S3 (simple storage service):

---------------------------

Q)what is S3?

s3 stands for simple storage service, s3 provides to the developers and it teams with secure, durable, high-scalable object storage.

amazon s3 is easy to use with simple web service interface to store and retrive any amount of data from anywhere on the web.

->s3 is a safe place to store your files. founder of dropbox help to amazon in a webpage interactive mode.

->it is object based stored. objects are photos, videos, all text based files(pdf,xl,txt, etc..).

->the data is spread across multiple devices.

s3-Basics:

---------

->s3 is a object based storage service i.e allows you to upload files.

->files can be from 0bytes to 5TB.

->there is unlimited storage.

->files are stored in buckets.

->s3 is universal name space, that is names must be unique globally.

->if you create a bucket it has a dns name

eg:https://s3-eu-east-2.amazonaws.com/(bucket name)

->when you upload a file to s3 you will receive a HTTP 200 code if the upload was successful.

s3 is simple key, value store

->s3 is object based, object consist of the follwoing

i)key (this is simply name of the object)

ii)value(this is simply the data what we upload)

iii)version ID (to maintain versions of the data)

iv)metadata (data about data what you are storing)

v)Access control list (who can access the list)

->built for 99.99% availability for s3 platform.

->amazon guarantee the 99.9% data availability.

->amazon guarantee the 99.999999999% durability for s3 information.

Data consistency model for s3:

-----------------------------

-> read after write consitency for puts of new objects

->eventual consitency for overwrite puts and deletes (can take some time to propogate) (update)

Q) what we learn in s3?

->Tired storage classes

->versioning

->Life cycle management

->s3 tranfer acceleration

->cross region replication

->security & encryption

->static website

s3- storage tires/classes:

-------------------------

->s3 - Standard 99.99% availability , 99.999999999% durability stored redundanty across multiple devices in multiple facilities and is designed to sustain the loss of 2 facilities concurrently.

-> S3 - IA(infrequently access) for data that is accessed less frequently, but requires rapid access when needed. lower fee than s3 standard, but you are charged a retrived fee.

->RRS (reduced redundancy storage) - design to provide 99.99% durability and 99.99% availiability of objects over a given year (very cheap)

eg: image in s3, thumbnail in rrs

->glacier - very cheap to store the data, but used for archival only. it takes 3-5 hours to restore from glacier

(for retriving data in minutes,hours,days, you choose your service depends upon cost)

Q)what is glacier?

glacier is an extremely low-cost storage service for data archival.amazon glacier stores data for as little as $0.01 per 1GB per 1 month, and is optimized for data that is infrequently accessed and for which retrival times of 3 to 5 hours suitable.

S3-charges:

----------

amazon charge for

->storage

->requests

->storage management pricing (associative dveloper)

->data tranfer pricing (move one region to another)

->tranfer acceleration

Q)how to create a bucket in s3?

go to aws management console->storage->s3

->s3 is in global region

->bucket name is unique

i)Name : technicaladda

region:mumbai

ii)configure options

a)versioning

b)logging (who can log to bucket)

c)tags (add key,value)

iii)set set permissions

iv)review

then you can upload your files. here you can set permissions,and properties

->by default bucket is a private.

->every bucket and object in bucket has overview, properties, permissions, management is their.

->bucket properties has versionong, logging, website hosting, encryption, transfer accelaration, events, tags,

->object properties has encryption, storage type, tags

S3 - Versioning Lab:

--------------------

consloe->services->storage->s3->bucketname->properties->versioning

->using this property we can mainatain the versions of the objects.

->stores all versions of an object.

->generate backup tool.

->once enabled versioning cannot be diabled, only suspended

->integrates with life cycle rules.

S3 Transfer acceleration:

-------------------------

s3 transfer acceleration utilises the cloud front edge n/w to accelarate your uploads to s3. instead of directly uploading to your s3 bucket,

you can use a distinct url to upload directly to an edge location which will then transfer that file to s3.

->s3->bucket->properties->transfer acceleration

we will get a distint url

Q)what is cdn?

cdn means content delivery n/w is a system of distributed servers. that deliver webpages and other web content to a user based on the geographic locations of the user.

here different terminology

i)edge location: cache location seperate to aws regions.

ii)origin : where your content stored.

iii)distribution : colection of edge locations to distribute data.

Learn briefly about CDN in cloud front topic

S3 static website hosting:

-------------------------

static website means which will not change with time.

we can create a static website using this property in s3.

->s3->bucket->properties->static website hosting

use this bucket we can host a website

index.html

------------->make the files are public

error.html

URL: <bucketname>.s3-website.<region>.amazon.aws.com

s3 - security & encryption:

---------------------------

->by default all newly created buckets are private.

->you can setup access control list to your bucket using

i)bucket policies

ii)access control list

->s3 buckets can be configured to create access logs which log all requests made to the s3 bucket. this can be done to another

encryption:

-----------

in transit: when you uploaded the from your local machine to aws data is encrypted (ssl/tls)

----------

at rest:

--------

i)server side

a)s3 managed

keys-SSE-S3

b)aws key management service

keys-SSE-KMS

c)server side encryption with customer provided

keys-SSE-C

ii)client side:

you can encrypt the data and upload to s3

Cross-Region replication:

------------------------

if you want move bucket data from one region to another region use the replication method.

eg:bucket1 (mumbai)

bucket2 (London)

from mumbai to london

source to destinaion

->s3->bucket->management->cross-region replication

->you need to enable the versioning of the bucket.

->under bucket management replication is their

->here objects are copied not objects permissions.

->versioning must be enabled on both sides source and destination buckets.

->regions must be unique.

->files in an existing bucket are not not replicated automatically, all subsequent updated files will be replicated automatically.

->you can not replicate to multiple buckets.

->delete objects are replicated.

->delete individual versions or delete markers will not be replicated.

->understand what cross region replication is at high level.

->you can copy the content of the bucket through aws command line.

->create a user with full s3 access.

in command line

aws configure

aws accesskey id:

aws secret accesskey :

default region name:

default-output format:

let's check bucket list

aws s3 ls

->to copy of one bucket to another bucket

aws s3 cp --recursive s3://technicaladda77 s3://londontechnicaladda

Life cycle management:

---------------------

life cycle means we can move an object from place to another such as standard to IA to Glacier then expire

->s3->bucket->management->lifecycle

->enable the life cycle.

->can be applied to current versions and previous versions.

->following actions can be done

a)transition to the standard - IA storage class

30 days after creation don

b)archive to the glacier storage class

30 days after IA

c)permanently delete

Storage Gateways:

----------------

->aws storage gateway is a service that connects an on-premises software appliance with cloud based storage to provide seamless and secure integration b/w an organizations. this service enables you to securely store data to the aws cloud for scalable and cost-effective storage.

->aws storage gateway's software appliance is available for download as a vm that you install on a host in your datacenter. storage gateway supports either vmware esxi or microsoft hyper-v. once you have installed in your gateway and associated it with your aws account through the activation process, you can use the aws management console to create the storage gateway option that is right for you.

gateway types:

--------------

i)file gateway (NFS)

text based, videos, picture

ii)volume gateway

block storage,vm,os

a)stored volumes (entire blocked)

b)cached volumes (recently activity store)

iii)virtual tape gateway (VTL)

for archival

Note: we can attach these gateways as iscsi devices from your on-premises application server.

i)file gateway:

---------------

files are stored as objects in your s3 buckets. ownership, permissions, and time stamps are durably stored in s3 in the user meta data of the objects associated with the file. once objects are transfered to s3 they can be managed by s3.

ii)volume gateway:

-----------------

to store block based data use this gateway.

a)stored vloumes:

-----------------

we can store entire os in aws s3 in the form ebs. so we can use this gateway we can store to aws

b)cached volumes:

-----------------

store frequently accessed data. it is cached here.

iii)virtual tape gateway:

-------------------------

to upload archival dat to to aws.