**IAM & S3**

**IAM**

1. IAM is universal – applies to all the regions
2. The “root” account is the sign-up account, that has all the accesses, including billing and payment
3. New user has no permission
4. Access key / Secrete access key for program access (Regenerate if lost)
5. User-name / password for console access – Account alias for easy remember that account ID
6. Recommend Multi-factor authentication on “root” account
7. Create and customize password and password rotation policy

**Billing Alarm**

1. Cloudwatch
2. Billing alarm -> SNS topic -> Email

**S3 101**

1. S3 is object based – Files
2. Files 0 – 5 TB (Remember 0)
3. S3 has unlimited storage
4. Buckets -> Top lever folder -> Names are globally unique -> Access URL like <https://bucket-name.s3.amazonaws.com/> (Northern Virginia US-EAST-1). Other regions will have the region name in the URL
5. Not possible to install OS or Database
6. Get HTTP 200 code when successfully upload a file
7. You can turn on MFA Delete (Protect the objects)
8. S3 Object (5 items) => Key + Value + Version ID + Metadase (Data about the data storing) + Sub-resources (Access Control Lists (Permissions, bucket lever / object level), Torrent)
9. Read after Write consistency on PUT new Objects ⬄ Eventual consistency for over-write PUTS and Delete (It can take some time to propagate)
10. Storage Classes
    1. S3 Standard => 11/9 and 99.99%
    2. S3 – IA (Infrequent Access) => Retrial fee
    3. S3 – One Zone -IA => Retrial fee and lower availability
    4. S3 – Intelligent Tiering
    5. S3 Glacier => Retrial time minutes to hours
    6. S3 Glacier Deep Archive => Lowest cost, retrieval time 12 hours

**Create a Bucket**

1. Bucket names are global, but created in a region
2. Bucket can be replicated among regions => Cross region replication
3. Change storage classes (6 of them) and encryption on the fly
4. Transfer acceleration => Edge location and AWS backbone network for speed upload
5. Restrict Bucket Access
   1. Bucket polices – Applies across the whole bucket
   2. Object Policies – Applies to individual files
   3. IAM Policies to users and groups (Roles to services) – Applies to Users and Groups and Services

**S3 Access Control (Song)**

Bucket Level

* Block public access (bucket settings) – Public access means URL through HTTP
* Bucket Policy
* Access control list (Grantee – Objects – Bucket ACL) (Like a folder, write means PUT/DELETE files) => Grant only

Object Level

* Access control list (Grantee – Object – Object ACL) => Grant only

**S3 Pricing Tiers**

**S3 Security and Encryption**

1. S3 can be configured to create **access logs** that logs all requests made to a bucket. This can send to another bucket or even another account

Graphical user interface, text, application, email

Description automatically generated

1. Encryption in transit
2. Encryption at rest (Server side / Client side)
   1. SSE-S3 => AWS Manage both master key and data key
   2. SSE-KMS => AWS Manage data key and you manage master key
   3. SSE - C => You manage both master and data key
   4. <https://acloud.guru/forums/aws-certified-solutions-architect-associate/discussion/-KWgkIT3lpX-oekUXCRU/sse-s3-and-sse-kms>

**S3 – Versioning**

1. Stores all versions of an object, including all writes and even if you delete an object => Delete marker
2. Great backup tool
3. Once Versioning is enabled, it cannot be disabled, only suspended
4. Integrates with Lifecycle rules
5. Versioning’s MFA Delete capability provides an additional layer of security. When Versioning disabled, no MFA Delete available

**S3 Lifecycle Management**

1. Automates moving objects between **storage tiers** (classes)
2. Can be used together with Versioning, although Versioning is not required
3. Can be applied to current version and previous versions
4. Lifecycle rules can be set on a pattern to match sub-folders

**S3 Object Lock & Glacier Lock**

1. S3 Object Lock => WORM Model. Not overwritten or deleted for the retention period
2. Bucket level or object level
3. It is possible to put a legal hold on an object (version). It prevents object change until hold removed (Need special permission to place / remove legal hold)
4. **Governance mode** or **compliance mode**
   1. Governance => Giving exceptions to some users
   2. Compliance => No exception even to the root user
5. Glacier Vault Lock => Once locked, it can no longer changed

**S3 Performance**

1. Prefixes – Subfolder names => Improve performance by adding prefixes. Each prefix 3500 PUT/COPY/POST/DELETE and 5500 GET/HEAD per second. 2 prefixes => 11000 GET/second
2. SSE-KMS sets hard limit on the KMS quota. Region specific => 5500, 10000, 30000 per second depending on region. No quota increase possible currently
3. Multi-part upload => Must use for files > **5 GB**
4. Use S3 byte-range fetches to increase performance when downloading file from S3

**S3 Select & Glacier Select**

1. S3 select is used to retrieve subset of data from an object using SQL expressions
2. Get data by rows or columns using SQL expressions
3. Save money on data transfer and increase speed up to 400 times
4. Glacier select is similar to S3 select, but on the Glacier storage tier

AWS Organizations & Consolidate Billing

1. Always enable MFA / STRONG password on the root account
2. Paying account is used for billing purpose only. Do not deploy resources into the paying account
3. Enable/Disable AWS Service using Service Control Policies (**SCP**) either on OU (Organization Unit) or on individual accounts

**S3 Bucket Cross Account Access**

**Bucket policy + IAM** => Applies to whole bucket and program access only

**Access Control List + IAM** => Applies to the individual object and program access only

**Cross Account IAM Role** =>Both programmable and console access

* Role ARN (Can be used by the other account to assign permissions to assume this role) => arn:aws:iam::660079349745:role/A-Cross-Account-S3
* Role URL (Easy to assume at the other account) => <https://signin.aws.amazon.com/switchrole?roleName=A-Cross-Account-S3&account=huge-head-li-2021>
* Trusted relation => A role has at least 2 policies. One of them is the ("sts:AssumeRole" that defines the principle allowed to assume the role). Other policies defines the permissions.
* An admin in the other account can then assume the role. But we can further limit who can assume the role (Non-admin and only allow the user to assume the Role by adding the ARN to the resource)

**Example Policies**

Text

Description automatically generated

**Cross-Region Replication**

1. **Versioning** must be enabled for both source and destination
2. Files in already in the bucket is not automatically replicated, only the newly added ones
3. All subsequent updated files will be replicated
4. Delete markers can be replicated (Enable / Disable) - <https://aws.amazon.com/about-aws/whats-new/2020/11/amazon-s3-replication-adds-support-for-replicating-delete-markers/>
5. **Deleting individual versions will not be replicated**
6. Replicated object is not automatically public when the source object is made public

**S3 Transfer Acceleration**

1. Upload files to Edge locations to take advantages of AWS back-bone network for speed
2. There is a tool to help to compare the speed to use Transfer acceleration from the current location - <https://s3-accelerate-speedtest.s3-accelerate.amazonaws.com/en/accelerate-speed-comparsion.html>

**AWS DataSync**

1. Used to move large amount of data from on-premises to AWS
2. Used with NFS (Network file system) and SMB (Server Message Block – a common internet file system) compatible file systems
3. Replication can be done hourly/daily/weekly
4. Install DataSync Agent to start the replication
5. Can be used to replicate EFS to EFS (Within AWS environment by installing DataSync Agent on an EC2)

**Cloudfront**

1. Cloudfront is a CDN (May be at multiple edge locations for different clients at different locations)
2. Edge location – The location where the content is cached for **TTL** (in seconds) (Cache can be disabled for dynamic content, but only take advantage of the Amazon back-bone network for speed). This is different from AWS Region/AZ. Edge locations is read / write => (Write for transfer acceleration for example)
3. Origin – The origin of the files, S3 Bucket. EC2, Elastic Load Balancer, or Route 53
4. Distribution – This is the name for the CDN which may consist a collection of Edge locations (I think a distribution is an instance of Cloudfront CDN)
5. Distribution Types – **Web Distribution / RTMP** (**REAL TIME MESSAGING PROTOCAL** Used for Adobe Media Streaming)
6. Cached Objects at Cloudfront can be Invalidated, but you will be **charged**

**Cloudfront – Lab**

1. Cloudfront is under Networking
2. You can specify a default root object - <https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/DefaultRootObject.html>

**Cloudfront Signed URLs and Signed Cookies**

1. Use Signed URLs/cookies when you want to secure content so that only the people you authorize are able to access it (Attach Policy URL expiration / IP ranges / Trusted signers – AWS accounts that can created Signed URLs) – **OAI** - <https://docs.aws.amazon.com/AmazonCloudFront/latest/DeveloperGuide/private-content-signed-urls.html>
2. OAI - An **Origin Access Identity (OAI)** is used for sharing private content via CloudFront – It is **NOT Limited to S3 Origin**
3. A **signed URL** is for individual file
4. A **signed cookie** is for multiple files. 1 cookie used on multiple files
5. If you do not want the user to access the S3 or the origin is EC2 (for instance), use Cloudfront signed URL. If you want the user to **access S3 directly**, you can use S3 presigned URL – (Use IAM User credential) - <https://docs.aws.amazon.com/AmazonS3/latest/userguide/ShareObjectPreSignedURL.html>
6. If the origin is **EC2** (**NOT S3**), then CloudFront signed URL is the only option.

**Snowball**

1. Snowball is Petabyte data transport solution
2. Snowball Edge is a mini-aws. It has storage and compute capability
3. Snowmobile is a truck – up to 100 PB
4. Snowball – Import to S3 / Export from S3

**Storage Gateway**

1. File Gateway -> For flat files stored directly on S3
2. Volume Gateway
   1. **Stored Volumes** – Entire dataset it stored locally and asynchronously backed up to S3
   2. **Cached Volumes** – Entire dataset is stored on S3 and most frequently accessed data is cached on site
3. Gateway Type Library – For types
4. DataSync vs Storage Gatewat - <https://acloud.guru/forums/aws-csa-2019/discussion/-MB1MIgdt9ZxnKYajtub/whats_is_the_difference_betwee>

**Athena vs Macie**

1. Athena
   1. Athena is an interactive query service
   2. It allows you to query data located in S3 using standard SQL
   3. Serverless
   4. Commonly used to analyze log data stored in S3
2. Macie
   1. Macie is a security service.
   2. Macie uses **AI** to analyze data in S3 to identify PsII
   3. Can also be used to analyze CloudTrail logs for suspicious API activity
   4. Includes Dashboard, reports and Alerting
   5. Great for **PCI-DSS** compliance and preventing ID theft
3. Athena vs. S3 Select - <https://stackoverflow.com/questions/49102577/what-is-difference-between-aws-s3-select-and-aws-athena> (Athena is a fully managed analytical service / S3 select allows to retrieve subset of data)

**S3 Static Web Hosting**

Graphical user interface, text, application, email

Description automatically generated

1. A Bucket can be enabled for Website Hosting
2. A index.html and error.html can be set on the hosting bucket
3. Need to disable “Block public access” and add a bucket policy to allow the files accessible by public URLs - <https://stackoverflow.com/questions/27912462/how-do-i-make-all-files-in-an-s3-bucket-public>

Text

Description automatically generated

A

A