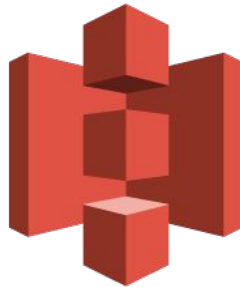# Amazon
Simple Storage Service (S3)

April 2018

What is Amazon S3 ?

- Data in the cloud

- Secure, durable, highly-scalable object storage

- Bottomless storage

- Store & retrieve any amount of data, at any time, from anywhere on the web

- Designed for 99.999999999% durability

- Designed for 99.99% availability

Why to use S3 ?

- Static website hosting

- Store & serve content of site

- Versioning

- Low cost data archiving

- Backup and Archiving snapshots

- Control object lifecycle

- Content Delivery Network (CDN)

- Logs of AWS services

Amazon S3 – Key Concepts

- Bottomless Storage

- **Bucket**
  - Container for objects stored in S3
  - limit of 100 buckets (soft limit)

- **Object**
  - Metadata
  - File
  - No limit on number of objects

- **Keys**
  - Unique identifier for an object
  - Combination of bucket, object and version id

- **Versioning**

Amazon S3 – Key Concepts

- **Cross-Region Replication**
  - Replication of object across Region
  - Minimize latency

- **Event Notifications**
  - Notifications in response to actions taken on objects
  - Notification through either Amazon SNS or Amazon SQS

- **Low-cost Reduced Redundancy Storage**
  - Low cost storage
  - For noncritical, reproducible data
  - Provide 99.99% durability and 99.99% availability

bucket

bucket with objects

object

Limits and Restrictions

- **Bucket Name Restrictions**
  - must be globally unique
  - should comply with DNS naming conventions
  - minimum of 3 or maximum of 63 characters long
  - must be small case (can only contain numbers , hyphen , dots)
  - must start with a letter or number i.e. cannot start with . (dot) or hyphen
  - can not be of the IP format 1.2.3.4
  - periods and hyphens cannot follow each other (e.g. cloudmanthan-.com)

  Some exceptions exists for `us-east-1` region
  - *The rules for bucket names in the US East (N. Virginia) region allow bucket names to be as long as 255 characters,*
  - *bucket names can contain any combination of uppercase letters, lowercase letters, numbers, periods (.), hyphens (-), and underscores (_).*

Amazon S3 – Hosting a Website

- Static Web Site

- Client Side Script

- Configure Amazon S3 bucket for website hosting

- Upload your content on web site

  - Index Document (index.html)
  - Error Document
  - Redirection (example.com and www.example.com)
  - Advanced conditional redirects

Amazon S3 – Hosting a Website

com.training.3dplm.s3-website-us-east-1.amazonaws.com

- For website to work objects must be publicly readable

- Grant everyone s3:GetObject permission

```
{
  "Version":"2012-10-17",
  "Statement":[{
      "Sid":"PublicReadGetObject",
      "Effect":"Allow",
       "Principal": "*",
    "Action":["s3:GetObject"],
    "Resource":["arn:aws:s3:::example-bucket/*"
    ]
   }
  ]
}
```

Managing Access

- All Amazon S3 resources are PRIVATE by default
  - Owner can access it
  - Owner can grant access to others by policy

- Policies
  - JSON based policies
    - Bucket Policy
    - User Policies

  - Resource based policies
    - Access Control List (ACLS)

Request

Permissions

Bucket with Objects

Managing Access with ACLS

- Manage access to buckets and objects
- Bucket permissions specify who is allowed access to the objects in a bucket and what permissions you have granted them. For example, one person might have only read permission while another might have read and write permissions.
- Grant basic read/write permissions to other AWS Accounts
- Limitations
- Grantee - to whom the permissions are assigned
  - Can grant to permissions to other AWS Accounts
  - Cannot grant permissions to IAM users in your account
  - No conditional permissions
  - No explicit 'Deny' permission
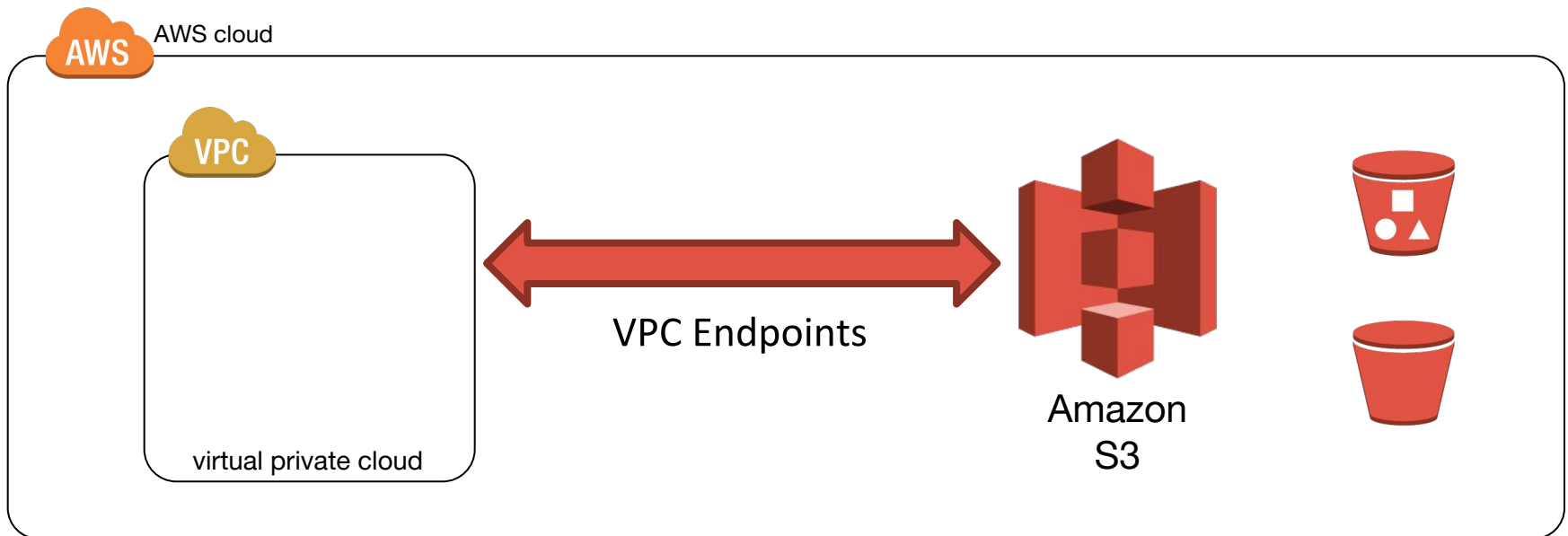
Bucket Policy

- Grant Read Only permission to anonymous user

```
{
  "Version":"2012-10-17",
  "Statement":[{
      "Sid":"AddPerm",
      "Effect":"Allow",
        "Principal": "*",
    "Action":["s3:GetObject"],
    "Resource":["arn:aws:s3:::examplebucket/*"
      ]
    }
  ]
}
```

VPC End point

- **VPC Endpoints**
  - Secure connection to S3 within VPC
  - Easy to configure
  - Highly reliable
  - Does not require a gateway or NAT instances

VPC endpoint - IP Ranges

```
C:\Users\Administrator>aws ec2 describe-prefix-lists
You must specify a region. You can also configure your region by running "aws co
nfigure".

C:\Users\Administrator>aws  configure
AWS Access Key ID [******************T7VQ]:
AWS Secret Access Key [******************OclB]:
Default region name [None]: us-east-1
Default output format [None]:

C:\Users\Administrator>aws ec2 describe-prefix-lists
{
    "PrefixLists": [
        {
            "PrefixListName": "com.amazonaws.us-east-1.s3",
            "Cidrs": [
                "54.231.0.0/17"
            ],
            "PrefixListId": "pl-63a5400a"
        }
    ]
```

S3 Features

S3 Transfer Acceleration

- Bucket level feature
- **When should you consider transfer acceleration ?**
  - Customers across globe are putting data to single bucket
  - GB and TBs of data to be uploaded globally across continents
  - You have not utilized enough bandwidth at your infrastructure
- Bucket name must comply with DNS naming convention and should not contain **'.' (dot)**
- In general, the farther away you are from an Amazon S3 region, the higher the speed improvement you can expect from using Amazon S3 Transfer Acceleration. If you see similar speed results with and without the acceleration, your upload bandwidth or a system constraint might be limiting your speed
- There is an additional charge for transfer acceleration /GB

S3 Transfer Acceleration

- Once enabled S3 provides different endpoints

    - *bucketname*.s3-accelerate.amazonaws.com

    - *bucketname*.s3-accelerate.dualstack.amazonaws.com (for
      IPv6)

Virtual Hosting of Buckets

- S3 buckets can be referred in 2 ways
- Virtual Hosting based
  - http://*bucketname*.s3.amazonaws.com
  - Can be mapped to DNS with alias to S3 bucket name
  - http://my.bucketname.com/
  - Provides the ability to publish to the "root directory" of your bucket's virtual server. This ability can be important because many existing applications search for files in this standard location. For example, favicon.ico, robots.txt, crossdomain.xml are all expected to be found at the root.
  - http://johnsmith.eu.s3-eu-west-1.amazonaws.com/homepage.html
- Path based
  - http://**s3-eu-west-1**.amazonaws.com/**buckentname**/puppy.jpg

Lifecycle Rules

Lifecycle Rules   ✕

Step 1: Choose Rule Target

Step 2: Configure Rule

Step 3: Review and Name

Lifecycle rules will help you manage your storage costs by controlling the lifecycle of your objects. Create Lifecycle rules to automatically transition your objects to the Standard-Infrequent Access Storage Class, archive them to the Glacier Storage Class, and remove them after a specified time period.

Choose different options below to see what works best for your use case. No rule will take effect until you activate them at the end of this wizard.

Action on Objects

☑ **Transition to the Standard-Infrequent Access Storage Class**   `30`   Days after the object's creation date
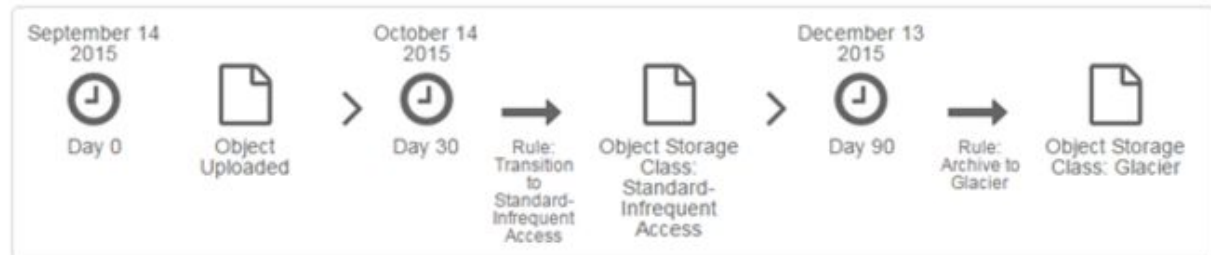
   Standard-Infrequent Access has a 30-day minimum retention period and a 128KB minimum object size. Lifecycle policy will not transition objects that are less than 128KB. Refer here to learn more about Standard-Infrequent Access.

☑ **Archive to the Glacier Storage Class**   `90`   Days after the object's creation date

   This rule could reduce your storage costs. Refer here to learn more about Glacier pricing. Note that objects archived to the Glacier Storage Class are not immediately accessible .

☐ **Permanently Delete**   `_____`   Days after the object's creation date

EXAMPLE:

September 14 2015			October 14 2015				December 13 2015		
🕐 Day 0	📄 Object Uploaded	>	🕐 Day 30	➡ Rule: Transition to Standard-Infrequent Access	📄 Object Storage Class: Standard-Infrequent Access	>	🕐 Day 90	➡ Rule: Archive to Glacier	📄 Object Storage Class: Glacier

Cancel   < Set Target   Review >

Pricing

- Storage Pricing
  - Standard Storage - $0.0300 per GB
  - Reduce Redundant Storage - $0.0240 per GB
  - Glacier Storage - $0.0100 per GB
- Request Pricing
  - PUT, COPY, POST, or LIST Requests - $0.005 per 1K request
  - Glacier Archive and Restore Requests - $0.05 per 1K request
  - GET and all other Requests - $0.004 per 10K request
  - Delete Requests - Free
  - Glacier Data Restores – Free
- Data Transfer Pricing
  - Data In – Free
  - Data Out – Max $0.090 per GB

Best Practices

- Choosing Bucket Location
  - Region Selection

- Utilizing Compression
  - http://www.http-compression.com/

- RRS non-critical data storage

- S3 with CloudFront CDN service
  - Low latency
  - Secure data access using signed URL

End of Module