazon S3 to host static assets for a global website. Users from Asia are reporting high latency when accessing these assets. The company wants to reduce latency and ensure secure content delivery using HTTPS, with minimal operational overhead.
What is the most appropriate solution?

A. Enable S3 Transfer Acceleration

B. Deploy the content on an EC2 instance in Asia

C. Use AWS Global Accelerator

D. Create a CloudFront distribution with the S3 bucket as the origin.

A company has deployed its application in multiple AWS Regions using Application Load Balancers. They want to improve performance for global users by routing them to the nearest healthy Region. Additionally, they want to use static IP addresses to simplify firewall configurations and ensure automatic failover if a Region becomes unavailable.
Which solution meets these requirements?

A. Use Route 53 with latency-based routing and health checks

B. Use CloudFront with regional origins and geo restriction

C. Use AWS Global Accelerator with the ALBs as endpoints

D. Deploy a Network Load Balancer in front of all ALBs with failover routing

AWS Snow Family

The AWS Snow Family includes physical devices designed to move large amounts of data in and out of AWS, especially when network transfer is slow, costly, or unavailable.

Snowcone

- Smallest device (8 TB of usable storage).
- Portable, rugged, and can run edge computing apps using AWS IoT Greengrass or EC2.
- Can be carried, shipped, or even used with a drone or backpack.

Snowball

- Available in two options:
    - Snowball Edge Storage Optimized: 80 TB of usable storage.
    - Snowball Edge Compute Optimized: 42 TB + GPU support, runs EC2 instances and Lambda functions.
    - Designed for edge computing, machine learning, video analytics, and storage transfer.

Snowmobile

- Massive-scale data transfer (up to 100 PB).
- A shipping container delivered by a truck – ideal for large data center migrations.

Snow Usage Process

1. Create a Job in the AWS Snow console (define data source and destination).
2. AWS prepares and ships the device to your location.
3. Connect the device to your local network and transfer data using AWS CLI or Snowball client.
4. Ship it back to AWS (prepaid label included).
5. AWS uploads your data into S3 (or desired destination) and securely erases the device.

Amazon FSx for Windows File Server

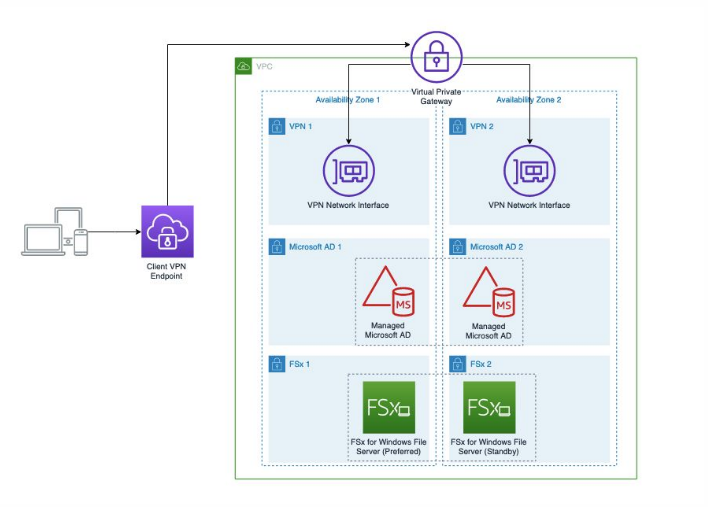
Amazon FSx for Windows File Server provides a fully managed Windows file system built on Windows Server, accessible via the SMB protocol.

Key Features:

- Native support for Windows ACLs, Active Directory, and DFS namespaces.
- Integrates with Microsoft applications like SQL Server, SharePoint, etc.
- Supports multi-AZ deployment for high availability.
- Scales up to tens of thousands of concurrent users.

Use Cases:

- Windows-based workloads requiring shared file storage.
- Migrating on-prem Windows file servers to AWS.
- Applications that require NTFS and SMB features.



Amazon FSx for Lustre

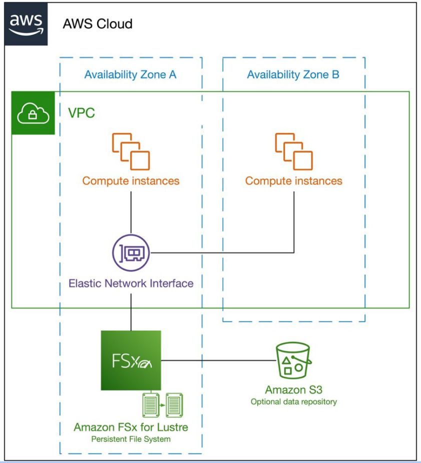
Amazon FSx for Lustre is a fully managed high-performance file system optimized for fast processing of large datasets, often used in HPC (High-Performance Computing), machine learning, and financial simulations.

Key Features:

- Built on Lustre, an open-source parallel file system used in supercomputers.
- Can scale to hundreds of GB/s throughput and millions of IOPS.
- Integrates natively with Amazon S3: you can link an FSx for Lustre file system to an S3 bucket.
- Supports bursting I/O for fast processing, especially useful in analytics pipelines.

Use Cases:

- Big data analytics (Spark, Hive, Presto).
- Machine learning model training.
- Video rendering and simulations.
- Genomics and scientific computing.



Amazon FSx for NetApp ONTAP

Amazon FSx for NetApp ONTAP is a fully managed shared storage service that provides the full capabilities of NetApp's ONTAP file system in AWS.

Key Features:

- Supports multi-protocol access: NFS, SMB, and iSCSI.
- Built-in data deduplication, compression, and thin provisioning.
- Snapshots and replication using NetApp SnapMirror.
- Can integrate with on-prem NetApp systems for hybrid storage architectures.
- Supports multi-AZ deployments for high availability.

Use Cases:

- Migrating existing NetApp workloads to AWS.
- Hybrid cloud storage with on-prem NetApp environments.
- Shared storage for enterprise applications (e.g., SAP, Oracle, VDI).
- Running legacy apps that need advanced data management features.

Amazon FSx for OpenZFS

Amazon FSx for OpenZFS delivers a fully managed file system based on the popular OpenZFS open-source technology, ideal for Linux-based workloads requiring advanced storage features.

Key Features:

- Built on ZFS, known for data integrity, snapshots, and replication.
- Supports NFSv3 and NFSv4 protocols.
- Includes compression, deduplication, and instant snapshots.
- Easily scales throughput and IOPS independently from storage.
- Designed for low-latency and high-performance workloads.

Use Cases:

- Linux applications requiring ZFS features.
- Migrating on-prem ZFS-based storage to AWS.
- Analytics, CI/CD pipelines, and container storage (e.g., Docker, Kubernetes volumes).

AWS Storage Gateway

AWS Storage Gateway is a hybrid cloud storage service that connects on-premises environments with AWS cloud storage. It allows local applications to use cloud-backed storage with low-latency access and standard storage protocols.

Gateway Types:

File Gateway

Provides access to objects in Amazon S3 as files using NFS or SMB protocols.

- Local caching for frequently accessed data
- Ideal for backups, analytics, and file sharing
- Data is stored in S3 as objects, with local file system access

Volume Gateway

Presents cloud-backed iSCSI block storage volumes to on-premises apps.

Two modes:

- Cached Volumes: frequently accessed data cached locally, rest in AWS
- Stored Volumes: full dataset stored locally, asynchronously backed up to AWS

Supports point-in-time snapshots stored as EBS snapshots

Tape Gateway

A virtual tape library interface for backing up to AWS using existing tape-based backup software.

- Integrates with popular backup tools (e.g., Veeam, NetBackup)
- Replaces physical tape infrastructure
- Virtual tapes stored in S3 or archived in Glacier

AWS Transfer Family

AWS Transfer Family is a fully managed service that enables you to transfer files directly into and out of Amazon S3 or Amazon EFS using traditional file transfer protocols.

Supported Protocols:

- SFTP (Secure File Transfer Protocol)
- FTPS (FTP over SSL)
- FTP (File Transfer Protocol)

Key Features:

- Fully integrated with Amazon S3 and EFS
- Supports user authentication via IAM, Active Directory, or custom identity providers using API Gateway & Lambda
- Serverless architecture – no need to manage infrastructure
- Offers logging, compliance, and security via CloudWatch, CloudTrail, and encryption options

Use Cases:

- Replacing legacy FTP servers with a secure, scalable alternative
- Partner or customer file uploads/downloads
- Secure, auditable file exchange pipelines

AWS DataSync

AWS DataSync is a data transfer service that automates moving data between on-premises storage and AWS, or between AWS services.

Key Features:

- Supports transfers to/from:
    - Amazon S3
    - Amazon EFS
    - Amazon FSx (Windows, Lustre, ONTAP, OpenZFS)

- Can transfer millions of files rapidly with built-in parallelism.
- Data is encrypted in transit and at rest.
- Incremental sync: only changed data is moved.
- Integrated with CloudWatch and CloudTrail for monitoring and auditing.

Use Cases:

- Migrating file systems to AWS.
- Periodic synchronization for hybrid storage.
- Large-scale backup or archival.
- Moving data between AWS Regions or services.

Storage Comparison

- S3 – Object storage for general-purpose data. (Keyword: Scalable, Durable)
- S3 Glacier – Long-term archival storage with retrieval delays. (Keyword: Archival)
- EBS volumes – Block-level storage for a single EC2 instance. (Keyword: Persistent Storage)
- Instance Storage – High-speed temporary block storage physically attached to EC2. (Keyword: Ephemeral, High IOPS)
- EFS – Shared file system for Linux instances with POSIX compliance. (Keyword: Scalable File Storage)
- FSx for Windows – Managed SMB file storage for Windows-based applications. (Keyword: Windows-native, AD Integration)
- FSx for Lustre – High-performance parallel file system for HPC and ML. (Keyword: High Throughput)
- FSx for NetApp ONTAP – Advanced enterprise file system with SnapMirror and iSCSI. (Keyword: Hybrid Cloud, iSCSI)
- FSx for OpenZFS – Managed ZFS file system with snapshots and compression. (Keyword: ZFS)
- Storage Gateway – Hybrid storage bridging on-prem to AWS via File, Volume, or Tape Gateway. (Keyword: Hybrid, Backup Integration)
- Transfer Family – Fully managed SFTP/FTP/FTPS interface to S3 or EFS. (Keyword: SFTP, Legacy Integration)
- DataSync – Accelerated, automated data transfer from on-prem or AWS-to-AWS. (Keyword: Migration, Sync)
- Snowcone / Snowball / Snowmobile – Physical devices for bulk offline data transfers. (Keyword: Petabyte Scale Migration)

Storage Extras Exam Questions

A company needs to transfer 800 TB of video surveillance data from a remote facility with very limited internet connectivity to AWS for long-term storage. Which AWS service provides the most efficient and cost-effective solution?

A. AWS DataSync

B. Amazon S3 Transfer Acceleration

C. AWS Snowball

D. AWS Storage Gateway – File Gateway

A company runs a Linux-based big data analytics application that requires access to large datasets stored in Amazon S3. They need a high-performance file system that supports sub-millisecond latencies and parallel processing. Which service best meets this requirement?

A. Amazon EFS

B. Amazon FSx for Lustre

C. Amazon FSx for NetApp ONTAP

D. Amazon S3 Glacier Deep Archive