

AWS Certified Cloud Practitioner

Gi Wah Dávalos Loo

2020



1 Cloud Computing

1.1 What is

Cloud computing is the **on-demand delivery** of it resources (database storage, compute power, etc), which you can access and **provision the right type and size of computing almost instantly** and have a **pay-as-you-go** pricing.

1.2 Deployment Models of the Cloud

1. Private:

- Complete control, meet specific business needs
- Used by single organization
- Cloud owned by third-party

2. Public:

- Cloud resources owned and operated by third-party.

3. Hybrid:

- Keep the control of some servers and extend some capabilities to the Cloud

1.3 Five Characteristics of Cloud Computing (Amazon POV)

- **On-demand self service:** Users can provide it resources without human interaction.
- **Broad network access:** Can access the AWS panel from the internet.
- **Multi-tenancy and resource pooling:** Multiple users share the same physical resources yet their it resources are isolated with security and privacy.
- **Rapid elasticity and scalability:** Automatically and quickly add or remove it resources.
- **Measured service:** Usage is measured and you pay what you use.

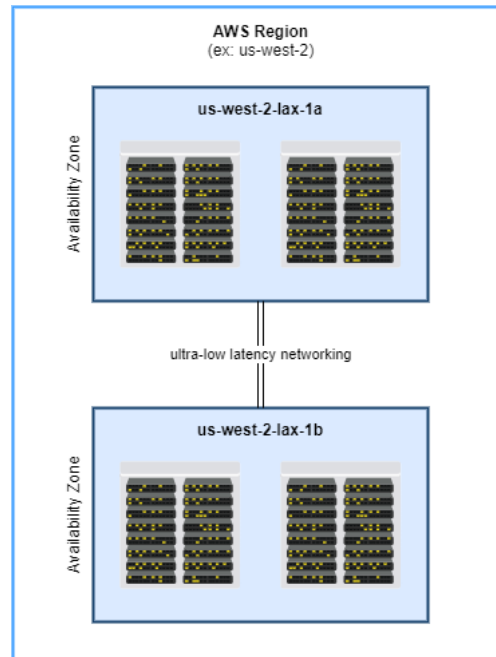
1.4 Advantages of Cloud Computing

- **Trade CAPEX for OPEX:** Pay on-demand don't own hardware.
- **Benefit from economies of scale:** If more people use AWS, AWS will acquire more hardware and become more efficient, then the pricing goes lower.
- **Use what you need:** Scale based on actual usage measure.
- **Increase speed and agility:** Instance resources almost immediately.
- **Globality:** Instance IT resources in many geographical locations.

1.5 Problems solved by the Cloud (Business with IT needs POV)

- **Flexibility:** Add, remove and change IT resources when needed.
- **Cost-effectiveness:** Pay as-you-go.
- **Scalability:** Increase IT resources when receiving larger loads.
- **Elasticity:** Scale-in and scale-out when needed.
- **High-availability and 'Fault-tolerant':** You gotta trust AWS.
- **Agility:** Quick development process, test and launch software applications.

1.6 AWS Global Infrastructure



2 IAM

2.1 What is

IAM (**I**dentify and **A**ccess **M**anagement) is a **global** AWS service that helps you manage the permissions of the access to AWS services and resources.

2.2 Users, Groups, Roles, Policies

2.2.1 Users

People within the organization. One AWS user account per physical user. You can assign policies to a user.

2.2.2 Groups

Can contain users only (not other groups). You can assign policies to a group.

2.2.3 Roles

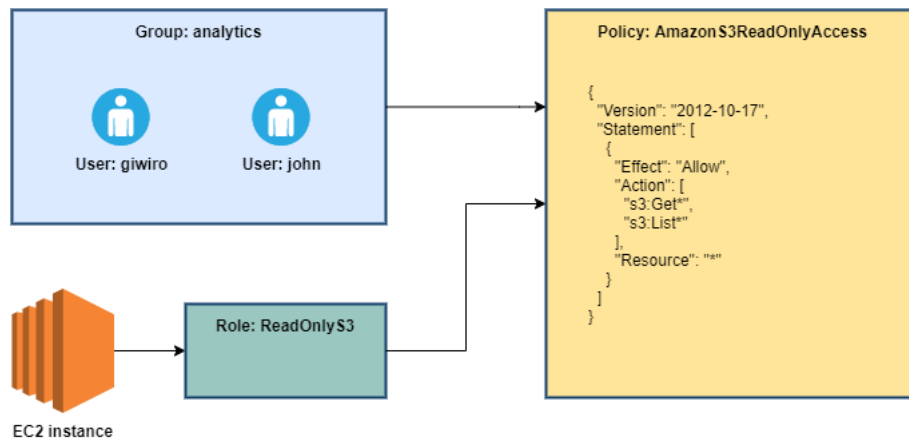
In order to bind a policy to an AWS service or resource (EC2, Lambda, etc), we need to it through a role.

2.2.4 Policies

JSON document that defines the access level of AWS services and resources. Example:

```
// AmazonEC2ReadOnlyAccess
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:Describe*",
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": "elasticloadbalancing:Describe*",
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "cloudwatch:ListMetrics",
        "cloudwatch:GetMetricStatistics",
        "cloudwatch:Describe*"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": "autoscaling:Describe*",
      "Resource": "*"
    }
  ]
}
```

2.3 IAM Elements Relationships



2.4 Access AWS

- **Management Console:** Web portal. Protected by password + MFA (Optional, but recommended)
- **CLI:** Command Line tool. Protected by access keys.
- **SDK:** Code library. Protected by access keys.

2.5 IAM Security Tools

- **IAM Credentials Report:** Account level tool that lists all users and the status of their credentials.
- **IAM Access Advisor:** User level tool. List of all permissions (policies) per service and when it was last accessed.

2.6 IAM Share responsibility model

- Management and monitoring of users, groups, roles and policies
- Enable MFA on all accounts
- Rotate all the keys often
- Analyze access patterns & review permissions

3 EC2

3.1 What is

EC2 (**Elastic Compute Cloud**) is a IaaS service which spawns an instance of selected OS. It can be configured with a custom script at start in the **User Data** option in Step 3 of the creation of an EC2.

3.2 Security Groups

Firewall security rules that control how traffic is allowed in and out the EC2 instance. By default all ports are closed, except for those ports allowed in the security groups.

3.3 Connecting to EC2

- **SSH:** Using a ssh client, the public EC2 ip and the private key from the key pair creation.
- **EC2 Instance Connect:** Web platform that creates a disposable key-pair for access to the EC2. If you EC2 ssh ports are not allowed, then you won't be able to connect through Instance Connect.

3.4 EC2 Instance Purchase options

- **On Demand:** Billing per second, **highest cost** but no long commitment or upfront payment.
- **Reserved: 75% discount** but has reservation period commitment of 1 or 3 years (more years more discount).
 - **Convertible:** Every characteristic of the EC2 (family, instance type, platform, etc) can be changed.
 - **Standard:** Some attributes such as size, can be changed but only on the existing instance (family, instance type or platform, for instance, cannot be changed).
- **Scheduled:** Reservation that recur on a daily, weekly or monthly basis with **fixed start time and duration** for **one year term**.
- **Spot Instances: 90% discount** but **it can be terminated at any time** (if your max price is bigger than the spotted price)
- **Dedicated Hosts:** Physical server with EC2 instance capacity dedicated to your use. Helps to address **compliance requirements** and allow to **use your existing server-bound software license**. Allocation is a **3 year commitment** and is more expensive.
- **Dedicated Instances:** Instances running on **hardware dedicated to you**.

3.5 EC2 Shared responsibility model

- Manage Security Groups rules
- OS patches & updates
- Software installed in the EC2
- IAM Roles assigned to the EC2
- Data security on your instance

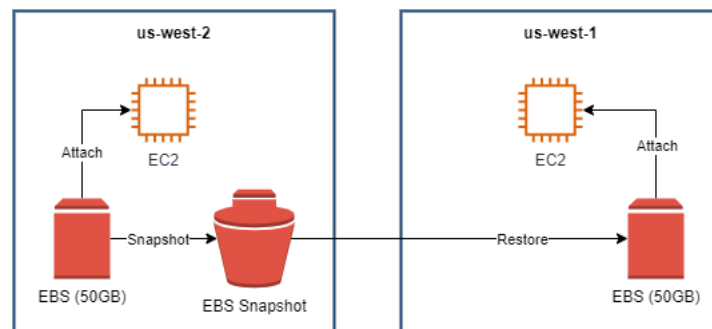
4 EBS

4.1 What is

EBS (Elastic Block Store) Volume is a network drive that can be attached to a EC2 while they run. EBS can only be **assigned to one EC2 at a time**, they it is bound to a **specific availability zone** and it has a provisioned capacity in GBs and IOPS.

4.2 EBS Snapshots

Snapshots are **backups a EBS Volume at a point in time**, and can be copied across regions/AZ. Since you cannot move an EBS across regions/AZ, you can create a Snapshot in the region, and then create a new EBS from the snapshot in the desired region/AZ.



4.3 AMI

It stands for Amazon Machine Image and are a **customization** of an EC2 instance (With additional software, config, OS, etc). they are built for a **specific region** but can be copied across regions. You can launch EC2 instances from:

- Public AMI (AWS provided)
- Your own AMI (maintained by you development team)
- AWS Marketplace AMI (Someone else's AMI)

4.4 EC2 Instance Store

Ephemeral storage (Data is lost upon restart) that is located in the same physical location of the EC2 instance and provides a **better I/O performance**. It's good for buffering and caching but it's **more expensive**.

4.5 EFS

Managed NFS (network file system) that **can be bond to multiple EC2 across multiple AZ**. It is **HA, scalable and very expensive**.

4.6 EC2 Storage Shared responsibility model

- Setting backup / snapshot procedures
- Setting up data encryption
- Responsibility of any data on the volumes
- Understand the risk of using EC2 Instance Store

5 ELB & ASG

5.1 Scalability

Ability to accommodate a larger load by making the hardware stronger (vertical) or by adding nodes (horizontal).

5.1.1 Vertical scaling

By adding more hardware capacity, the instances can handle greater load of requests. For example, change from a *t2.micro* to a *t2.large*. It is very common for non distributed systems such as databases. You **scale up** when the hardware capacity increases, and **scale down** when decreases.

5.1.2 Horizontal scaling

By adding the number of instances you can handle greater load of requests, and it implies using distributed systems, such as modern applications. You **scale out** when the number of instances increases, and **scale in** when decreases.

5.2 High Availability

Instances are across multiple Availability Zones (AZ). Using multi AZ ELB and ASG. (Usually related to horizontal scaling)

5.3 Elasticity

Once a system is already scalable, elasticity means that you can dynamic and automatically scale based on the load.

5.4 Agility

(NOT RELATED, MEANT TO DISTRACT) Cloud capacity to provide IT resources without human interaction (one click away).

5.5 ELB

Elastic Load Balancer (ELB) are servers that forward internet traffic to multiple EC2 instances downstream.

5.5.1 Why?

- Spread load across multiple downstream instances
- Expose single point of access to your application
- Seamlessly handle failures of downstream instances
- Do regular health check on instances
- SSL termination
- Is a managed load balancer (AWS oversees)

5.5.2 Kinds of load balancers

- Application Load Balancer - Layer 7 (HTTP Only)
- Network Load Balancer - Layer 4 (Raw TCP)

5.6 ASG

Auto Scaling Group (ASG) are a collection of EC2's that ensure to have a minimum and a maximum number of machines running. It automatically register new instances to a load balancer. When it need more instances (configured logic based on hardware statistics) it **scales out**. It also replace unhealthy instances.

