CS 79A **Date: 04/21/21**

Class Project

For CS 79A class project, I would like to discuss about: Amazon S3 and Amazon EC2 services

AWS Service: Amazon S3

Amazon S3: Simple Storage Service

1. Legacy storage service. One of the first service offered in 2000s
2. Provides access to reliable and inexpensive(very cheap) data storage infrastructure
3. You can store your personal data, company data, store backups of EC2 instance(called snapshots), store applications, used for s/w dev in cloud, backup configuration files for your amazon web services
4. Amazon EC2 uses Amazon S3 to store EBS(Elastic Block Store) snapshots and instance store-backed AMIs (Amazon Machine Images)
5. Amazon S3 is available for anyone: myself, store lots of pictures, companies, developers, personal storage, etc
6. Amazon S3 can store and retrieve any amount of data from the web. S3 is highly durable, scalable, secure storage.
7. We can accomplish these tasks using AWS Management Console.

1. S3 Use Cases:

a. **Backup and storage:** Provide data backup and storage services

b. **Application hosting:** Provide services that deploy, install, and manage web applications.

c. **Media hosting:** Build a redundant, scalable, and highly available infrastructure that hosts video, photo, or music uploads and downloads.

d. **Software delivery:** Host your software applications that customers can download.

2. Charges for S3 service:

S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data

	First 50 TB / Month	$0.026 per GB
	Next 450 TB / Month	$0.025 per GB

Over 500 TB / Month	$0.024 per GB

S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns

Frequent Access Tier, First 50 TB / Month	$0.026 per GB
Frequent Access Tier, Next 450 TB / Month	$0.025 per GB
Frequent Access Tier, Over 500 TB / Month	$0.024 per GB
Infrequent Access Tier, All Storage / Month	$0.019 per GB
Archive Access Tier, All Storage / Month	$0.005 per GB
Deep Archive Access Tier, All Storage / Month	$0.002 per GB
Monitoring and Automation, All Storage / Month	$0.0025 per 1,000 objects

S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access

All Storage / Month	$0.019 per GB

S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access

All Storage / Month	$0.0152 per GB

S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours

All Storage / Month	$0.005 per GB

S3 Glacier Deep Archive ** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours

All Storage / Month	$0.002 per GB

3. **Is it part of Free-tier: No**

4. **URLs to documentation of the service:**
   a. https://docs.aws.amazon.com/whitepapers/latest/aws-storage-services-overview/aws-storage-services-overview.pdf#welcome
   b. https://docs.aws.amazon.com/AmazonS3/latest/userguide/S3-gsg-CommonUseScenarios.html
   c. https://aws.amazon.com/s3/pricing/
   d. https://docs.aws.amazon.com/whitepapers/latest/aws-storage-services-overview/aws-storage-services-overview.pdf#welcome

5. **Provide URLs to AWS videos about this service:**

   Instead of video I went through the documentation:
   https://docs.aws.amazon.com/whitepapers/latest/aws-storage-services-overview/aws-storage-services-overview.pdf#welcome

6. **Watch the videos you located. For one of the videos you listed, answer the following questions:**
   Instead of video I went through the documentation:
   https://docs.aws.amazon.com/whitepapers/latest/aws-storage-services-overview/aws-storage-services-overview.pdf#welcome

Summarize a few key points made in the video.

1. S3 is cloud storage
2. We should know about bucket when working with S3.
3. Initially we create containers i.e called buckets. You can have as many containers as you want.

4. S3 treats files as Objects. Object contains Data and Metadata. Metadata is file type, size, date creation, etc. It's the name-value pair that describes the Object.
5. There is a key-value. Key is a unique identifier for the Object in a bucket.
6. AWS has an encryption management service. You can also use your own encryption key. You can encrypt before or during.
7. AWS supports versioning Control. We are paying for each version :-)
8. S3 is highly available with 99.99%(four 9's) availability
9. S3 is durability with 11 9(11 9's). Multiple copies in multiple regions, data centers
10. S3 is highly scalability. Limitless storage
11. File size is 0-5 TB. If 15TB file, they split into 3 parts and store in each separate buckets.
12. S3 bucket should be a unique name globally, not just in account.
13. User permissions: public/private/specific users read/write, etc.
14. AWS implemented multiple security measures.
15. Bucket ACLs
16. Static website in S3

S3 OBJECT:

1. S3 treats files as Objects
2. Key is the name of Object/file
3. Value is data in file i.e data in file, bytes
4. Metadata - version, file size, permissions, upload time, etc
5. Version ID - S3 offers version control

S3 Storage Classes/TIERS:

S3 Standard - 99.999(4 x 9s) Availability and 11 x 9s Durability, Redundantly store data in different data centers, low latency. Multiple AZs in a Region.

S3 Reduced Redundancy(RRS) - non-critical, reproducible data with less redundancy than S3 standard. For durability, RRS objects have an average annual expected loss of 0.01 percent of objects. If an RRS object is lost, when requests are made to that object, Amazon S3 returns a 405 error. **Note: Do not use RRS. Instead use S3 Standard**

S3 IA Infrequently Access - ideal for less frequent but rapid when needed. Typically used for long term storage backup. Low latency and throughput performance. 99.9 Availability(3 x 9s). Data is resilient

S3 One Zone-IA - Data is stored in one AZ in a Region. Less accessed data, 20% less cost

S3 Intelligent Tiering - helps to manage. Allows to optimize your data. For archiving, long-term storage, Glacier, Cost saving

Identify two interesting quotes that were made.

S3 Standard - 99.999(4 x 9s) Availability and 11 x 9s Durability, Redundantly store data in different data centers, low latency. Multiple AZs in a Region.

S3 IA Infrequently Access - ideal for less frequent but rapid when needed. Typically used for long term storage backup. Low latency and throughput performance. 99.9 Availability(3 x 9s). Data is resilient

What new facts did you learn from watching this video.

1. S3 is highly available with 99.99%(four 9's) availability
2. S3 is durability with 11 9(11 9's). Multiple copies in multiple regions, data centers
3. S3 is highly scalability. Limitless storage
4. File size is 0-5 TB. If 15TB file, they split into 3 parts and store in each separate buckets.
5. S3 bucket should be unique name globally, not just in account.
6. User permissions: public/private/specific users read/write, etc.
7. AWS implemented multiple security measures.
8. Bucket ACLs
9. Static website in S3
10. S3 treats files as Objects
11. Key is the name of Object/file
12. Value is data in file i.e data in file, bytes
13. Metadata - version, file size, permissions, upload time, etc
14. Version ID - S3 offers version control

What was the best part of the video? Why?

All the documentation provided by this whitepaper is informational

S3 Glacier Deep Archive:

1. It takes time to pull the data.
2. Cost: 0.00099/GB/Month (< ¼ cost of S3 Glacier)
3. Cheaper than backup magnetic tapes
4. 12 hours data retrieval

The following table summarizes the archive retrieval options.

Storage class	Expedited	Standard	Bulk
Amazon S3 Glacier	1–5 minutes	3–5 hours	5–12 hours
S3 Glacier Deep Archive	Not available	Within 12 hours	Within 48 hours

What questions remain in your mind after watching the video? Why? **None**

AWS Service: Amazon EC2

What is EC2?

1. EC2 (Amazon Elastic Compute Cloud) provides computing services in Amazon Infrastructure.
2. EC2 is a server in Cloud. We need not buy any infrastructure, no upfront cost, and save a lot for money for companies including startups. They are VMs or called instances and we can configure the security, network, volume/partitioning.
3. With EC2 instances you need not take care of wiring cabling, setup, etc. AWS takes care of it.
4. A.EC2 can scale up/down depending on how busy the season in a year is. We can easily shut down/terminate instances, startup instances as needed.

1. EC2 Use Cases:

Hosting environments:

EC2 is used for hosting a variety of applications, software and websites on the Cloud.

Development and test environments:

The scalable nature of EC2 means that organizations now have the ability to create and deploy large scale testing and development environments with unprecedented ease.

Backup and disaster recovery:

Companies are leveraging EC2 as a medium for performing disaster recovery for both active and passive environments.

Banking and financial sector:

These are areas that demand the utmost in security and scalability — both are factors that Amazon EC2 provides in droves.

Marketing and advertisement:

Low costs and rapid provisioning capabilities that EC2 allows have resulted in the platform increasingly being used to host marketing and advertising environments on the fly.

High performance computing:

The need for HPC is exponentially on the rise, and EC2 provides specialized virtualized servers that provide both high performance networking and compute power.

2. Charges for EC2 service:

Pricing:

1. There is a free tier once you sign up for AWS.
2. AWS provides following pricing options:
    a. **On-Demand:** Once you start configuring EC2 instances, you are by default with On-Demand service. Pay as you go service
    b. **Reserved Instances:** One or 3 year term. It's cheaper but if you don't use it, you are still paying the money. It's a low, one-time payment.
    c. **Spot instances:** Requesting for unused EC2 instances. Bidding for it. It's good for people who are not time sensitive, need not do work immediately, researching, etc
    d. **Scheduled Instance pricing:** Using the EC2 instance on a scheduled day and time for a year. It's cheaper than Reserved Instances.
    e. Dedicated Instance Physical Server: Instance(s) are located on physical servers and billing is per instance. This physical server is dedicated to you and you can run 1 instance or 100 instances. https://aws.amazon.com/ec2/dedicated-hosts/
    f. **Dedicated Hosts Physical Server:** The physical server host is dedicated to you, you have more control on Hardware, how instances are placed on server. Billing is per host. You might use such type of host for compliance or security purposes. https://aws.amazon.com/ec2/dedicated-hosts/

3. Is it part of Free-tier: Yes

4. URLs to documentation of the service:

a. https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-ug.pdf#EC2_GetStarted
b. https://aws.amazon.com/ec2/pricing/
c. https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-instance-lifecycle.html
d. https://www.youtube.com/watch?v=b42C3uDNh_c

5. Provide URLs to AWS videos about this service:

Instead of video I went through the documentation:
https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-ug.pdf#EC2_GetStarted

6. Watch the videos you located. For one of the videos you listed, answer the following questions:

Instead of video I went through the documentation:
https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-ug.pdf#EC2_GetStarted

Summarize a few key points made in the video.

EC2 features:

1. Virtual machines or servers are termed as instances in AWS.
2. Preconfigured templates for your instances are called Amazon Machine Images (AMIs), that consists of preconfigured operating system, software, etc.
3. WE need to consider configurations of CPU, memory, storage, networking capacity for the instances
4. Securely logging into Windows & LInux instances using Public and Private keys (saves the private key in a secure place)
5. For Windows authentication into Win instance: we have username as “administrator’ and password, we get decrypted off their website via key
6. For Linux authentication into Linux instances: we have public and private key to authentication.
7. EC2 has multiple EBS instances
8. EBS STORAGE: It is temporary storage and on stop and reboot or restart or reboot the data is lost and it’s also lost when instance is terminated.
9. Elastic Block Store (EBS) is attached to EC2 instances. It requires an operating system for a few different types, magnetic standard, fast IOPs, there are 5 different options for elastic block storage
10. When you configure your instance, you can configure near the customer’s Region and AZ so there will be less performance issues, less latency lag.
11. If yours is a global company, you have web servers all over the world and set up DNS (Domain Name System) and route to local instances of that AZ or Region.
12. When setting up EC2 instance, we have the option of setting up IP version 4 and IP version 6 configuration. During EC2 instance configuration, you can set a static IP address, a non changing IP address and it’s referred to as Elastic IP.
13. When setting up an EC2 instance, you set an IP address that will not be the same always. It can change. When we get an Elastic IP address, it will not change. It remains the same but they are not free.
14. When configuring EC2 instance, we need to configure in Virtual Private Cloud i.e selecting network and subnet you want instance running. We can use default VPN if you haven’t created any Public or Private Virtual Private Cloud.

Identify two interesting quotes that were made.

Amazon Machine Images (AMI):

1. AMIs are Pre-configured templates with operating systems, application servers, applications, etc.
2. One AMI can launch 1 or more instances of different processor types such as t2.micro, D2, M4, etc

3. Once an instance is launched, it behaves like a normal computer. You have complete control of the instance. You can use command sudo that give root privileges
4. There is a limit on no: of instances you are running. You might need additional permission from Amazon if you need to run more instances.

What new facts did you learn from watching this video.

Storage for your Instance:

1. The root device of your instance is where Amazon Machine Image is installed. In Windows, it's C:\ drive and on LInux, it's "/sda". This is where the operating system(or Root Volume) of the instance is installed.
2. **The Root Volume (OS) can be in Instance Store Backed or Elastic Block Storage (EBS)**
3. **Instance Store Backed:** The root volume is in physical drive connected to the host instance. Driver is connected to the host directly and I/O performance has low latency. It's Ephemeral (non-persistent) i.e data will be lost when instance is terminated.
4. **EBS:** The Root Volume is in EBS along with other volumes. Data gets loaded back and forth, we lose a little bit of performance but gain persistence of data. The volume can persist even after instance is terminated. Once disconnected, it can be connected to new instances.
5. EBS can be stopped and make configuration changes.
6. Instance Store Backed cannot be stopped. Once terminated, it's deleted.

What was the best part of the video? Why?

All the information about how we can use EC2 instance, Storage, AMIs, Security Groups configuration, etc

What questions remain in your mind after watching the video? Why? None