CLOUD COMPUTING

Practical 2: Overview of AWS EC2 and S3

School of Computer Science
University College Dublin

Description
In this practical, we will explore a cloud computing environment though AWS EC2 and S3. This will bring you through AWS EC2 and S3 VM to explore some of the cloud computing concepts.

Part I: AWS EC2 **(35 marks)**

1. What is Amazon EC2?

Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the AWS cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy application faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage.

2. List main components of AWS. Some key components of AWS:
   - Simple Email Service: It allows you to send emails with the help of regular SMTP or by using a restful API call.
   - Route 53: It's a DNS web service.
   - Simple Storage Device S3: It is a widely used storage device service in AWS Identity and Access Management.
   - Elastic Computing Cloud (EC2): It acts as an on-demand computing resource for hosting applications. EC2 is very helpful in time of uncertain workloads.
   - Elastic Block Store (EBS): It allows you to store constant volumes of data which is integrated with EC2 and enable you to data persist.
   - Cloud watch: It allows you to watch the critical areas of the AWS with which you can even set a reminder for troubleshooting.

3. List main features of Amazon EC2.
   - Virtual computing environment, known as Instance.
   - Preconfigured templates for your instances, known as Amazon Machine Images (AMISs) that includes all you need for the server (Operating system and additional hardware)
   - Various configuration of CPU, memory, storage, and networking capacity for you instances, known as instance types.
   - Secure login information for your instances using key pair (public key stored at AWS and private key stored at your computer).
   - Storage volumes for temporary data that's deleted when you stop or terminate your instance, known as instance store volumes.
   - Persistent storage volumes for your data using Amazon Elastic Block Store (Amazon EBS), known as Amazon EBS volumes.
   - Multiple physical locations for your resources, such as instances and Amazon EBS volumes, known as Regions and Availability Zones.

- A firewall that enables you to specify the protocols, ports, and source IP ranges that can reach your instances using security groups.
- Static IPv4 addresses for dynamic cloud computing, known as Elastic IP addresses.
- Metadata, known as tags, that you can create and assign to your Amazon EC2 resources.
- Virtual networks you can create that are logically isolated from the rest of the AWS cloud, and that you can optionally connect to your own network, known as Virtual Private Clouds (VPCs).

4. What are **region** and **Availability Zones** in AWS Cloud?
   - Each Amazon EC2 Region is designed to be completely isolated from the other Amazon EC2 Regions.
   - Regions are separated into geographical areas, each region has multiple, isolated locations known as Availability Zones (AZ). Each AZ is represented by a Region code followed by a letter identifier; for example us-east-1a.

5. Discuss some benefits of **Availability Zones**?
   - Amazon EC2 provides you the ability to place resources, such as instances, and data in multiple locations. Failures can occur that affect the availability of instances that are in the same locations. If you host all your instances in a single location that is affected by such a failure, none of your instances would be available.

6. What is Amazon Machine Image (AMI)?

AMI is template that contains a software configuration (e.g., an Operation System, an application server, and applications).

7. What is Instance in AWS?

From an AMI, an instance can be launched, which is a copy of the AMI running as a virtual server in the cloud. User can launch multiple instances of an AMI.

8. How is relation between Instance and AMI?

Using a single AMI, you can download as many instances as you can. An instance type is used to define the hardware of the host computer for your situation. Each instance is unique and provides the facilities in computational and storage capabilities. Once you install an instance, it looks similar to a traditional host with which we can interact in the same way we do with a computer.

9. What is Key pair used in AWS?

Amazon EC2 uses both public and private keys to encrypt and decrypt the login information. The sender uses a public key to encrypt the data and the receiver uses a private key to decrypt the data. Private and public keys are known as key pairs. The public key enables you to access the instance securely and a private key is used instead of a password.

10. Cloud computing model for EC2?

It is IaaS.

Part II: AWS S3 **(35 marks)**

Using online resources, explore the AWS S3 (through registering an AWS Educate account).

Questions

1. What is Amazon AWS S3?

Amazon Simple Storage Service (S3) = storage service over the Internet.

2. What can you do with Amazon S3?

User can use Amazon S3 to store and retrieve any amount of data at any time, from anywhere on the web.

3. What is bucket in Amazon S3?

Amazon S3 stores data as objects within buckets. An object consists of a file and optionally any metadata that describes that file. To store an object in Amazon S3, user upload the file to a bucket.

Buckets are the container for objects. For each bucket, user can control access to it (who can create, delete, and list objects in the bucket), view access logs for it and its objects, and choose the geographical region where Amazon S3 will store that bucket and its contents.

4. What is maximum size of S3 bucket?

Range from 1 byte to 5 terabytes.

5. What is the maximum number of S3 buckets you can create?

By default = 100 for each AWS account. If you need additional buckets, you can increase your account bucket limit to maximum 1000.

6. How to upload a file that is greater than 100 MB in Amazon S3? Using Multipart Upload.

7. How is Amazon S3 data organized?

Data is stored as objects within resources called “buckets”, and a single object can be up to 5TB in sizes. Objects are organized with shared names called prefixes. You can also append up to 10 key-pairs called S3 object tags to each object, which can be created, updated and deleted throughout an object’s lifecycle. To keep track of objects and their respective tags, buckets, and prefixes, you can use an S3 Inventory report that lists your stored objects within an S3 bucket or with a specific prefix, and their respective metadata and encryption status.

8. What is S3 versioning?

Versioning is a means of keeping multiple variants of an object in the same bucket.

9. What are benefits of using versioning in S3?

You can use versioning to preserve, retrieve, and restore every version of every object stored in your Amazon S3 bucket. With versioning, you can easily recover from both unintended user actions and application failures.

10. Cloud delivery model for S3?

It is Storage as a Service (STaaS). STaaS is a cloud services model in which a company leases or rents its storage infrastructure to another company or individuals to store either files or objects. Small companies and individuals often find this to be a convenient methodology for managing backups, and providing cost savings in personnel, hardware and physical space.

In general, STaaS operates through a web-based API that is remotely implemented through its interaction with the client application's in-house cloud storage infrastructure for input/output (I/O) and read/write (R/W) operations. If the company ever loses its local data, the network administrator could contact the STaaS provider and request a copy of the data.

For enterprise-level cloud storage, Amazon S3 is a popular-provider.

Part III **(30 marks)**

1. IaaS is an important requirement.

2. Discuss the suitability of AWS AC2 for this company.

Required feature	Satisfied by EC2	Comments
Ability to scale	Yes	The autoscaling offering of AWS EC2s can be leveraged to scale instances up or down based on utilization
Security	Yes	AWS EC2 is an IaaS solution that gives users the ability to configure every aspect of the platform. This means that the application team can take full control of the security configurations of the EC2 instances to be able to apply all required safeguards to enable data security
Governance compliance wrt. retaining data within geographical boundaries	Yes	With EC2 you can choose the region within which the instances are launched. This feature can be leveraged to ensure that data is processed within the geographical boundaries as prescribed by the law. For EU data can be processed in EC2 instances in the Europe region to ensure data does not leave the borders of the EU

3. Discuss the suitability of AWS S3 for this company.

Required feature	Satisfied by S3	Comments
Ability to scale	Yes	There is no restriction on the number of objects that can be stored in an S3 bucket. Hence scalability of data is not an issue with S3
Security	Yes	S3 offers strong data encryption mechanisms supporting both client side and server-side data encryption using KMS keys. This can be leveraged to securely store confidential data on the cloud

Governance compliance wrt. retaining data within geographical boundaries	Yes	AWS buckets can be created in specific regions and access can be controlled at each bucket level. This can be leveraged to maintain data in specific locations geographically and govern access to the same.
Availability	Yes	S3 promises an availability of 11 9's which means that availability aspect of data has been taken care of by the platform itself
Geo-proximity of data	Yes	The feature of being able to create buckets in specific regions can be leveraged for this purpose too